Empirical Approaches to

CONSUMER PROTECTION ECONOMICS

PAULINE M. IPPOLITO
DAVID T. SCHEFFMAN
Editors

Proceedings of a Conference
Sponsored by the
BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
April 26-27, 1984

FEDERAL TRADE COMMISSION

ACKNOWLEDGEMENTS

As with any conference of this type, there are many individuals at the FTC who contributed to its overall success. In particular, though, I owe a special debt of gratitude to Ieva P. Cucinelli, who cheerfully and painstakingly oversaw the conference manuscripts through revision, editing and typesetting; Kevin P. Leonard, who designed and typeset the entire volume despite his otherwise heavy workload and despite his total unfamiliarity with economists' jargon and technical requirements; and Don Cox, for his assistance with the preparation of the graphics contained in the volume.

P.M.I.

FEDERAL TRADE COMMISSION
BUREAU OF ECONOMICS
Consumer Protection Conference

Table of Contents

Consumer Protection Economics: A Selective Survey

*Pauline M. Ippolito**
Federal Trade Commission

The papers in this volume were presented at a conference on Empirical Approaches to Consumer Protection Economics held at the Federal Trade Commission in Washington, D.C. on April 26 and 27, 1984.[1] The conference was sponsored by the FTC's Bureau of Economics; the early stages of conference preparation were under the guidance of David T. Scheffman, then Associate Director for Special Projects; subsequent preparation for the conference and the production of this volume were handled by this Editor.

I. Introduction

Consumer protection regulation is one of the primary activities of the Federal Trade Commission. Until the early 1970s, the Commission solicited little economic advice on consumer protection policy and little was offered by economists. Even economists employed at the FTC spent virtually no time on consumer protection matters before 1974, in contrast to the substantial commitment to competition case work and research. The first small budgetary allocation for economic analysis of consumer protection policy within the FTC was made in the mid-1970s; a division of economists to support this function was established in 1978. Before this time, decisions dealing with "unfair or deceptive acts or practices" were apparently judged to be issues that would not benefit from an economic perspective.[2]

In many ways this is not surprising. The economics of consumer protection regulation is essentially contained in the economics of information[3]; twenty years ago there was no "economics of information." The traditional economic view

*I would like to thank Ronald Bond, Gerard Butters, John E. Calfee, Peter Huang, Richard Ippolito and Michael Lynch for helpful comments. The opinions expressed here do not necessarily represent the opinions of the Federal Trade Commission or any individual Commissioner.

[1] In most cases, papers in this volume are revised versions of papers presented at the conference.

[2] The legislative authority for consumer protection activities at the Federal Trade Commission is contained in the 1938 Amendments to Section 5 of the FTC Act. The operative sentence of Section 5 states that "Unfair methods of competition in or affecting commerce and unfair or deceptive acts or practices in or affecting commerce are declared unlawful." The "unfair or deceptive acts or practices" clause is generally regarded as the legal basis for the Commission's consumer protection function. Like its antitrust authority, this legislative language maximizes the agency's flexibility in defining the scope of consumer protection law.

[3] The development of the law and economics literature has also had a substantial influence in the changing economic views of consumer protection regulation. This is especially true in evaluations of regulatory and judicial approaches to perceived information problems. Because of space limitations, I will not review many of the related developments in this area.

of consumer protection issues at the time was one of indifference or outright hostility. The economic consensus of twenty years ago is probably fairly summarized by the position that as long as there are a sufficient number of competitors, consumers receive an optimal mix of goods and services.[4] The issue of contract enforcement and the possibility of fraud would probably have been acknowledged, but not as problems of interest to economists.

This is not to imply that economists did not recognize the importance of information in markets. In the *Wealth of Nations* (1776), for instance, Adam Smith's observation that "the wages of labour vary according to the small or great trust which must be reposed in the workmen" reflected the special economic response required for one type of situation where information about the quality of a service is not available before purchase. However, with very few exceptions,[5] recognition of the issue did not lead to any serious analysis of the implications of imperfect information for individual decision making or for overall market performance. This trivializing treatment of information's role in markets provides an easy explanation for the long-standing indifference of economists towards consumer protection policy.

In the last twenty years the situation has changed dramatically. Today's economics journals are literally filled with articles detailing the many possible effects of costly information.[6] While models based on assumptions of perfect information are certainly appropriate for many — and maybe even for most — economic problems, there is a growing recognition that in some cases information *matters*. Many market arrangements and many government policies are today seen to be fundamentally shaped by the information environments in which they exist. Contracts, liability rules, retailing organizations, advertising, the degree of vertical integration, the nature of investment patterns and industry structure are just a few examples.

These developments have certainly been of interest to those trying to evaluate or to guide consumer protection policy. At institutions like the FTC it is now routine for perceived consumer protection problems to be discussed in terms of possible market solutions, the likely effects of alternative regulatory approaches on consumer and firm behavior, and the associated benefits and costs of these interventions.[7] The theoretical developments of the last twenty years have brought the analytical tools of economics to consumer protection regulation.

Despite these substantial developments, consumer protection regulation remains a topic of great controversy. This is certainly the case in the political and

---

[4] The monopolistic competition view of markets might be the exception here in that excess competition in some activities such as advertising was expected.

[5] Stigler's 1961 paper is certainly an anomaly in the literature of the time as is the earlier paper by Scitovsky (1950) which also focused on the implications of information in consumer markets.

[6] See, for instance, Arrow (1974), Barzel (1977), the symposium described in Spence (1976), and Stiglitz (1979).

[7] For a policy discussion of these issues, see, for instance, the *Policy Review Session on Consumer Information Remedies*, FTC (1979).

regulatory arenas.[8] Here the fundamental debate is often framed as a contest between the "pro-consumer" groups, who essentially argue that consumers do not have access to sufficient information to influence market performance adequately, and the "pro-business" groups, who argue that market discipline is a better arbiter of product quality and consumer choices than government regulation.

The issues are no more settled in the economics profession, and surprisingly, the tension here is fundamentally the same as in the policy arena. As briefly reviewed in the next section, the economics of information literature has developed along two major lines: the first is the identification of the variety of ways in which asymmetric information can affect market performance, and the second is the detailing of the many mechanisms that the market or government can use to reduce these information problems. While this literature certainly supports the view that information asymmetries can be a substantial force in market performance, it is inconclusive on the appropriate role for consumer protection policy. Policy responses to information problems usually vie directly with market responses to these problems. The literature to date has very little guidance to offer to policymakers who ideally seek to implement policy remedies only when they are more efficient than private responses.

In light of this situation and the needs of the FTC in formulating consumer protection policy, this conference was organized with two primary goals:

(1) To highlight the fact that the economics literature currently contains very little empirical research that attempts either to measure the magnitude of information problems in consumer markets or to test alternative theories of the precise nature of these problems and the effectiveness of possible solutions;

(2) To bring together a capable and diverse group of economists with varied interests in policy and the economics of information to explore empirically the effects of current consumer protection policy as well as underlying market behavior.

A selective review of the economics of information literature as it relates to consumer protection policy issues is presented in Section II. The conference papers are briefly reviewed in Section III. As is clear from the broad range of topics covered by these papers, this collection includes some very interesting work. Nevertheless, in many cases the papers here are only tentative first steps in an effort to understand behavior in markets with costly information; and as such, the papers often raise as many questions as they seek to answer. Concluding remarks follow in Section IV.

We hope that the conference and this volume will interest others in the economics of consumer information problems. We especially hope that economic researchers will attempt to develop methods that will lead to better measure-

[8] Averitt (1981), for instance, reviews some of the legal and political controversies surrounding the FTC's unfairness authority. See also Beales, Craswell and Salop (1981) and some of the papers cited there.

ment and understanding of both market discipline and alternative regulatory policies. Policy can be improved only if there is a realistic understanding of the magnitude of the problems themselves and the effectiveness and limitations of alternative remedies in actual use.

II. The Economics of Information in Consumer Markets: A Brief Review

The economics literature relating to consumer protection regulation is large and growing rapidly; it is impossible to review it adequately here.[9] Instead I would like to discuss briefly some of the major developments in this literature and to highlight their relationship to problems in consumer protection policy. The work dealing with information about prices is covered first; the range of research exploring issues created by asymmetric information about product quality is reviewed in the second section.

A. Information About Prices

Lack of information about price is generally taken to be a simpler problem to analyze than information about product quality. Historically, it is this problem that was addressed first. Stigler's 1961 article, which was seminal in the modern literature, observed simply that if consumers did not have costless information about firms' prices, some degree of price dispersion would persist even in otherwise competitively-structured markets. Much work has followed Stigler's paper, refining the modeling of the consumer's search process, more carefully addressing the seller's role in disseminating information and setting prices, and explicitly dealing with the effects of individuals' search cost differences.[10] However, these refinements have not changed the primary result of this line of literature: if information is costly to acquire or to disseminate, prices will be higher than at competitive levels, and price dispersion may persist in the market. The magnitude of these effects depends fundamentally on the cost of informing consumers.

The primary impact of this literature on policy has been to increase the scrutiny paid to regulations and to private devices that restrict the flow of price information. This has been most significant in the area of occupational regulation where restrictions on price advertising are gradually being removed. For instance, in 1978 the FTC passed a trade regulation rule prohibiting states and trade organizations from restricting price advertising for eyeglasses and related services. The published basis for this rule[11] was essentially the conclusion of this literature:

[9] Reviews of various segments of this literature include Hirshleifer (1973), Hirshleifer and Riley (1979), Salop (1978), and Stiglitz (1979).

[10] See, for instance, Butters (1977), Rothschild (1973), Salop (1976) and Stiglitz (1979).

[11] See Statement of Basis and Purpose and Final Trade Regulation Rule for the Advertising of Ophthalmic Goods and Services, FTC, 16 CFR Part 456, Federal Register, June 2, 1978, Volume

(footnote cont'd)

that restricted price advertising increased the consumer's cost of acquiring price information and thus increased the average price and the dispersion of prices in the market. Removal of regulatory prohibitions on advertising by lawyers, by dentists and other medical professionals, and by drug stores are other recent examples of this movement.

This focus on the costliness of price information and its role in setting market prices has also had an effect on antitrust policy.[12] It has contributed to the general diminution of the structure-conduct-performance view of markets that had long been a foundation of antitrust policy. If in some markets information about prices is more efficiently conveyed by large sellers than small sellers (through advertising or reputation, for instance), then a growth in concentration could be beneficial to consumers. This would explain the often-observed reality that markets characterized by many small sellers frequently seem to perform rather poorly, while markets with just a few sellers seem very competitive (see Stiglitz (1979), for instance). One of the primary changes in the eyeglass market since the removal of restrictions on price advertising, for instance, has been the substantial growth of chain retailers — who advertise regularly and who appear to have been a substantial force in bringing prices down.[13]

While these developments in the literature have been helpful in understanding regulatory policy, several price information issues have not been addressed. A basic premise of the literature to date is that price is immediately and objectively verifiable upon inspection. In this sense, these price information analyses are equally applicable to all quality characteristics of goods which are similarly verifiable — "inspection characteristics" in the Nelson (1970) terminology. On the other hand, this line of literature is not applicable to more complex price information. I would like to mention two such situations that have arisen in regulatory issues at the FTC: prices of multi-product sellers and "uncertain" prices.

Multi-product sellers, such as grocery stores, provide the efficiency of collecting a wide variety of goods in one location. This localization of so many goods with one seller and the consumer's cost of moving between sellers make price competition in these markets fundamentally different from that analyzed in the literature to date. The magnitude of the computational task and the stochastic nature of purchases from such sellers makes direct price comparisons difficult.

43, No. 107. This portion of the rule was eventually remanded to the Commission by the review courts for evidentiary reasons; however, many state legislatures had enacted the essence of the rule in the meantime.

[12] The usual distinction made between "consumer protection problems" and "antitrust problems" is that the former involve consumer information deficiencies while the latter involve problems that would be eliminated by sufficient competition. The issue of price advertising strains this already-strained distinction, as does any consumer information issue that has implications for the structure of markets or the institutional nature of competition in them.

[13] In some states other explicit or implicit restrictions on chains were also removed, making an assessment of causality difficult. In November 1984, the FTC opened a new rulemaking that would override remaining state laws that continue to restrict commercial practices of opticians.

These information difficulties, I suspect, are largely responsible for the pattern often observed in these markets: large amounts of price advertising, significant variation in item pricing over time, features, coupons, and substantial variation in retail margins across categories and brands of goods. The multidimensional and stochastic nature of the problem makes it similar to some of the quality information problems that have been addressed, but the ease with which prices can be changed by the seller and the large number of dimensions (creating the need for summary measures) make it essentially different from the usual quality problem. At a minimum, this pricing information problem defies the conventional wisdom that pricing issues are much simpler to analyze than quality information issues.[14]

Uncertain prices, that is, prices that are not fixed and known at the point of sale, also raise unexplored issues. It is traditional in the purchase of life insurance, for instance, that the buyer pays a nominal price at the point of purchase but that part of this price is later refunded as a "dividend" at the discretion of the seller. Similarly, the purchase of durable goods often requires the subsequent purchase of replacement parts at prices that are set by the seller. The fact that these institutions survive and are not replaced by full pricing contracts at the point of sale suggests that some type of reputation mechanism operates to discipline sellers.

The case of uncertain prices seems to fall much more directly into the reputation and implicit contracts literatures that have developed for product quality and labor issues.[15] Yet even with reputational constraints, it would seem that this pricing problem could still be colored by Stigler-type information costs leading to the same type of supra-competitive prices and price dispersion. Whether these information costs would have differential effects on the different parts of the price seems to depend fundamentally on the nature of the reputation mechanism. For instance, if the consumer's assessment of the seller's reputation for dividends is completely determined by his own experience with dividends, pricing would be less competitive for the second part of the price and dividends would fall as the consumer ages. Other reputational assumptions would predict different pricing patterns. Because of the concreteness with which the *ex post* price can be measured, this type of market might allow for cleaner tests of both

[14] At the FTC, for instance, issues involving this type of market are repeatedly addressed: what deception standard should be applied to grocery stores attempting to claim "lowest prices"? how significant is it if too strict a policy is adopted? what merger policy should be adopted for grocery chains? are the information issues here significant and how should they enter the analysis? should firms be allowed to restrict access to "price checkers" who wish to publish the data? how important is it that private property concerns be balanced against the value of improved information in these markets? do grocery stores have appropriate incentives to stock sufficient merchandise when they advertise price specials (Unavailability of Advertised Specials Rule — currently being reconsidered)?

For evidence on grocery retail margins across categories, for instance, see Albion (1983). See also Lynch (1983) and Steiner (1984) for discussions of retail pricing issues.

[15] See, for instance, Allen (1984), Carmichael (1984), Klein and Leffler (1981) and Shapiro (1983).

the Stigler-type information phenomenon and the effectiveness of reputations and the speed with which reputational adjustments are made.

B. Information About Quality

The behavior of markets with imperfect information about the quality characteristics of goods and services has been the focus of a great volume of recent literature. Here it is much more difficult to summarize the major findings neatly. The literature has a number of branches, and there is little empirical testing to rank these developments.

The basic "problem" is generally agreed upon: if it is difficult for consumers to assess the quality of goods sold by individual sellers in a market, there is an opportunity for sellers of low quality goods to attempt to pass their goods off as high quality goods. If successful, competition will drive sellers of high quality goods from the market. This is true even if consumers can judge the average quality available. In the case where it is impossible to assess the quality of individual sellers' goods, this result is typified by the Akerlof (1970) "lemons" model in which, in his example, owners of the best used cars find that their cars are worth more than the prevailing market price and therefore do not offer them for sale. As a result only the lowest quality cars ("lemons") are sold. More general models, as those in which it is costly (rather than impossible) for consumers to assess the quality of goods offered by different sellers or where information costs differ across consumers or sellers, would modify the Akerlof result as in the case of price information above; at a given price, average quality would be lower (though not necessarily at the minimum level) and quality dispersion might remain when compared with perfect information results.

The first two sections below review the purely private approaches to remedying this quality problem: those involving information provided by the producer directly and those relying on information provided by others. The third section briefly discusses contractual approaches which can be purely private or which can rely on judicial enforcement. Finally, the literature on policy approaches is very briefly discussed, including the literature on liability rules, policies towards deception and fraud, and direct information and product quality policies.

1. *Information Provided By Producers: Signals, Bonds and Reputations*

A premise of the quality degradation "problem" is that consumers do not have accurate quality information. One possible source of quality information is, of course, producers themselves, since they usually know the expected quality they deliver. In the abstract, claims by producers are suspect because of their incentive to exaggerate the value of their goods. If producers' claims could be made credible, market performance could be substantially improved.

A primary focus of the recent literature in this area has been the identification of conditions under which producers can credibly make quality claims. The essential finding of this effort is that manufacturers' claims can be relied upon if manufacturers possess or can purchase other observable characteristics that are economically associated with the hidden quality characteristics.[16] This association can be derived from some inherent cost or productivity relationship between the observable characteristic and the hidden quality, or from some bonding mechanism which gives the producer an economic stake in providing the promised quality. In the literature, these issues are discussed under the somewhat overlapping topics of signals, bonds and reputation.

In information theory, a *signal* is any bit of information that can improve the predictability of a second bit of information. In an economic setting, this predictability is derived from economic forces: for an activity to serve as an economic signal of quality, it must be less costly (or more productive) for high quality sellers to undertake the activity than for low quality sellers to do so. Spence (1974) initially introduced this idea in a labor market setting where higher quality workers were able to reveal this fact through an investment in education, because for them education was less costly to obtain than it was for lower quality workers. The idea has broad potential applicability: the use of warranties by sellers of high quality goods who would expect to pay less under the warranty; the amount of advertising used to attract new customers by sellers who depend on repeat purchase or referrals by satisfied customers (Nelson 1970); and the higher deductible amount chosen by low risk insurance buyers who expect to lose least from this choice (Rothschild and Stiglitz (1976)) are standard examples of economic signals.

Bonding devices, or *bonds*, are capital assets or secured monies which are forfeited if the bonded party does not perform as promised. In many market settings, voluntary bonding devices act as signals of quality: if the bond is sufficiently large, the presence of the bond reveals that the firm or individual does not plan to offer a low quality good or service; the loss of the bond value, once the low quality is discovered, is sufficient to make cheating uneconomic. Viewed in this way, quality-specific investments become information devices (Klein and Leffler (1981)). For example, designing and furnishing retail establishments so that they cannot be easily transfered to other uses acts as an assurance to customers that the firm can be relied on to provide the promised quality. Investment in durable brand name recognition, through advertising or other means, has the same effect (Kihlstrom and Riordan (1984)).

In general, quality bonds are signals of quality, but not all signals are necessarily bonds. To see this, it is important to make a distinction between two different types of quality issues: those where cheating can be detected (at least to some extent) *ex post* and those where it cannot. Bonding devices can be used only in the first case where "cheating" is detectable, since bonds depend fundamentally on the "punishment" inherent in the loss of the bond value. In contrast, non-bonding signals can arise even if low quality can never be detected in in-

[16] The arguments in this section apply equally to hidden price characteristics as discussed above.

dividual units as long as there is some (known) cost or productivity advantage in acquiring the signal for high quality goods or sellers. In the Spence labor example, for instance, education is a non-bonding signal that screens higher ability workers from the pool of workers on the basis of a cost advantage in acquiring education; it has no bonding effect to prevent shirking by any type of worker once employed.[17]

The term *reputation* is used in the literature, and in common language, to capture the idea that a firm's quality claims can be relied upon. In economic usage, a firm's reputation is a bond which usually requires some investment to acquire. The stock of goodwill inherent in the firm's reputation can be lost if cheating is detected. For reputation to induce a firm to continually provide high quality goods, the firm's goodwill (like any bond) must be able to generate a stream of price premiums which will be lost (at least in part) if the firm cheats (Klein and Leffler (1981) and Shapiro (1983)). Moreover, this loss must be greater than the short-term gain from cheating. The premiums are the market return to the reputation asset, and in a competitive setting must be secured on the margin by firm-specific sunk costs sufficient to justify this return.

Policy Issues

From the perspective of consumer protection policy, the most pressing issue in the quality assurance area is the development of a better understanding of the market conditions under which producer signals or bonds will be a reasonably effective check on quality information problems, and ideally some empirical evidence that supports that understanding.[18] From the theoretical developments to date, we can draw a few principles to guide policy. I will mention two:

i) The availability of information about cheating is a critical component of the market's use of bonds.

An essential feature of any quality bond, like reputation, is that cheating must be sufficiently discoverable. Otherwise, the bond will have little or no disciplining effect on the firm's behavior.[19] Because the ability to punish the firm is distributed among individual consumers, it is the summation of individual reactions that is ultimately important in disciplining cheating. The ease with which cheating is detected and the speed with which this information spreads to future consumers should influence the size of the bond (and therefore the price premium) necessary to secure performance.

---

[17] Under different economic assumptions, of course, educational expenses could serve a bonding function; for instance, this might be the case where education is specific to an occupation and where poor performance or ability is at least somewhat detectable.

[18] Two recent empirical tests of the "lemons" model in product markets are Bond (1982) and Lacko (forthcoming).

[19] On this count, Klein and Leffler (1981) assume that any cheating is immediately known by all consumers. Shapiro (1983) allows for a lag in discovery and some averaging over time, but again the knowledge is held by all consumers.

Characteristics of the market and of the particular type of quality involved should be relevant here. For instance, for stochastic quality issues involving differences in a low rate of defect, only a small percentage of consumers will actually get a defective good, and they will not be able to determine easily whether their breakdown is symptomatic of a higher breakdown rate overall. In contrast, a quality issue involving the use of inferior materials and a shorter useful life for all units of a good should be more easily diagnosed by a much larger segment of the market. To achieve the appropriate incentives in both cases, the bond and the premium would have to be much higher in the first case than in the second (assuming the same total reduction in value from cheating).

More generally, reputation or other bonding devices will be more effective in cases where information about cheating will spread broadly and clearly; for instance, where the quality degradation is widespread among consumers, where its cause is easily diagnosed, and where *ex post* information from other sources is widely available. This is the reason that the literature has so consistently focused on the frequency of repeat purchase as an important determinant of effective reputations: more frequent purchase by all consumers improves the development and spread of information, should cheating occur. However, it is important to an understanding of reputation and other quality bonds that the focus be kept on the relevant issue — the spread of accurate information if the firm cheats — rather than on one particular way in which that information will spread.

ii) The availability of sunk cost intensive production and distribution processes is necessary for bonding in a competitive setting.

For reputations or other quality bonds to survive competitive pressures, they must be secured on the margin by sufficiently large firm-specific sunk costs. In particular, this implies that the firm will adopt a production and selling process that is not necessarily cost-minimizing in the narrow sense. Thus, reputation or other quality bonds will be more effective quality controls when the magnitude of the cost implicit in this shift to a more sunk cost intensive process is relatively small. In markets where alternative production or distribution technologies are readily available which require sufficiently high sunk costs, bonding is a low cost alternative for dealing with this type of information problem; but bonds will be a high cost alternative in markets where both production and distribution inherently involve no sunk costs.

The rapid development in understanding these market approaches for dealing with quality information problems has had some impact on consumer protection policy, especially on economists' views of it. Certainly, the suspect and almost hostile view that was held towards advertising and other visible selling expenses in the past (see Scitovsky (1950), for instance) has changed dramatically in recent years. Proposals that explicitly or implicitly reduce the quantity of advertising, such as mandatory disclosures in advertisements, are treated more cautiously today. The focus of advertising regulation is held more tightly to issues of deception in the advertisement itself and the costs and benefits of correcting that deception. Also, accounting measures of profit in consumer goods industries are more generally questioned because of the difficulty in accounting for reputation assets. In the practice of antitrust, however, the quality assurance role of sunk assets has yet to achieve much recognition;[20] here sunk costs and reputation are generally dealt with in the context of the "entry barriers" debate with its focus on the non-informational role of sunk costs in determining market structure and innovative activity.

Finally, in consumer protection policy involving product quality issues directly, more attention is now given to the issue of whether the market is likely to discipline firms adequately,[21] and if the judgment is that it will not, to whether narrow information remedies might be sufficient to support otherwise private mechanisms. For instance, if quality measures of performance were made available even with a substantial lag, the sunk cost requirements to support reputations might fall enough to be supportable.[22]

2. *Information Provided By Others: Standards, Certification and Retail Distribution*

Information about quality can also be provided by private parties other than the producer. Privately-developed standards and certification procedures and independent information providers are the primary examples in this class.

In the U.S., there are a large number of private organizations that develop standards of quality, sizing and compatibility for many product categories. These organizations are often non-profit groups set up as part of industry trade associations and are funded in a wide variety of ways, including fixed membership fees and direct quantity assessments.[23] Sizing and grading of lumber, model codes for building construction, and toxicity standards for children's crayons are just a few examples of privately produced standards. Many of these standards are invisible to consumers, effecting their quality improvements through intermediaries, but some, such as the Underwriters Laboratory certification, are aimed directly at consumers. Despite their prevalence and widespread use, private standards have been virtually ignored in the economics literature; there are no theories predicting when industry agreements will arise and how the standards produced will compare with efficiency criteria.

Certification procedures are similarly widespread. Independent testing laboratories test and certify that products meet specified quality standards. In

[20] The most significant exception might be antitrust cases dealing with vertical issues.

[21] For instance, see the recent FTC policy statement on deception which incorporates this view (reprinted at 45 *BNA Antitrust and Trade Regulation Reporter* 689 (Oct. 27, 1983)).

[22] For instance, FTC rules and cases in recent years have involved a number of information approaches: mandatory information disclosures of quality indices as in the R-Value Rule for insulation and the Octane Rule for gasoline, direct measurement of quality as in the measurement of tar and nicotine for cigarettes, and broader distribution of private information as in cases involving auto defects where the firms were required to more widely distribute information about problems as they are discovered (Calfee and Ford (1985)).

[23] For a description of some of these organizations, see Hemenway (1975) and FTC (1978).

some cases, the producer secures this certification directly as an assurance to his customers that the product meets certain quality standards. For instance, most electrical product manufacturers display the Underwriters Laboratory seal directly on their products. In other cases, retailers (or other large buyers) purchase the testing themselves before allowing the goods on their shelves.

More generally, it has long been understood (see Stigler (1961), for instance) that multi-product retailers can serve a quality assurance role by selecting and screening products. In this case, the certification is less formal than that of the independent laboratories, but the function is essentially the same.

Finally, independent information sellers do exist who sell information directly to potential buyers. *Consumer Reports*, *Good Housekeeping* and some of the auto magazines are prominent examples of mass-marketed, quality information providers. House inspection services for would-be purchasers and antique and jewelry appraisers are examples of information providers who give more direct quality assessment advice.

In each of these cases, the information provided is subject to the same types of quality problems they are designed to solve. Quality assurance devices, like reputation, thus become essential to guarantee the quality of the information provider himself. Moreover, the public good nature of the information when it is sold separately makes it very difficult for the seller to collect much of the value of the information. Some of the public good problem is circumvented when information is provided by the producer of the good directly.[24]

One of the relatively unexplored economic topics raised by these issues is the allocation of the quality assurance role between the producer and the seller, and the effect this allocation has on empirical work done in consumer product industries. For instance, the retail margin on the generic (or non-leading brand) version of a good is often larger than the margin on the leading brand.[25] If this difference is determined by the differential quality assurance roles played by the retailer and the producer in the two cases, welfare implications related to the determinants of manufacturer or retailer returns would have to be carefully considered. Much of the advertising-price literature that demonstrates that manufacturers' prices are increased by advertising, as summarized in Comanor and Wilson (1979), for instance, would be subject to reinterpretation: higher advertising levels by manufacturers could simply reflect a shifting of the quality assurance role to the manufacturer, requiring a corresponding shift in the price premiums that quarantee that quality. Welfare conclusions from other branches of the advertising literature would be similarly affected.

Policy Implications

The improved understanding of the importance of information providers has

[24] See Beales, Craswell and Salop (1981) for a more thorough discussion of these issues.

[25] See Masson and Steiner (forthcoming) for ample evidence of this margin difference in the case of prescription drugs. See also Albion (1983) for evidence related to grocery products.

yet to have substantial effect on policy, though it is contributing to the growing controversy in some areas of antitrust policy and is opening up new areas of investigation. The active policy debate on the current *per se* prohibition on a manufacturer's ability to control retail prices through resale price maintenance (RPM)[26] is being fueled in part by an understanding that retail margins can affect the quality assurance role provided by certain retailers. For instance, if high profile department stores provide a fashion or other quality certification to a manufacturer's goods, the manufacturer has a direct stake in ensuring that retail margins are sufficient to persuade these retailers to carry his goods. RPM laws, in conjunction with the Robinson-Patman prohibition on price discrimination, currently make this quality certification difficult.

In the standards and certification area directly, the FTC has recently dropped a rulemaking that proposed to regulate all private standard setters and certifiers.[27] In its place, the agency has opened a program to investigate particular standard setters to determine if some individual standards have anticompetitive effects in prohibiting entry or limiting innovation. More generally, the public good nature of standards might be used to justify direct public support for standard creation (as in the support provided to private standard organizations by the National Bureau of Standards) and the development of standards by the regulatory agencies (for example, the mileage ratings by the Environmental Protection Agency and the tar and nicotine ratings by the FTC). This public good justification for government support and development of standards must, of course, be balanced against the concomitent political economy problems that inevitably arise. Once developed, for instance, government standards often effect a near monopoly on the measurement of relative quality, creating powerful incentives for affected parties to attempt to influence the development of the standards for their benefit.

3. *Contracting Approaches to Quality Problems*

Contracts are one of the long-standing approaches for dealing with trading problems where the quality of the good or service is not apparent at the point of sale. The law and, more recently, the economics literature make a distinction between explicit and implicit contracts.

Explicit contracts are usually written promises that specify either what will be delivered or the parties' responsibilities in such events as product failure or late delivery. In consumer goods markets, explicit contracts are generally warranties, although in the services area contracts specifying the good itself do arise. Credit contracts and health club agreements are examples of the latter.

Explicit contracts are usually taken to be legally enforceable in the sense that if disputes arise about performance under the contract, the injured party has the option of appealing to the courts to enforce the contract. In some consumer

[26] See Overstreet (1983) for a review of the RPM area.

[27] See FTC (1978) or *Federal Register* of December 7, 1978 — 43 FR 57269 for the statement of the proposed rule.

contracts, a more private approach is taken where the contract specifies that disputes will be first submitted to a private arbitration board, like that run by the Better Business Bureau.

In the economics literature, *implicit contracts* are usually contrasted with explicit contracts in two important ways: implicit contracts are not written, and they are not taken to be legally enforceable. The recent use and development of the implicit contract idea in economics has taken place primarily in the labor literature in an effort to understand employment relationships,[28] but the ideas and fundamental issues are directly applicable to product quality problems. It is usually argued that implicit contracts are used in cases where the conditions and responsibilities under the contract are too difficult or too costly to specify for inclusion in an explicit contract.[29] Because the parties to an implicit contract do not have legal recourse in the event of breach, implicit contracts must be self-enforcing or must be enforced by reputations. The ease with which a breach is discovered and that information spread to future customers again becomes a critical issue.

In consumer product markets, the idea of an implicit contract is therefore simply an alternative way of conceptualizing the promises a producer or seller can credibly make to consumers. In this sense, the discussion of information provided by producers in section 1 becomes directly applicable. Intuitively, the "implicit contract" label seems more appropriate for performance issues over time, as in specifying producer responsibility in the event of product failure, but substantively, the producer's credibility for product quality claims and for future performance claims involves the same issues.

The distinction between explicit and implicit contracts is more apparent than real in consumer product markets. Consumer lawyers would be quick to point out that some implicit contracts could be legally enforced, for instance, under the implied warranty of merchantability or the implied warranty of fitness for a particular purpose of the Uniform Commercial Code. More important, I think, is the "implicit" nature of most explicit contracts offered for consumer products. Most of these explicit contracts are not worth enforcing should a breach occur: the cost of raising the issue legally is much larger than the injury from the breach.[30] In these cases, few consumers would be expected to exercise

[28] See, for instance, Azariadis and Stiglitz (1983) and the other articles in that volume for recent developments.

[29] While plausible, this argument does raise issues, since the enforceability of contracts often depends on a reputation mechanism, which in turn depends on consumer reaction to cheating. If contract terms are complex, cheating may be difficult to detect and consumer reactions weak or haphazard. See Newbery and Stiglitz (1983) for further discussion of this point. Implicit contracts which are simply too costly to specify, for example, because there are many combinations of conditions which could arise in which behavior would have to be (and could be) specified, do not raise this problem.

[30] Warranties on toasters and coffee pots illustrate this point dramatically, but even most automobile defects would fall short of covering the legal and time costs of actually enforcing the warranty on an unwilling manufacturer. See Priest (1981) for an interesting discussion of consumer warranties.

their legal rights if the contract is not honored, making it difficult to argue that legal enforcement is a significant disciplining mechanism in these markets or a significant explanation for the explicit contract.[31]

The widespread prevalence of explicit warranties and other consumer contracts in these situations is thus somewhat puzzling in light of the standard theories of explicit contracts. Possibly these explicit contracts are offered not because they are legally enforceable, but because they are superior information devices for the development and efficacy of reputations. A written warranty, for instance, is a more specific articulation of a minimum that the manufacturer is willing to promise the purchaser in the event that problems arise with the product. If the producer does not honor the contract, all consumers who experience the problem will be more certain that a breach has actually occurred. This improves the clarity and diffusion of information about cheating and increases the effectiveness of reputations as a disciplining device in the market.

Viewed in this way, warranties or other consumer contracts are subject to the same quality problems as other goods. The more contracts are made clear and explicit, the more they take on the characteristics of experience goods (in the Nelson (1974) terminology): the quality of the contract becomes apparent after the purchase. Reputation and other bonding devices are more effective for experience goods than for goods with more credence-like characteristics.

Policy Issues

Our limited understanding of when and how consumer contracts are enforced and the almost total lack of evidence on these issues makes policy decisions in this area difficult. Even in the simplest case of explicit contracts, for instance, the best policy is uncertain. There might be a role for public enforcement of consumer contracts in cases where a seller systematically fails to live up to his contract and where the cost of each individual bringing a case is large relative to the loss. If a public agency can aggregate the claims in a way that reduces legal costs sufficiently, public enforcement might be economically justified.[32] However, this enforcement tends to undermine the value of reputation, since consumers no longer need to rely on reputation as much. Whether an equilibrium that depends more on public enforcement is better than one that depends totally on private reputations is an empirical question about the efficiency of alternative market institutions — a question on which we have virtually no evidence.

In cases where explict promises are not made, it is even more difficult to sup-

[31] The economics literature on warranties generally assumes that all warranties are costlessly enforced. The extension of these analyses to costly enforcement is straightforward, and in cases where the cost of enforcement is large relative to the loss, this extension would completely undermine the value of warranties in these models. See, for instance, Courville and Hausman (1979), Grossman (1981), Palfrey and Romer (1982), and Matthews and Moore (1983).

[32] At the FTC, for instance, warranty performance cases are usually justified on this ground. This is also the basis on which class action suits were established.

port policy initiatives. In particular, the major question of what the implicit contract is must be added to the explicit contract questions. Let me mention two types of cases where this issue has arisen in recent FTC actions: the first deals with systematic post-warranty failures and the second with firm's post-sale information obligations.

In a series of cases beginning in the late-1970s,[33] the FTC issued legal complaints against a number of auto and other durable goods manufacturers who were alleged to have knowingly sold goods with systematically higher failure rates in major components of the good and who were not voluntarily accepting liability for the "unexpected" post-warranty failures. The primary legal basis for the complaints was a deception charge: namely, that the manufacturer failed to disclose material facts that would have affected purchase decisions and use and care decisions. Essentially this charge assumes that, absent a disclosure to the contrary, consumers expect components of the product to exhibit "normal" failure rates and "normal" maintenance costs and that they expect "abnormal" failures to be the responsibility of the manufacturer.

In support of this implicit contract theory, proponents point to the fact that the major auto manufacturers sometimes extend *ad hoc* warranty coverage beyond the legal warranty on unusual problems and that consumer complaints show that consumers expect manufacturers to share in the costs of such repairs. However, even if the existence of such an implicit contract is accepted, there is currently no way to determine the terms of that contract for enforcement purposes and no explanation for why the contract was breached in a particular instance when it is honored in other cases. Further, a policy of public enforcement fundamentally alters the concept of an implicit contract as one that is self-enforcing or disciplined by reputation. To support such a policy requires a move into the more nebulous realm of implicit contracts that would be agreed upon if enforcement were feasible.

The second type of case involves the manufacturer's failure to disclose certain information discovered after the sale. Here the issue typically involves the post-sale discovery of a design defect and the identification of some related maintenance action that would reduce the expected cost of the problem. In this type of situation it is puzzling that manufacturers do not contract with consumers to provide the maintenance information to reduce the expected costs of operation.[34] Of course, a major problem with such a contract is that consumers and third parties would find it very difficult to determine when the contract has been breached, since the information is discovered and held privately by the manufacturer. The argument implicit in the FTC cases of this type is that public enforcement is needed to provide the discipline to support otherwise efficient con-

[33] See Calfee and Ford (1985) for a brief description of the program and some of the cases in it.

[34] A recent example here is a defect case against International Harvester (FTC Docket No. 9147) in which the company was found in violation of the FTC's unfairness statute because it failed to disclose a safety problem with the fuel covers on its equipment, and there were cost-effective actions that owners could have taken to reduce the likelihood of injury.

tracts for post-sale maintenance information. There has been very little systematic study of issues of this type and no relevant empirical work. The incentives for producers to supply information on a continuing basis are not well understood.

Beyond these enforcement and implicit contract issues, a primary focus of traditional consumer protection policy on contracts has been the concern over particular contract provisions themselves, often attacked under the legal doctrine of unconscionability.[35] Prohibitions on the required use of manufacturer's parts as a condition of the warranty (Eisenach, Higgins, and Shugart (1984)), restrictions on creditor remedies (FTC Credit Practices Rule), and limitations on the disclaimer of implied warranties (UCC) are all examples of specific contract term regulations.

When contracts are viewed as goods, the concern over particular contract provisions is essentially a concern over the "quality" of the contract. In this sense, the regulatory prohibitions on contract terms could be viewed simply as minimum quality standards designed to preclude "low" quality contracts from the market. However, the types of contract provisions at issue in these regulations often do not follow unambiguous quality dimensions. Moreover, the move away from direct regulatory approaches has been somewhat slower in the case of contracts when compared to that for products. In current consumer protection regulation, direct minimum quality standards are viewed as a rather restrictive regulatory approach for product quality problems (in fact, they are generally used only for drugs, food and safety issues). But minimum quality standards are still commonly considered for contract provisions in the regulatory setting and are implicitly adopted in the judicial determination of unconscionability. Possibly the fact that there has been very little formal analysis of contracts that would parallel even the theoretical analyses of the quality and price of goods in markets with information asymmetries has contributed to this more static view.

4. *Policy Approaches to Quality Problems*

(a) Liability Rules

The economic literature on liability standards has developed substantially in the last 10 years. Much of this work has focused on legal liability as a solution to third party injury problems, like automobile accidents, and to externality problems, like pollution.[36] However, liability rules can also be used in some market settings to increase product quality. In cases where information problems result in products with too low quality, legal liability standards are another policy option to change the equilibrium market quality.

The economics literature in this area has generally focused on the relative

[35] See Kornhauser (1976) for a discussion of the history and current questioning of legal views.

[36] See, for instance, John Brown (1973) and Steven Shavell (1980) for examples of the accident literature and Polinsky (1980) for the externality literature.

merits of alternative standards of liability, including strict producer liability, negligence standards and no liability in cases where product quality is measured as a failure rate.[37] Under the premise that consumer beliefs are an exogenous function of true quality (in particular, that perceptions are not influenced by other actions of sellers), the literature finds that liability standards that shift more liability to the producer improve welfare in cases where consumers underestimate true product quality and underreact to changes in actual quality.

This result is in marked contrast with much of the standard policy/legal discussion of product liability where consumer beliefs are assumed to be endogenous. In particular, the consumer is assumed to be led to expect a certain product quality only to be disappointed with the product that is actually provided — that is, liability rules are required because the consumer (justifiably) overestimates product quality. In this view, the producer is implicitly assumed to be able to influence consumer perceptions of quality and to profit if he can raise beliefs above the quality he will actually deliver. With its roots in contract law, liability for product failures (or any other measure of low quality) is viewed as a means of reducing producer incentives to deceive and of inducing him to provide the quality that is actually promised. In contrast with economic models, for instance, the policy/legal view would impose no liability if the producer adequately disclosed that his product was of low quality.[38]

These divergent perspectives on the role of liability rules for quality problems highlight the fact that it is not consumer beliefs *per se* that are important in determining appropriate liability rules, but the entire process by which these beliefs are formed. The economics literature on liability rules implicitly assumes that sellers of higher quality products cannot make credible claims about their higher quality. As a result quality supplied in the market falls to too low a level reflecting consumers' (exogenously determined) belief process. In this setting, liability is used to raise the sustainable minimum quality level.

In the policy/legal view, liability stems more from the violation of an explicit or implicit quality claim than from some derived notion of efficient quality itself. In this sense, liability rules are seen to improve the overall credibility of quality claims, rather than to increase product quality directly. Thus, in contrast with the static consumer belief formation assumption of current economic liability models, the legal view is fundamentally premised on the idea that consumer beliefs can (and should) be influenced by producers. Liability rules are conceptualized as a method for improving the results of this belief formation process.

Liability Policy Issues

The idea of tying liability to the quality claims made for a product has significant appeal in designing liability standards for policy use. At least theoretically

[37] The primary article in this area is Spence (1977). Articles by Epple and Raviv (1978), Shavell (1980) and Polinsky and Rogerson (1982) are closely related.

[38] See, for instance, Posner (1972).

this improves the market's ability to offer a range of qualities to satisfy different consumer preferences. At a practical level, however, there are many unresolved problems. In complex products, it is not feasible for a manufacturer to meaningfully disclose the "quality" of all component parts, and in a relevant sense, the quality of each component does not matter — the overall quality of the product does. Yet when "problems" occur, they are usually problems with components. It is difficult for the legal system to evaluate a design defect in the context of the "whole product" — for example, should a better than average steering mechanism and suspension system be considered in deciding liability for a defective drivetrain in a car? Yet, if a producer claims that his product is "better than average," failure of any single component means little in terms of the overall quality claim. More importantly, if explicit claims are not made for the product or particular component, liability must be assigned on the basis of an implicit claim. This raises essentially all the issues connected with designing a minimum quality liability standard.

More generally, moving from a minimum care liability approach to one tied to the producer's explicit or implicit claims is essentially a move to legally enforceable implicit contracts with all the problems discussed in the previous section. Because of the difficulty in effectively aggregating most consumer claims and because of the cost of litigating the issues involved, it seems doubtful that a large level of activity could be economically justified under either approach. However, there is no evidence to support this (or any other) position on the merits of product liability laws, and there appears to be a growing volume of product liability litigation in the U.S.

(b) Policies Towards Deception and Fraud

An alternative policy approach to quality problems is to focus directly on the lack of credibility of sellers' claims that leads to the quality problems. Laws against deception and fraud do this by making it more costly for producers to communicate false information (because of the risk of prosecution), and therefore, presumably these laws increase the truthfulness and credibility of claims that are made. Thus, the laws might be viewed as providing sellers with a more credible communications channel to consumers. At a basic theoretical level, the issues here are relatively straightforward, and there is general support for some policy against deception and fraud.[39]

Nevertheless, deception law has become a controversial area of consumer protection activity. Most of the recent debate has focused on the FTC's policies towards deceptive advertising: a number of bills have been introduced in Congress that would change the statutory standards applied to deceptive advertising,[40] and there has been sharp disagreement both inside and outside the Com-

[39] See, for instance, Darby and Karni (1973), Posner (1979), Pitofsky (1979) and Grady (1981).

[40] *Hearings on the FTC's Authority Over Deceptive Advertising*, Senate Commerce Committee, (footnote cont'd)

mission about which changes, if any, should be made.[41] Currently the review courts require only a showing that an advertisement has the "tendency or capacity to deceive" to be in violation of the FTC Act;[42] for instance, there is no legal requirement to show either actual deception or consumer injury from the deception. Moreover, in understanding the current policy debate, it is important to note that most advertising cases involve implicit rather than explicit claims in the ads; in fact, the primary dispute in a deceptive advertising case is usually about whether the questionable claim was actually made by the advertiser.[43]

At the heart of this policy debate are some fundamental issues. Those resisting the move towards a more stringent deception standard consider a strong deception law essential to restrain what they see as firms' powerful incentives to deceive consumers. A broad mandate to control deception, subject only to limiting exceptions worked out through the judicial process of precedent and appeal,[44] seems to them to put the burden of law where it should be — on the firms who can control deception. Even when they agree that current policy does not (and should not) take advantage of the full rein of the law, they oppose efforts to narrow the legal mandate, preferring to rely on prosecutorial discretion and voting requirements at the Commission to control excesses.[45]

In contrast, critics of the Commission's past deceptive advertising activity generally point to the large number of cases generated under this legal arrange-

97th Cong., 2d Sess., Ser. No. 97-134 (1982). The enabling legislation for deceptive advertising cases is again Section 5 of the FTC Act, which states that "... unfair or deceptive acts or practices in commerce are hereby declared illegal." See Beales, Craswell and Salop (1981) for a more thorough background discussion of deceptive advertising policy; also Craswell (1984).

[41] See, for instance, the FTC majority's "Policy Statement on Deceptive Advertising," reprinted at 45 *BNA Antitrust and Trade Reg.* Rep. 689 (Oct. 27, 1983) and the dissenting FTC minority view in "Analysis of the Law of Deception By Commissioners Patricia P. Bailey and Michael Pertschuk," reprinted at 46 *BNA Antitrust and Trade Reg. Rep.*, 372 (March 1, 1984). See also "Dingell Assails FTC Chief on Deceptive Ad Issue," New York Times, Oct. 27, 1983, p. 16. For a review of the recent debate, see Ford and Calfee (1984).

[42] Though most of the debate has focused on the FTC law, some of the same issues are raised by private deceptive advertising suits brought under Section 43(a) of the Lanham Act (15 U.S.C. 1125(a) 1976). For a discussion of private remedies, see, for instance, Thomas J. Donegan, Jr., "Deceptive Advertising May Be Remedied By Courts," *Legal Times of Washington*, September 7, 1981, p. 14.

[43] It might be tempting to argue that deception policy should be limited to explicit claims because of the inherent difficulty of judging implied claims. However, this restriction is a very strong one legally. For instance, consider an advertisement like: "You have all heard a lot of talk recently about the importance of a high fiber, high nutrition diet. To get the most out of life, eat a well rounded diet. Next time you visit your favorite grocery store, look for our new Gran-O brand cereal." Even if "Gran-O" contains no fiber and is less nutritious than every other cereal on the market, there are no false explicit claims in this ad; it is an implicit claim that "Gran-O" is a nutritious cereal.

[44] For instance, there are legal precedents that protect non-factual exaggerations or opinions ("puffery") from scrutiny under deception law. See, for instance, the FTC's *Consumer Information Remedies Policy Review Session* (1979) for a discussion of these issues.

[45] See, for instance, the testimony of Commissioners David Clanton, Patricia Bailey and Michael Pertschuk at the Senate Commerce Committee Hearings (1982) as cited above.

ment where consumer benefits seem quite implausible.[46] While earlier critics argued for more reasoned case selection at the Commission, some current critics favor an explicit narrowing of the FTC's legal authority by including more economically-based criteria in the deception standard itself.[47] In addition to reducing the number of questionable cases actually brought, this increased and more explicit evidentiary burden for the Commission[48] is seen as necessary to correct the undesirable damping of truthful advertising claims created by the breadth and uncertainty of the current legal standard. After an unsuccessful effort to narrow the FTC's deception authority in the legislature, in 1983 the Commission (by a 3 to 2 vote) issued a policy statement on deception that incorporated some of these more explicit criteria.[49]

Compared to those favoring the *status quo*, critics of the current legal standard generally share a more confident view of the market's ability to control most deceptive advertising. In keeping with the recent economic literature's analyses of bonds, signals, reputations and competitive information incentives, they see much less opportunity for firms to engage in profitable deceptive advertising. They do not share the fear that a narrowing of the Commission's authority will unleash strong incentives for firms to deceive the consuming public. Moreover, they have a much more sanguine view of consumers' ability to effectively identify and react to many types of deception that the FTC has pursued in the past.

Thus, the fundamental issues in the deception policy debate are the unanswered empirical questions currently dominating the entire product quality information area. There is little systematic empirical evidence that confirms the market's ability to police deceptive claims in the circumstances predicted by theory. On the other hand, the questionable nature of many past Commission cases may indicate that little significant deception occurs. Further, there is no evidence with which to judge the deterrent effects of current advertising policy or to judge the likely increase in deception claims if the standards are relaxed. More subtly, the relationship of policy to the formation of consumer beliefs is not well understood; for instance, if deception that caused only limited injury

[46] Commonly cited examples are past FTC interpretations that consumers are deceived by celebrity endorsements if there is no disclosure that the celebrities are paid for their efforts, that consumers interpret claims of savings "up to" a specified amount to indicate the average savings for consumers, and that advertisements that report the opinions of reputable publications are deceptive unless the manufacturer has independent tests confirming those opinions.

[47] See, for instance, Posner (1973), Jordan and Rubin (1979), Grady (1981), Craswell (1984), Ford and Calfee (1984), testimony by Commissioners Miller, Douglas and Clanton before the Senate Commerce Committee (1982) as cited above.

[48] Judges have consistently given the Commission the broadest latitude in judging the deceptiveness of advertising claims. Findings of deception by the Commission are rarely overturned.

[49] For instance, the former Bureau Director for Consumer Protection Timothy Muris proposed a legal standard of deception that would find an act or practice deceptive "if it would mislead consumers, acting reasonably in the circumstances, to their detriment." See "Memorandum ... on Definition of Deceptive Advertising" at 699 *BNA Antitrust and Trade Regulation Report* 42 (Sept. 25, 1982).

was allowed to persist unchallenged by policy, would the credibility of more significant advertising claims suffer?

Finally, an issue that is not at the forefront of the current policy debate, but is growing in significance, is the policy distinction between affirmative deception (that is, when explicit or implicit claims are made which are false) and deception by omission (as when a negative feature of a product is not disclosed). The trend in Commission activity in recent years has been towards increased scrutiny of deception by omission issues,[50] and — in what comes very close to omission cases — broader interpretation of implied claims in ads. This movement in the law deviates substantially from the more neutral view of deception law as an institution designed to provide sellers with a credible communications channel to consumers. Instead, this shifts deception law towards the more active role of forcing information into a market that would (presumably) produce too little on its own. This issue deserves closer examination on economic grounds. In particular, such a trend appears to require acceptance of the proposition that even when producers can make (only) credible claims, market incentives are not sufficient to induce producers of better products to distinguish themselves from lower quality producers.[51] The usual situation offered as problematic in this regard is one in which the whole product class has a hidden risk or flaw (e.g., the health risks of cigarettes in the 1950s). However, even here there seem to be substantial profit opportunities for firms who offer higher quality products.[52]

(c) *Direct Government Regulation: Mandatory Information and Minimum Quality Standards*

Among the policy options for dealing with product quality problems are the more direct regulatory approaches: government-mandated quality disclosures and minimum quality requirements for the products themselves. Mandated quality disclosures address the underlying information deficiency by legally requiring firms to reveal the hidden quality dimension. In contrast, minimum quality standards focus on the equilibrium quality itself. Theoretically, these standards either raise the single quality level offered in a "lemons" equilibrium, or, under some market conditions where producers have different cost functions, raise the

[50] In a significant departure from this trend, in December 1984 the Commission overturned an administrative law judge's finding of deception against International Harvester (FTC Docket No. 9147) for failing to disclose a safety hazard. In a 3 to 1 decision, the Commission ruled that deception required some form of false affirmative representation, which may be either expressly stated or implied by the general circumstances surrounding the sale. International Harvester was found guilty under the FTC's unfairness authority.

[51] See Grossman (1981), for instance, for a model of how this quality unfolding can lead to full disclosure.

[52] In the cigarette case, advertising in the 1950s had begun to address the relative health benefits of different brands before the FTC issued "guidelines" that made it exceedingly difficult for a seller to advertise relative health benefits.

average quality in the market by directly limiting low quality products, and by attracting higher quality products because of the resulting increase in the equilibrium price.[53]

There are, of course, a host of practical problems in deciding on a measure of quality and in selecting an appropriate minimum quality standard when these approaches are adopted.[54] More fundamentally, however, there is currently only very limited discussion in the literature of economic rationales or evidence to support the choice of these direct regulatory approaches over the more market-based approach inherent in the fraud and deception laws. For instance, in requiring a disclosure of quality, there is an implicit rejection of the market's efficiency in inducing this information from the higher quality firms. Yet if the deception laws prevent false claims and if there is a credible measure of quality available, firms willing to provide higher quality should be able to effectively convey their quality to consumers and to profit in doing so.[55]

Casual empirical evidence gives some support to the market's effectiveness in such circumstances. Margarine sellers appear to have been quite successful in communicating their superiority to butter on the cholesterol issue. Low tar and nicotine cigarette sellers have been vigorous in distinguishing themselves from the higher tar brands (going far beyond the mandated disclosures in advertisements). High mileage automobiles often feature this fact in their advertisements. Lower calorie foods (especially in the diet soda and frozen food categories) have been very successful in conveying their superiority to higher calorie counterparts. The same is true for high fiber foods. In this type of case, the (admittedly casual) empirical evidence seems to bear out the theoretical work which suggests that there is no need for the government to mandate the disclosures.[56]

The availability of a clear and credible index of quality is an important component of the market's provision of information. Without a credible measure of quality, it is more difficult for firms to make quality claims[57] and for deception

[53] See Akerlof (1970) for the classic "lemons" model and Leland (1980) for the quality distribution case.

[54] See Beales, Craswell and Salop (1981) for a thoughtful discussion of many of these issues.

[55] See Grossman (1981). Here I am assuming that a measure of quality is available to the firms. The incentives to develop such a standard, where it is not available, are more questionable.

[56] It is tempting to argue that there is also little cost to such a mandate and that the potential for improved information in the market may make such an approach worthwhile. While this argument appears plausible for a single quality dimension, the problems of information congestion and overload limit its applicability. The choice of which dimensions to disclose, how aggressively that information should be spread, and how the dynamics of the market should be incorporated are not currently accommodated by government regulatory mechanisms.

[57] An example here might be the crashworthiness of automobiles. Crashworthiness is a complex quality that is very difficult to summarize in a simple measure. It is thus difficult for individual firms to convincingly advertise their superiority along this dimension. General Motors has recently experimented with advertisements for some of their models based on auto insurance statistics for fatal accidents, but these numbers are colored by driver selectivity factors. Other firms resort to

(footnote cont'd)

law to discipline misleading claims. However, the public good nature of such standards makes it less likely that firms will individually develop and promote quality standards in cases where there are significant costs to doing this. Private industry groups appear to overcome this problem to some extent, but we know very little about the economics of these standards setting groups. Some government support for the development of quality indices, where feasible, thus may be a more economically justified area of activity.

Of course, even with clear measures of quality, there might be a problem initially if consumers are not broadly educated on the quality issue; the health hazards of sodium or cholesterol consumption might be recent examples of this.[58] Firms in such a case might be reluctant to initiate a promotion campaign until consumers are aware of the basic issue; other firms could free ride on the initial, costly education process. This problem might justify a government-sponsored education campaign or some type of limited duration disclosure program, but usually not on-going mandatory quality labeling.[59] Other more limited approaches may suffice, however; the recent Kellogg's advertising campaign on its high fiber cereals is an example where a company simply used government-sponsored studies and findings to make its point in what appears to be a successful advertising campaign.

In setting minimum quality standards, policy goes further, either rejecting consumers' ability to understand the information if it were given, or judging the costs of information disclosure to exceed the inefficiencies usually inherent in a minimum quality standard.[60] Minimum quality standards are more appropriate policy choices in cases where quality is particularly difficult to convey to consumers and where there would be little (informed) demand for the level of quality below the minimum standard. This, of course, presumes that liability rules and market devices like reputations and bonds are less efficient in delivering the requisite quality.

In regulatory matters, there has been some movement away from the use of minimum quality standards in favor of information remedies. This is certainly

---

advertisements describing structural factors and engineering features that presumably improve crashworthiness. The National Highway Traffic Safety Commission has created an index for front-end crashes, but this index is often criticized as being very sensitive to precise test specifications. The overall difficulty in measuring this quality does appear to limit effective competition along this dimension.

[58] In the case of cigarette smoking, for instance, there is clear direct and indirect evidence that it took years for many consumers to understand the risks of smoking. See Ippolito and Ippolito (1984).

[59] Contrary to the theory, most cases where the government has actually mandated disclosures have been situations in which clear indices of quality already existed (sometimes developed by the government). EPA mileage ratings, octane rating for gasoline, R-ratings for insulation, tar and nicotine content for cigarettes are all examples. The exceptions have been the general warnings and contraindication information mandated in cases like cigarettes, saccharin and drugs.

[60] Optimal consumer reactions to newly discovered hazards can be quite varied depending on the nature of the risk (Ippolito (1981)). It is quite possible for these reactions to differ substantially from policymakers' views of appropriate reactions to risk information. There are also paternalistic rationales offered for minimum quality standards. I am restricting my attention to efficiency criteria.

of minimum quality standards in favor of information remedies. This is certainly the case in consumer protection activities involving product quality at the FTC. However, minimum quality standards are still the dominant approach in many areas of regulation; for example, they are common in the labor area (e.g., OSHA standards, minimum wage laws, and much of the ERISA pension regulation), in food and drug issues (e.g., the FDA's minimum ingredient requirements for foods as well as the FDA's drug certification program), in the regulation of professionals (e.g., minimum training requirements), and in automobile safety and fuel economy matters (e.g., NHTSA's mandatory passive restraints rulemaking and EPA's maximum fleet mileage rules for auto manufacturers).

A few of these cases (drugs, for example) may meet economic criteria supporting a minimum quality approach to an information problem,[61] but many would appear to reflect forces having little to do with efficiency. Moreover, even in cases where there are plausible economic arguments for minimum quality regulations, the results are often disappointing. Regulation of professional services, for instance, has generally been characterized by minimum requirement regulations on both credentials and on providers' actions, and there may be a theoretical basis for such an approach. Yet many of these strictures have come under strong criticism from economists in recent years.[62] While the focus of this criticism has been the removal or relaxation of the particular standards under study, the broader implications of this body of work may be to reveal the greater susceptibility of minimum quality standards to capture by the regulated industry. The political economy of information remedies has not been explored by the literature to date, and in many cases, this issue may be the most important factor in dictating more efficient outcomes when government regulation is adopted.

III. Overview of the Conference Papers

The volume is organized into four sections that parallel the conference sessions: *Quality Issues*, *Advertising Issues*, *Experimental Studies*, and *Econometric Studies*. The first section begins with the opening remarks of James C. Miller III, Chairman of the Federal Trade Commission, who highlights the Commission's need for more evidence, including more empirical evidence on the magnitude of actual information problems and on the effects of alternative regulatory policies. The two papers in this section are econometric studies of particular markets where information problems are often suspected to occur.

John C. Weicher uses an FTC/HUD-sponsored survey of new home purchasers to investigate whether quality problems in new homes are reflected in house prices. This data set is particularly well-suited to this question since a subsample of the data was verified by professional home inspectors to check for any owner response bias in reporting quality problems. Using both deter-

[61] Though even here the level of the standards has been questioned by many. See Peltzman (1973), for example.

[62] See, for instance, Benham and Benham (1975), Bond et al. (1980), Cady (1976) and FTC (1984).

ministic and stochastic hedonic techniques, the study supports the view that, in the new house market, price does reflect quality, and thus that any information problems must be generally corrected by market mechanisms of some type. Possible mechanisms include builder reputation and the scrutiny of banking intermediaries.

James E. Anderson and Frank M. Gollop study the effects of state warranty laws on the used car market. The study uses an extensive transaction-specific data base that includes prices for more than 120,000 retail sales in 1983 collected by the National Automobile Dealers Association. While the study does not address the ultimate consumer welfare questions behind the laws, it does find that stronger warranty laws increase the price of used cars; preliminary estimates from the study indicate that the major warranty laws add about $200 to the price of the average used car. An additional feature of the study that may be of use to other researchers is its collection of a substantial amount of information about state warranty laws, published as an appendix to the study.

The second section, on *Advertising Issues*, consists of four papers on advertising. The first by Yehuda Kotowitz and Frank Mathewson attempts to formulate testable implications of the Nelson, signalling and persuasive theories of advertising. Using published data from the automobile market of the 1960s and from the Canadian whole life insurance market of the 1960s and 1970s, the authors find that the evidence does not support a pure signalling theory but is consistent with a persuasive theory of advertising. The authors then develop a model of consumer learning where the content of the ads is at issue and show that exaggeration can occur at least during the period when consumers are learning about the quality of the good.

The paper by Timothy Bresnahan looks at advertising as a part of the broader distribution system in which selling services can be provided in a number of ways. Specifically Bresnahan formulates the hypothesis that mass media advertising (whether informative or persuasive) is a substitute for high-service retailing and that the growth of mass media advertising over time reflects relative price changes in these two approaches to distribution. Using cross-industry data and firm-level data from the brewing industry, the study confirms this hypothesis with a slight modification: nationwide media (magazines and network TV) are substitutes for high-service retailing, but local media (spot TV, newspapers and radio) are not. The study also finds that changes in the relative price of (national) mass media advertising and high-service retailing have created economies of scale in distribution, leading to the increased growth of national brands.

The last two papers in this session attempt to examine the effects of the FTC's advertising substantiation program initiated in the early 1970s. Richard Higgins and Fred McChesney consider a political economy explanation for the change in policy, arguing that support for the program derives from large advertisers' ability to gain at the expense of their smaller competitors. They provide preliminary tests of this theory with analysis of stock market data for large advertisers and with an analysis of the incidence of FTC substantiation cases. In both tests, they find support for their view.

. Keith Leffler and Raymond Sauer attempt to examine the effects of the advertising substantiation program on advertising agencies. Their preliminary research reported in this paper finds that substantiation was largely a program of the mid-1970s. Beyond a small positive effect on mid-sized agencies, they find little structural effect of the program. In particular, they find no evidence to support the Higgins-McChesney view of a large firm advantage. An analysis of account retention by agencies who face a substantiation complaint also fails to find any significant effect of the program.

The third section, entitled *Experimental Studies*, includes three studies that use experimental techniques to examine problems in markets with imperfect information. The first paper by Charles Holt and Roger Sherman investigates the use of bundling to mitigate the effects of quality uncertainty. In a model in which individual items of a good have exogenously determined quality, bundling by the seller is shown to be an efficient response under certain market conditions. Experiments under two sets of assumptions show a tendency for both buyers and sellers to prefer bundled goods in cases where this is efficient.

The study by Michael Lynch, Ross Miller, Charles Plott and Russell Porter examines experimental markets where buyers are uninformed about quality, but sellers are not. The study shows that "lemons" equilibria are easily produced in this experimental setting. A variety of mechanisms to resolve the quality degeneration are also examined. Among the major results are: truthful, voluntary disclosure and enforced warranties both lead to full information equilibria; identification of the seller (to allow for one type of reputation) improved market performance but not to full information levels — perhaps because of two externalities: first, "cheating" in the market with seller identification tended to taint other sellers in addition to the seller who actually cheated, and second, sellers seemed to be able to "free ride" on high quality markets developed by pioneers.

Thomas Palfrey and Thomas Romer study a market with uncertain quality in which disputes over warranty obligations can arise. In this setting, the rules of a dispute resolution mechanism can change the equilibrium outcome in the product market itself. Using such an experimental market, Palfrey and Romer report evidence that the allocation of the cost of the mechanism affects the market price through effects on both the buyer and the seller. Despite the relatively complex nature of this experimental market, the "rational" Bayesian equilibrium model predicts prices very well when compared to a model with more myopic or limited consumer understanding.

The final section of the volume, *Econometric Studies*, contains three empirical studies that deal with policy or behavior in markets with potential information problems. Gregg Jarrell and Sam Peltzman examine the effects of product recalls on the stock market value of firms and their competitors. This provocative study finds that for both drug and auto recalls, the reduction in stock value is surprisingly large compared to plausible estimates of the observable direct and indirect costs to the firm. For drug recalls, for instance, the mean cumulative excess return (two week interval) is -6.13%. The study also finds evidence of negative spillovers of recalls to other firms in the industry. Reputational damage from the recall seems an unlikely explanation for this large stock market reaction: evidence from the study suggests that the probability of future recalls is not affected by a past recall, and where it could be tested in autos, that the reduction in sales is too small and too temporary to be consistent with large reputational damage. It is unclear whether this implies that there is some other hidden cost to recalls, some basic problem with the "event study" methodology, or problems with the efficient market hypothesis.

The study of the retail coffee market by Michael Katz and Carl Shapiro examines a variety of ways in which consumers and firms respond to costly information, including the firms' use of coupons, in-store promotions, price ads in local print media, non-price ads on spot and network TV and trial sizes, and consumers' responses to these promotional techniques. Such a detailed study was feasible, in large part, because the authors had access to cross-section, time series data on individuals' coffee purchases collected from grocery stores with electronic scanners. The study finds that price ads, which reduce the cost of acquiring information, increase the consumers' response to both price increases and decreases; that the evidence on coupon usage is consistent with both price cutting that limits stockpiling by regular customers and price discrimination; and that in-store displays are taken as a signal of price cuts by consumers, but that this signal is only weakly accurate. Overall, the evidence here demonstrates that in markets where consumer shopping and dealer promotion are important, aggregate demand estimates that do not account for key nonprice variables may be seriously misspecified. Moreover, this study illustrates a case where information about price is a more significant and difficult issue than information about quality.

The final paper by Robert Porter is an econometric study of the cigarette market from 1947 to 1982 based on aggregate time series data. The study examines the effects of the health information shocks and related government policies on the cigarette market. Unlike previous studies that focused exclusively on demand reactions, Porter attempts to determine the effects of the information on manufacturers' decisions by estimating a simultaneous equations model for the market during this period. Overall, he finds that advertising and pricing decisions were probably not materially affected by the policies; the exception might be the effects of the FCC advertising ban, which is estimated to have reduced advertising's effectiveness and thus to have increased prices by 3-6%.

IV. Conclusion

As illustrated by the wide range of issues covered in this review and by the conference papers themselves, consumer protection regulation raises a great many economic questions for which we currently have only limited answers. The theoretical developments in the economics of information area and the emerging micro data sets of purchasing behavior (generated by electronic scanners, for instance) present opportunities for a broad agenda of research. To those interested in policy, the most pressing research topics are a better theoretic and empirical understanding of the market's ability to resolve information problems, and more insight into the actual consequences of regulation in this area. In sponsoring this conference, the FTC hoped to interest other economic researchers in the problems at the core of consumer protection policy and to stimulate them to conduct further research, especially empirical and factual research, on the results of current regulatory policies.

References

Akerlof, George A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, August 1970, 84, 488-500.

Albion, Mark S., *Advertising's Hidden Effects*, Auburn House, Boston, 1983.

Allen, Franklin, "Reputation and Product Quality," *Rand Journal of Economics*, Autumn 1984, 15, 311-327.

Arrow, Kenneth J., "Limited Knowledge and Economic Analysis," *American Economic Review*, March 1974, 64, 1-10.

Averitt, Neil W., "The Meaning of 'Unfair Acts or Practices' in Section 5 of the Federal Trade Commission Act," *Georgetown Law Journal*, October 1981, 70, 225-295.

Azariadis, Costas, and Stiglitz, Joseph E., "Implicit Contracts and Fixed Price Equilibria," *Quarterly Journal of Economics*, Supplement 1983, 98, 1-22.

Barzel, Yoram, "Some Fallacies in the Interpretation of Information Costs," *Journal of Law and Economics*, October 1977, 20, 291-307.

Beales, Howard; Craswell, Richard; and Salop, Steven, "The Efficient Regulation of Consumer Information," *Journal of Law and Economics*, December 1981, 24, 491-539.

Benham, Lee, and Benham, Alexandra, "Regulating Through the Professions: A Perspective on Information Control," *Journal of Law and Economics*, October 1975, 18, 421-447.

Bond, Eric W., "A Direct Test of the 'Lemons' Model: The Market for Used Pickup Trucks," *American Economic Review*, September 1982, 72, 836-840.

Bond, Ronald S., et al., *Effects of Restrictions On Advertising and Commercial Practices In the Professions: The Case of Optometry*, Staff Report to the Federal Trade Commission, Washington, D.C., September 1980.

Brown, John Prather, "Towards an Economic Theory of Liability," *Journal of Legal Studies*, June 1973, 2, 323-349.

Butters, Gerard R., "Equilibrium Distribution of Sales and Advertising Prices," *Review of Economic Studies*, October 1977, 44, 465-491.

Cady, John, *Restricted Advertising and Competition: The Case of Retail Drugs*, American Enterprise Institute, Washington, D.C., March 1976.

Calfee, John E., and Ford, Gary T., "The FTC's Product Defects Program and Consumer Perceptions of Product Quality," in *Perceived Quality: How Consumers View Stores and Merchandise*, Jacob Jacoby and Jerry C. Olsen (Eds.), Lexington Books, Lexington, Mass., 1985.

Carmichael, H. Lorne, "Reputations in the Labor Market," *American Economic Review*, September 1984, 74, 713-725.

Comanor, William S., and Wilson, Thomas A., "Advertising and Competition: A Survey," *Journal of Economic Literature*, June 1979, 17, 453-476.

Courville, Leon, and Hausman, Warren H., "Warranty Scope and Reliability Under Imperfect Information and Alternative Market Structures," *Journal of Business*, July 1979, 52, 361-378.

Craswell, Richard, "Interpreting Deceptive Advertising," University of Southern California Law Center Working Paper, July 1984.

Darby, Michael A., and Karni, Edi, "Free Competition and the Optimal Amount of Fraud," *Journal of Law and Economics*, April 1973, 16, 67-88.

Eisenach, Jeffrey A.; Higgins, Richard S.; and Shugart, William F., "Warranties, Tie-Ins, and Efficient Insurance Contracts," *Research In Law and Economics*, 1984, 6 167-185.

Epple, Dennis, and Raviv, Arthur, "Product Safety: Liability Rules, Market Structure, and Imperfect Information," *American Economic Review*, March 1978, 68, 80-95.

Federal Trade Commission, *Standards and Certification Proposed Rule and Staff Report*, Bureau of Consumer Protection, Washington, D.C., 1978.

Federal Trade Commission, *Consumer Information Remedies Policy*, Washington, D.C., June 1979.

Ford, Gary T., and Calfee, John E., "The FTC's 1983 Policy Statement on Deception," Unpublished paper, November 1984.

Grady, Mark F., "Regulating Information: Advertising Overview," in *The Federal Trade Commission Since 1970*, Kenneth W. Clarkson and Timothy J. Muris (Eds.), Cambridge University Press, Cambridge, 1981.

Grossman, Sanford J. "The Informational Role of Warranties and Private Disclosure About Product Quality," *Journal of Law and Economics*, December 1981, 24, 461-483.

Hemenway, David, *Industrywide Voluntary Product Standards*, Ballinger, Cambridge, MA, 1975.

Hirshleifer, J., "Where Are We in the Theory of Information," *American Economic Review*, May 1973, 63, 31-39.

Hirshleifer, J., and Riley, John G., "The Analytics of Uncertainty and Information — An Expository Survey," *Journal of Economic Literature*, December 1979, 17, 1375-1421.

Ippolito, Pauline M., "Information and the Life Cycle Consumption of Hazardous Goods," *Economic Inquiry*, October 1981, 19, 529-558.

Ippolito, Pauline M., and Ippolito, Richard A., "Measuring the Value of Life from Consumer Reaction to New Information," *Journal of Public Economics*, November 1984, 25, 53-81.

Jacobs, William W., et al., *Improving Consumer Access to Legal Services: The Case for Removing Restrictions on Truthful Advertising*, Staff Report, Federal Trade Commission, Washington, D.C., November 1984.

Jordan, Ellen R., and Rubin, Paul H., "An Economic Analysis of the Law of False Advertising," *Journal of Legal Studies*, June 1979, 8, 527-553.

Kihlstrom, Richard E., and Riordan, Michael H., "Advertising as a Signal," *Journal of Political Economy*, June 1984, 92, 427-450.

Klein, Benjamin, and Leffler, Keith B., "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, August 1981, 89, 615-641.

Kornhauser, Lewis A., "Unconscionability in Standard Forms," *California Law Review*, September 1976, 64, 1151-1183.

Lacko, James M., *Product Quality and Information in the Used Car Market*, Staff Report to the Federal Trade Commission, Washington, D.C., forthcoming.

Leland, Hayne, "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," *Journal of Political Economy*, December 1979, 87, 1328-1346.

Lynch, Michael P., "A Dual Stage Model of Retailing and Manufacturing," Unpublished paper, January 1983.

Masson, Alison, and Steiner, Robert, *The Effects of State Drug Product Selection Laws*, Staff Report to the Federal Trade Commission, Washington, D.C., forthcoming.

Matthews, Steven, and Moore, John, "Monopoly Provision of Product Quality and Warranties," Unpublished paper, September 1983.

Milgrom, Paul, and Roberts, John, "Price and Advertising Signals of New Product Quality," Unpublished paper, April 1984.

Nelson, Phillip, "Information and Consumer Behavior," *Journal of Political Economy*, March/April 1970, 78, 311-329.

Nelson, Phillip, "Advertising as Information," *Journal of Political Economy*, July-August 1974, 81, 729-754.

Newbery, David M., and Stiglitz, Joseph E., "Wage Rigidity, Implicit Contracts, and Economic Efficiency: Are Market Wages Too Flexible?" Draft, June 1983.

Overstreet, Thomas R., *Resale Price Maintenance: Economic Theories and Empirical Evidence*, Staff Report, Federal Trade Commission, November 1983.

Palfrey, Thomas, and Romer, Thomas, "Warranties, Performance, and the Resolution of Buyer-Seller Disputes," Carnegie-Mellon University, GSIA Working Paper, No. 32-81-82, July 1982.

Peltzman, Sam, "An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments," *Journal of Political Economy*, September/October 1973, 81, 1049-1091.

Pitofsky, Robert, "Mandated Disclosure in the Advertising of Consumer Products," in *Business Disclosure: Government's Need to Know*, Harvey Goldschmid (Ed.), McGraw-Hill, New York, 1979, 311-330.

Polinsky, A. Mitchell, "Strict Liability vs. Negligence in a Market Setting," *American Economic Review*, May 1980, 70, 363-367.

Polinsky, A. Mitchell, and Rogerson, William P., "Products Liability, Consumer Misperceptions, and Market Power," *Bell Journal of Economics*, Autumn 1983, 14, 581-589.

Posner, Richard A., *Economics of the Law*, Little, Brown and Co., Boston, 1972, 50-53.

Posner, Richard A., *Regulation of Advertising by the FTC*, American Enterprise Institute, Washington, D.C., 1973.

Posner, Richard A., "The Federal Trade Commission's Mandated-Disclosure Program: A Critical Analysis," in *Business Disclosure: Government's Need to Know*, Harvey Goldschmid (Ed.), McGraw-Hill, New York, 1979, 331-359.

Priest, George L., "A Theory of Consumer Product Warranty," *Yale Law Journal*, May 1981, 90, 1297-1351.

Rothschild, Michael, "Models of Market Organization with Imperfect Information: A Survey," *Journal of Political Economy*, November/December 1973, 81, 1283-1308.

Rothschild, Michael, and Stiglitz, Joseph, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, November 1976, 90, 629-650.

Salop, Steven, "Information and Monopolistic Competition," *American Economic Review*, May 1976, 66, 240-245.

Salop, Steven, "Parables of Information Transmission in Markets," in *The Effects of Information on Consumer and Market Behavior*, Andrew W. Mitchell (Ed.), American Marketing Association, Chicago, 1978.

Scitovsky, Tibor, "Ignorance as a Source of Oligopoly Power," *American Economic Review*, May 1950, 40, 48-53.

Shapiro, Carl, "Consumer Information, Product Quality and Seller Reputation," *Bell Journal of Economics*, Spring 1982, 13, 20-35.

Shapiro, Carl, "Premiums for High Quality Products as Returns to Reputation," *Quarterly Journal of Economics*, November 1983, 98, 659-679.

Shavell, Steven, "Strict Liability Versus Negligence," *Journal of Legal Studies*, January 1980, 9, 1-25.

Smallwood, Dennis E., and Conlisk, John, "Product Quality in Markets Where Consumers Are Imperfectly Informed," *Quarterly Journal of Economics*, February 1979, 93, 1-23.

Smith, Adam, *An Inquiry into the Nature and Causes of the Wealth of Nations*, The Modern Library, New York, 1937.

Spence, Michael, "Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution," *Journal of Economic Theory*, March 1974, 7, 296-332.

Spence, Michael, "Informational Aspects of Market Structures: An Introduction," *Quarterly Journal of Economics*, November 1976, 90, 591-597.

Spence, Michael, "Consumer Misperceptions, Product Failure and Producer Liability," *Review of Economic Studies*, October 1977, 44, 561-572.

Steiner, Robert L., "Basic Relationships in Consumer Goods Industries," in

Research in Marketing, 7, JAI Press, 1984.

Stigler, George J., "The Economics of Information," *Journal of Political Economy*, June 1961, 69, 213-225.

Stiglitz, J. E., "Equilibrium in Product Markets with Imperfect Information," *American Economic Review*, May 1979, 69, 339-345.

Wilde, Louis L., "On the Formal Theory of Inspection and Evaluation in Product Markets," *Econometrica*, July 1980, 48, 1265-1279.

Opening Remarks

*James C. Miller III, Chairman**
Federal Trade Commission

On behalf of the Federal Trade Commission — welcome to the Conference of the Bureau of Economics on Empirical Approaches to Consumer Protection Issues. As the first economist to chair this Commission I am especially interested in the outcome of a conference emphasizing empiricism.

You know, we economists have much to be proud about; but we also have much to make us humble. It is said of us, for example, that we have predicted ten of the last five recessions — or worse — that an economist is the only one who doesn't know why he is standing in the breadline. Yet, each of us, I am sure, has enough friends in the so-called "hard sciences" who can relate their own horror stories of predictions — and theories — gone awry. And, like it or not, today all scientists in every field swear by the empirical method — at least if they don't, they should!

What many don't know, moreover, is that the application of scientific testing to public policy issues began with Adam Smith and David Hume — two great economist/philosophers of the 18th century. Indeed, one of Smith's and Hume's greatest contributions to social progress may have been this substitution of the humility of skepticism for the arrogance of certitude — and in the physical as well as the social sciences. In so doing, they spawned a veritable explosion of knowledge — not only in economic and political science — but in every major field.

Darwin's theory of evolution, for example, was influenced by Smith's theory of market equilibrium. But perhaps it wasn't until Einstein's theories displaced Newtonian physics that the revolution begun by Hume and his good friend Smith was complete — that scientists would no longer take anything for granted in their search for knowledge.

In any event, I think we're at the same sort of juncture — though perhaps not quite as cosmic — in consumer protection theory. The accepted theory is that consumers are limited — often extremely so — in their ability to judge sellers' claims. But we're not quite sure which kinds of judgments are more or less likely to be correct and under what conditions.

For example, where there is clear-cut deception or fraud, everyone will agree that a sizable group of consumers may be misled to their detriment. And where there is clear scientific validation for an express claim — such as Diet Pepsi has only one calorie — few will contend any deception is involved, whether or not some consumers choose to believe that consuming Diet Pepsi can turn them into a model of physical fitness.

But it's in the borderline cases where it's not always immediately obvious

*The views expressed are those of former Chairman Miller (1981-1985). They do not necessarily reflect the views of the other Commissioners.

what the advertiser intended, how the claim should be substantiated, and, ultimately, what impact it might have on consumers, that reasonable commentators sometimes become confused and skeptical about their own ability to determine whether consumers will be misled. Let me cite just one example from a galaxy far, far away in a time long, long ago.

In that galaxy on a planet called *Terra* in the year 1942 a popular magazine, called *Reader's Digest*, published an article purporting to show scientific evidence that different brands of cigarettes have different levels of tar and nicotine content. The brand with the lowest level of tar and nicotine, "Old Gold," subsequently advertised that fact, but did not include the *Reader's Digest* statement that:

> The laboratory's general conclusion will be sad news for the advertising copyrighters but good news for the smoker, who need no longer worry as to which cigarette can most effectively nail down his coffin. For one nail is just about as good as another. Says the laboratory report: "The differences between brand [sic] are, practically speaking, small and no single brand is so superior to its competitors as to justify its selection on the ground that it is less harmful."...

In that country, there was an agency called Federal Trade Commission — and it found the omission deceptive and halted the brand's advertisements of the test results; or so the story goes. The agency's theory of deception was if Old Gold were to report the tar and nicotine figures it must also report the conclusion of the magazine article and the testing laboratory that any differences among the brands were small and insignificant.

The difference at that time between Old Gold at 2.04 mg. and the worst brand at 3.02 mg. was 0.98 mg. of nicotine. By contrast, back in the present era, in 1981 over half of the 200 brands tested at our FTC had nicotine content under 0.98 mg.

I'm not a physician or medical scientist, so I cannot say whether in this specific context 0.98 mg. is significantly different from Old Gold's 2.04 mg., or whether the two tests used different metrics that are simply not comparable. But I do find it troubling that one Commission found one set of differences highly useful and informative and another Commission found another, seemingly similar, set of differences highly harmful and deceptive.

I also find it troubling that our own agency established a set of guidelines in the mid-1950s which expressly forbade any mention of the health effects of cigarette smoking and any representation of the nicotine, tar, acid, resin or any other substance "when it has not been established by competent scientific proof applicable at the time of dissemination that the claim is true, and if true, that such difference or differences are significant."

Now it is well and good to require documentation of scientific claims. But the blanket prohibition against representations of tar and nicotine content seems to imply a very low respect for the ability of consumers to interpret information. What's more, it is difficult to believe that even if the variation of tar and nicotine content among cigarette brands was insignificant in the early years, that publication of test results would not have stimulated competition among cigarette companies to reduce their levels of tar and nicotine; and that the reduced levels would not have been significantly less hazardous; and that such competition would have occurred long before the Commission acted in 1966 to standardize test procedures and allow disclosure of information from those tests.

In other words, I believe the Commission inadvertently suppressed the development of a useful market for the dissemination of information whose knowledge might have stimulated considerable demand for and competition to produce less hazardous cigarettes. And I believe that this action might not have occurred if that earlier Commission had known a little more about how consumers respond to the dissemination of information that is subject to alternative interpretations.

I raise this example not to criticize decisions of a Commission under both Republican and Democratic, liberal and conservative, administrations in a bygone era. But to point up the clear need for continued research into how consumers evaluate the information they receive from sellers. And to express the fervent hope that what you are doing here today and tommorow will at least push us in the right direction — if not launch a valuable, on-going inquiry into the theory of how we should think about consumer protection issues, into the ***epistemology***, if you will, of consumer protection.

Let me say also that with the likes of Sam Peltzman in your midst, I can't help but be optimistic.

Again, welcome and good luck!

The Market for Housing Quality

John C. Weicher
American Enterprise Institute for Public Policy Research

Introduction

This paper analyzes the relationship between the quality of new home construction and the prices paid by new home buyers. In recent years there has been substantial and increasing public concern with the problems faced by buyers whose homes prove to have defects of various kinds, and several government agencies, including the Federal Trade Commission, have investigated the problem. While there has been very little research, most analysts of product liability and quality issues have regarded housing as one of the clearest examples of a good about which buyers have particularly little information before purchase, and therefore the housing market would be one where regulation would presumably be most beneficial to buyers, and contribute to a well-functioning market (*e.g.*, McKean [1970], Leland [1979]).

In response to the perceived problem, private firms have begun to offer warranties and inspection services for existing homes, and warranties for new homes. Surprisingly, these services have not found especially wide acceptance in the market, even though the purchase of a home is the largest financial decision made by the typical consumer. Quality problems appear to be more serious for existing homes than new ones, but only about 5 percent of existing homes are sold with a warranty against defects. In the new home market, the Home Owners Warranty program, established by the National Association of Home Builders in 1972, has gradually grown until it covers about 25 percent of new homes produced each year, but this has come only after heavy promotional activity by NAHB. This experience suggests that there is room for research into the importance of quality problems in the new home market.

The Relationship Between Prices and Repairs

The main purpose of this paper is to test the hypothesis that the new home market adjusts to the existence of defects in the homes, so that buyers of homes with problems pay less for them as a result. If they do, then the lower price suggests that the market works even when the individual buyer does not have full information about product quality. If they do not, then the market does not work efficiently in the absence of full information among home buyers, and some sort of mechanism to provide information or protection against problems may be appropriate.

Specifically, I examine the relationship between new home price and repair costs incurred by homebuyers after they buy the home and move in. I also investigate more briefly the relationship between price and two of the more common problems encountered in homes. The data source is the most extensive data set available concerning quality problems in new homes: a survey of recent buyers conducted by Mathematica Policy Research (MPR) for the Federal Trade Commission (FTC). This survey contains data on some 90 different defects, including descriptions of the problem, the process by which it was remedied (if it was), the cost of repairing it, and who bore the cost.[1]

Differences from Previous Work

The project builds on a previous project in which I studied new home defects and house prices using the Annual Housing Survey (Weicher [1981]). That study found that there were very few defects in new homes, and that those which did exist did not usually affect the owner's estimated value of a new home, his or her evaluation of its quality, actual repairs, or planned repairs. These findings suggested that popular perceptions of the problem are overstated. However, there were several limitations which can be addressed in the present study:

1. Few categories of defects were included in the AHS.
2. Owners were asked only about their estimate of the value of the home, not the price they paid. These are likely to be very close for recent home buyers, but the discovery of defects could have led owners to lower their subjective estimates of value, below the price actually paid.
3. Owners were asked only about problems they were experiencing at the time of the survey, not about any that had been discovered and already resolved.
4. The AHS reports value only within broad categories of at least $5,000. Many defects can be remedied at much lower cost, and thus may not affect price or value enough to change the estimate measurably.
5. Households were not asked about repair costs, but only whether repairs had occurred or were planned.

Another important difference between this paper and the earlier work is that I treat problems as stochastic phenomena; previously they were treated deterministically. That is, previously I assumed that the market reacted only to the existence of a defect in a particular house; now I investigate the possibility that it reacts to the probability of a defect in houses of a particular quality level.

Organization of the Paper

The next section briefly describes the recent history of private and public consumer protection activity. I then discuss the data source. The main body of the paper reports on the econometric analysis, which takes the form of hedonic estimation of new home prices as a function of the attributes of the homes. I

[1] This survey is described in detail, and the results summarized, in Kaluzny (1979).

first describe the variables included in the analysis, and then report the results, stressing the repair cost variables, of the deterministic analysis. Then I explain the technique used to construct the probabilistic measures of repair cost, and report the results of the econometric analysis in which they are used.

The Development of Protection for Homebuyers

Private Programs

Protection for home buyers against defects began to become available, apparently for the first time, during the 1970s. Around 1972, several firms began to issue insurance or to warrant existing homes against specific defects at the time of sale. By 1977, there were about a dozen firms in the industry.[2] Each of them provided coverage for a somewhat different range of problems, such as basic appliances, mechanical systems, or basic structural components, at premiums ranging from $140 to $300 on a $70,000 house, or about $2 to $4 per $1,000; rates generally were inversely related to house value. The individual firms are small, and the industry as a whole provided coverage for less then 5 percent of the existing homes sold annually. A trade association has been formed, the National Home Warranty Association.

In 1974, the National Association of Realtors promulgated standards for warranties which it would recommend to its members. Eleven firms were approved by NAR as of April 1980, at which point it ceased to recommend specific programs.

The process of developing a warranty program for new homes was also undertaken by a trade association, the National Association of Home Builders (NAHB), in 1972, with the impetus coming in large part from the experience of a similar program previously established in Great Britain. The first homes were insured under this Home Owners Warranty (HOW) program in 1974; as of December 1983, some 1.1 million were covered, with 250,000 being added within the preceding year.[3] This represented about 25 percent of the homes started during the year.

As these figures indicate, the program has been voluntary. Builders choose whether to apply for coverage; their application is reviewed by a local HOW Council, itself consisting of builders. A participating builder, however, must enroll all houses being built within the territorial jurisdiction of the local council, including town houses and low-rise condominiums (with four or fewer stories). About 12,000 builders participate in the program.

However, the voluntary nature of HOW may be changing. In 1978, New Jersey passed a law requiring each builder in the state to participate in HOW, or some similar state-approved program. Minnesota has also required builders to offer protection, but less specifically patterned after HOW.

[2] A description of the existing home warranty industry is contained in Brewster et al. (1977).

[3] Conversation with Jane Snow of the Home Owners Warranty Corporation, April 9, 1984. For a general description of HOW, *see* Snow (1979).

The program provides coverage for one year on workmanship and materials, for two years on mechanical and structural systems, and for ten years on major structural defects. It insures the builder's own warranty of the house. If the builder fails to satisfy a buyer's claim—or if the builder has gone bankrupt—the warranty comes into force. HOW first appoints a conciliator at the request of the buyer; it states that 84 percent of complaints have been resolved successfully at this stage. If the conciliator cannot persuade seller and buyer to agree, either of two procedures may be used: the buyer can request arbitration, which is binding on the builder; or the conciliator can himself become the arbitrator to resolve the remaining disagreements. The latter procedure is being used increasingly, to expedite settlement.

Through December 1983, HOW has incurred $94 million in claims. Of this, $51 million has been for major structural defects in the third through the tenth year of coverage; the remainder is for items covered during the first and second years, which may include structural defects if the builder has failed to perform on his own warranty. Some 6,000 claims for structural defects have been incurred, at an average cost of about $8500. The premium charged was originally $2 per $1,000. When this failed to cover claims, HOW adopted a sliding scale of discounts and surcharges varying by location, and now has established a system of differential premiums for individual builders, based on the builder's claim experience with HOW.

The Government Role

While there has been some federal government concern with housing defects since the mid-1950s, it has accelerated within the last nine years, at the same time that private programs have developed and expanded. Until very recently, federal actions were limited to homes where the mortgages were insured by the Federal Housing Administration. In 1964, FHA was authorized to pay the cost of repairing substantial defects in new FHA-insured homes, provided the claims were made within four years;[4] in 1970, similar protection was extended to buyers of existing homes insured under certain programs between 1968 and 1972, if the claim was made by 1976.[5] Few claims have been paid under either statute.

Congress began to consider a possible expansion of the federal role in 1976, when it directed the Department of Housing and Urban Development to study the need for a national home inspection and warranty system, for conventional as well as FHA-insured homes, to be operated by the government. This study, conducted by Mathematica Policy Research, concluded that demand would be

---

[4] Section 518(a) of the National Housing Act of 1937, enacted as P.L. 88-560, Housing Act of 1964, Sec. 121.

[5] Section 518(b) of the National Housing Act of 1937, enacted as P.L. 91-609, Housing and Urban Development Act of 1970, Sec. 104. This section applies only to homes insured under Section 235 (*i.e.*, subsidized homes for low-income families), and to those insured under Section 203 and 221 and "located in an older, declining urban area."

low for any of several kinds of home inspection and warranty programs, and that voluntary programs for FHA-insured buyers were not feasible without subsidy (Brewster, et al. [1977]).

At about the same time, the Federal Trade Commission began to be interested in new home defects, under the authority of the Magnuson-Moss Warranty-Federal Trade Commission Improvement Act of 1975, which extended its jurisdiction to matters affecting, as well as in, interstate commerce. In 1978, "housing-related issues" occupied 6 percent of total FTC attorney staff time, or between 40 and 50 attorney man-years.[6]

The Commission has proceeded on the basis of case-by-case enforcement, rather than rule-making. Its policy can perhaps best be illustrated by a speech given by retiring Commissioner Elizabeth Hanford Dole in January, 1979, to the NAHB convention. She urged the NAHB and the building industry generally to expand and strengthen the HOW program as the most desirable means of meeting the growing public demand for protection from quality problems in new homes. She also indicated that the FTC might decide to devote additional resources to specific instances of companies "engaging in especially disreputable homebuilding practices" (Dole, 1979).

The Commission's first case, against Kaufman and Broad of Chicago, concerned non-performance on the builder's warranty and was settled in 1978 by a consent decree in which Kaufman and Broad agreed to join HOW, as well as make repairs. In 1979, the FTC established a new home construction defects enforcement program and has since pursued a number of other cases. The only current case in which the FTC has entered a complaint is against Ward Homes of the Washington, D.C. area; it also involves non-performance on the builder's warranty. This case has very recently been withdrawn from adjudication, as the FTC staff and the company have agreed on the terms of a consent decree. The Commission is also pursuing other cases, on which complaints have not been filed; some have been closed without proceeding. Apparently the Commission has never seriously considered rule-making, such as requiring all builders to join HOW or establish similar programs, as New Jersey and Minnesota have done.

Data

The data source for this study is a survey conducted in the fall of 1979 by Mathematica Policy Research, Inc., under contract to the FTC and HUD.[7] This survey sought to ascertain the experience of recent new home buyers with defects in their homes: how many and what sorts of defects occurred, if and how the defects were repaired, how much it cost to repair them, and who paid for the repairs.

---

[6] For a discussion of federal involvement, with emphasis on the FTC, *see* Stanton (1979).

[7] For a full description of the survey, *see* Kaluzny (1979), chapters I and V and appendices A and C.

The study had two components. The first was a telephone survey of 1,812 households that had bought new homes between April 1977 and October 1978, between 12 and 30 months prior to the survey. The second was a subsequent inspection of the homes for 299 households participating in the survey. Inspectors from Certified Homes Corporation, Inc. (CHC) were sent to these homes. They were asked to verify the problems reported by the homebuyer, and estimate the cost of repairing those in which they concurred with the owner that a problem existed. They were also asked to look for additional problems which the homeowners had not found, and estimate the cost of repairing them. CHC is a large home inspection firm, located in Columbia, Md.

The smaller sample is used in the present study because of the greater information it contains about repair costs. It also has more information about the location of the home, which was provided by MPR separately to the FTC.

The telephone survey was a national sample of new home buyers that corresponds fairly closely to the geographic distribution of new home permits by Census region and division, with some underrepresentation of the West region. The subsequent field inspection included only homes in 28 states and the District of Columbia, nearly all of them being states in which CHC had inspectors. This smaller sample, as used in the present paper, overrepresents the East North Central and Mid-Atlantic Census divisions (from New York and New Jersey to Wisconsin and Illinois), as well as underrepresenting the West. The mean sales price is close for the two samples; both are about 10 percent higher than the national average sales price for 1977 and 1978 combined. The sample used in this paper also consists of somewhat larger houses than the national average, as reported by the Census Bureau, perhaps because I excluded town houses from the sample.

The MPR survey reports about two defects per home, on average. This is many more than HOW (two-thirds of a defect) or the AHS sample used in my previous paper (one-third of a defect). The differences between the MPR and AHS surveys may occur partly because the former is explicitly designed to collect information on defects, and does not limit the range of responses; some 90 categories are listed in the MPR report. The AHS includes ten specific questions about housing conditions that could be regarded as building defects. But this does not account for all of the difference. There are two defect questions which are identical (leaky roofs and wet basements), and in both cases MPR reports a much higher incidence: 6.1 compared to 2.5 percent for roofs, and 32.1 compared to 18.2 percent for basements. Both proportions are different at much better than the 0.1 percent level of statistical significance.

Households in both surveys are also asked to rate the overall quality of their house. Those participating in the MPR survey tend to rate their homes lower; the mean rating was 4.04 on a scale of 5, just above "somewhat satisfied," while AHS respondents gave an average rating of 3.66 on a scale of 4, closer to "excellent" than to "good."

There are a number of possible explanations for these differences between the two surveys, which I do not attempt to pursue in this paper.[8]

Hedonic Analysis of New Home Prices

Specification of Functional Form

This section describes the basic hedonic regression equation used to estimate the relationship between new home price, measures of quality, and other attributes. Although participants in this conference are probably not primarily interested in the hedonic estimation of home prices, nonetheless it may be useful to begin with a brief enumeration of the variables included in the analysis, other than quality and repair cost.

The form of the hedonic equation is:

$$\ln P = \sum_{i=1}^{n} b_i X_i + \sum_{j=1}^{m} b_j D_j \tag{1}$$

where lnP is the natural logarithm price paid for the house, adjusted to 1978 dollars, the X_i are the attributes of the house (such as measures of size, amenities, and location), and the D_j are measures of defects experienced, or costs of repairing the defect. Most of the attributes, and the defect measures, take the form of dummy variables; the size and repair cost variables are measured continuously.

The semi-logarithmic form of equation (1) was chosen because it has proven to be the most appropriate in previous analyses of the Annual House Survey and other housing microdata sets.[9] Others which were tested include the linear form, which did not work as well; the coefficient of determination was significantly smaller and fewer variables were significant. The semi-logarithmic form implies that the dummy variables, including the defects, have the same percentage effects on the prices of all houses; that is, they have a greater dollar impact on more expensive homes. It also implies that each additional dollar of repair cost has a slightly smaller effect on house price, but in the range of house prices and repair costs in this data set, the deviation from a linear relationship is negligible.

Independent Variables: House Attributes

The characteristics of houses included in the analysis can conveniently be grouped into eight categories. These are:

(1) Measures of size: number of rooms, number of bathrooms, square feet of floor space in the house, presence of a basement.

---

[8] For extensive discussion of the differences, *see* Weicher (1981), pp. 373-374.

[9] *See* Follain and Malpezzi (1978) for an extended discussion of the choice of functional form.

(2) Systemic attributes: type of heat (electric, gas, oil, "other," or none), presence of central air conditioning, connection with public sewers or septic tanks, presence of gutters and downspouts.

(3) Structure type: basic layout (one floor, two or more floors, or split level), exterior building material (brick, aluminium siding, concrete, wood or "other," which appears to be mostly stucco).

(4) Location: Census Division (nine categories), 1977 population of the Standard Metropolitan Statistical Area (SMSA) in which the house is located, distance from the centroid of the Zip Code in which the house is located to the center of the Central Business District of its SMSA.

(5) Date of purchase, measured as the month of closing, between April 1977 and October 1978.

(6) Terms of sale: cash purchase or mortgage type (FHA, VA, conventional or other).[10]

(7) Tract: whether the home is located in a tract.

(8) Warranty: whether the home carries a warranty and if so, the type (Home Owners' Warranty—HOW—of the National Association of Home Builders, builder's warranty, seller's/agent's warranty, other, or none).

Compared to other microdata sets, the FTC survey has some strengths and weaknesses which may be worth mentioning. Notably, it includes a more specific measure of house size than is customarily available, namely the number of square feet of space within the house. This is so rarely available in other data sources that it has been the subject of some concern and analysis (Ozanne, et al. [1979]). Other characteristics which are not available in the AHS, and were therefore excluded from my previous analysis, include presence of gutters and downspouts, presence of heat pump, all of the structure type variables, the location within the SMSA (distance from CBD), and the tract and warranty variables. (The sale terms were available but not used.) The distance variable has been available in microdata sets for individual SMSAs assembled by local researchers.[11]

The most important variable missing from the FTC data, based on its importance in other research, is the presence of a garage or carport. The second most important is the quality of the neighborhood in which the house is located. Others used in my previous work or that of other analysts of the AHS which are sometimes significant and important include the type of furnace (especially built-in wall furnaces in the South and California), the cooking fuel (gas or electricity), and, occasionally, specific neighborhood attributes, such as convenience of shopping facilities.[12] Size of the lot on which the house is located has proven

[10] There were no homes in the data sample reporting mortgage financing through the Farmers Home Administration.

[11] *See* Miller (1982) for a summary of hedonic studies.

[12] *See* for example Follain and Malpezzi (1978), Weicher and Hartzell (1982) and Malpezzi et al. (n.d.). Weicher and Hartzell is the only study analyzing only new home prices, and therefore the one used most to evaluate the importance of omitted variables. The presence of a garage or carport is important in all of these studies.

has proven useful when it was available, but that is not often.

An important difference between the present and the earlier study occurs with respect to measures of location. In my work with the AHS, I used samples in which all of the homes were located in 59 SMSAs, about 20 in each of three samples by year of survey (1974, 1975 or 1976). I then used dummy variables for the specific SMSA, and, where available, location inside or outside of the central city, and the county or group of counties in which the house was located. This approach was dictated by the Census requirement that information not be divulged for any observation within an area of less than 250,000 population.

The MPR data are much closer to a national sample, with few observations in any one SMSA. However, the data include the zip code in which the house is located, and that information has been used by MPR to identify distance from CBD, size of SMSA, and apparently also population of the zip code itself. I have used the location of the house by Census Division, the size of SMSA, and the distance from the CBD, to measure location, rather than the dummy variable technique of my earlier study.

Data

As mentioned previously, the data set used in this analysis is the file of 299 new homes which were inspected by CHC after participating in the telephone survey. However, the regressions include only 170 observations, rather than 299. The reason is that there are a large number of missing values, particularly for several key variables. Some 60 do not report the price of the home; an additional 33 do not report the size of the house in square feet; a further 26 do not have the locational variables, distance to CBD and population size.[13] Thus the sample used in this analysis is significantly smaller than the number of homes inspected by CHC. Kaluzny used almost the entire CHC data set (294 of the 299 observations) in his analysis of defects and repairs, and also weighted the sample to achieve the same distribution as the larger sample of the telephone survey. His weights were based on problem resolution experience and Census division. I have not attempted to weight the observations in my sample. It may be worth noting that mean unweighted repair costs in my sample are close to those reported by Kaluzny: within $50 or 5 percent for problems reported by the owners and resolved by them, and within $10 or 1 percent for all problems verified and discovered by CHC inspectors. Geographically, my sample under-represents that West Census region, and to a lesser extent the South, and over-represents the North Central and to a lesser extent the Northeast, compared to Kaluzny's weighted sample.

Regression Results

Table 1 contains the results of a hedonic regression which omits all measures

[13] This information comes from a separate file on the data tape provided to me by the FTC. The file omits 30 of the original 299 observations.

of defect and repair costs. This regression is presented in order to provide some information on the structure of the basic regression equation. The next section will present results for defect and repair cost variables. Most of the results reported in Table 1 are essentially unaffected when these variables are included; any noteworthy differences will be reported. In terms of equation (1), the results reported here are for a regression of the form

$$\ln P = \sum_{i=1}^{n} b_i X_i, \tag{2}$$

omitting the defect (D_j) variables.

I also experimented with several different specifications of the dependent variable: real dollar price, log of the midpoint of the $5,000 interval in which the real price fell (*e.g.*, if price was reported as $59,000, the dependent variable was coded as ln57,500 = 10.95954), and natural log of the midpoint of the price intervals used by MPR for those observations where the interviewee could not specify the price. Real dollar price proved less satisfactory than the natural log; the coefficient of determination was 6 percentage points lower, and most of the significant variables in Table 1 had smaller and less significant coefficients, including the size variables. The MPR price-interval variable was used because this is the only form of price data available in the AHS, and I wanted to see if the difference in data might explain the pattern of insignificance in my earlier study. Apparently it does have some effect; the coefficient of determination was 9 points lower for the interval-estimate prices, the coefficient of the square foot variable was cut in half, the coefficient of SMSA size was reduced by a third, and the air conditioning variable became insignificant. However, this appears to be an artifact of the large price intervals used; when the natural log of the $5,000 interval midpoint is the dependent variable, the results are virtually identical to those in Table 1.

Most of the coefficients in Table 1 are reasonable, or at least not unreasonable. The most significant variable is square feet of floor space in the house. Others which are conventionally significant include four Census divisions, presence of a heat pump, FHA insurance, and location within a tract development.

The importance of the square footage variable is manifested in two ways: its inclusion raises the coefficient of determination by 20 percentage points, and reduces the significance of the number of bathrooms, and particularly the number of rooms. In the absence of square footage, both room variables are conventionally significant and their coefficients much larger: twice as large for bathrooms, five times as large for rooms.

Most of the other variables take the form of dummies, and their coefficients can be interpreted as rough approximations to the percentage change in price that results from the presence of the characteristic.[14] Gas heat, oil heat, and electric heat with a heat pump have about the same value; electric heat in the absence of a heat pump reduces the price of the home by over 10 percent. (Only two homes report some other heating fuel, and only one reports no heat.) Central air conditioning is significant at the 10 percent level, adding almost 10 percent to the price. Structure type and building material have no effect on price.

The location variables gave a somewhat unexpected pattern. The omitted Census division is the Pacific, consisting of Washington, Oregon, and California; it has the highest prices. New England is nearly as high. The lowest, unexpectedly, occur in the Mid-Atlantic division, comprising New York, New Jersey, and Pennsylvania. Other low-cost areas are the West South Central (Texas, Oklahoma, Louisiana and Arkansas), the South Atlantic (Delaware, Maryland and West Virginia south to Florida), and the West North Central (the northern states between the Mississippi River and the Rockies). This pattern is mostly consistent with previous work, but with one glaring exception: the Mid-Atlantic region is generally one of high home prices. My best explanation for this is that most of the homes in this region are located near New York City, and the SMSA size variable may be picking up higher prices in this region. In my previous work, I also found Philadelphia to have rather low prices.[15] The size variable is significant at the 10 percent level; it implies that prices rise by about $105 for each additional 100,000 residents of the SMSA. The measures of distance from the CBD have the expected signs, and imply a negative-exponential rent gradient, but are clearly insignificant; my judgment is that the insignificance is not surprising. Studies of density gradients across cities find wide differences, which are only partly explained by SMSA size and region (*e.g.*, Muth [1960], Mills [1972]).

The insignificance of the date of sale implies that real home prices did not rise during the survey period. This is not consistent with other evidence. The Census Bureau's "Price Index of New One-Family Houses Sold" rose much faster than the GNP deflator or the CPI between the second quarter of 1977 and the third quarter of 1978; relative to the GNP deflator, for example, it rose by 8.7 percent. Even taking the point estimate in Table 1 at face value, it implies a real price increase of only 0.5 percent during the period.

The fact that FHA and VA homes are worth less than others of similar size and location is consistent with the mandate of these agencies to insure or guarantee lower-priced homes, but I am surprised by the magnitude of the FHA

[14] Halvorsen and Palmquist (1980) have shown that dummy variable coefficients in semi-logarithmic regressions cannot be interpreted as percentage changes. This effect, however, is practically important only for fairly large coefficients, of which there are few in Table 1. For example, the error is less than ½ a percentage point for coefficients up to about .10 in absolute value, less than 1 percentage point up to about .14, and less than 2 percentage points up to about .19. The actual error depends on the estimated house price as well as the individual coefficient.

[15] This assertion is based on the unpublished appendix to my earlier study (Weicher [1981]). Prices in Philadelphia were lower than in all but two SMSAs of the 21 included in the 1975 AHS wave of Selected SMSAs. The only ones lower were Kansas City and Colorado Springs.

coefficient in particular.

Warranties and Home Prices[16]

The coefficients of the home warranty variables are all insignificant. Furthermore, the coefficient for the HOW program is negative, with a value of about 5 percent of sales price. This is not significantly different from, but also not very close to, the rate of 0.2 percent that was actually being charged during 1977 and 1978 by HOW. Builders' warranties are worth literally nothing, compared to no warranty at all; in some regressions, not reported in Table 1, this coefficient turns very slightly negative. "Other" warranties do have a positive and rather large coefficient, but only 4 of the homeowners in the sample reported having such warranties. Taken at face value, these results do not suggest that there is much market demand for new home warranties.

There are several possible explanations for the results. In 1981 and 1982 the homes that HOW insured were less expensive than average, so they may have been of lower quality. If this was also true in 1977 and 1978 the coefficient may therefore simply measure a quality difference associated with participation in the HOW program, rather than a market evaluation of the warranty itself. This is similar to the FHA result mentioned previously. Essentially, there may be a problem of adverse selection for HOW. The unexpectedly high claims incurred during the early years of the program provide some support for this hypothesis.

An alternative explanation is that HOW appeals to those builders who are less well established—smaller, newer firms—and whose houses therefore are less valued in the market. If that is true, and if the warranty is not perceived as a "sufficient" guarantee of quality equal to the "best" builders, then the coefficient could be negative even though the warranty has value; the coefficient may measure the joint effect of the warranty and the fact that the builder has a lesser reputation. A further implication is that the model may be misspecified. The markets for warranties and for new homes may be interrelated. Either or both the demand and supply of warranties may depend on the price of the house, as well as the price at which a house is supplied being a function of whether it carries a warranty. My judgment is that any such bias in the results is small, but I have not tested for it.

It is also possible that the variables fail to capture the effects of warranties. Homebuyers may have statutory implied rights, in the absence of any warranty. A warranty may therefore only define the buyer's rights, rather than expanding them, and may not be of value to the buyer (although it might be to the seller). This hypothesis also suggests a more elaborate modelling of the relationship between warranties and house prices, though it does not necessarily imply simultaneous-equations bias.

[16] This section has benefitted from the comments of Robert Crandall of the Brookings Institution and Edward Manfield of the FTC. The second and third hypotheses were suggested by them at the conference.

Defects, Repairs, and Home Prices

This section reports the empirical results for the defect and repair cost variables. These results are taken from regressions which have all the variables reported in Table 1, plus the measures of defect and repair cost.

Incidence of Defects

As previously mentioned, the MPR survey includes some 90 types of defects. It would be possible to include each of these as dummy variables in a hedonic regression, but I did not choose to do so for several reasons. Most fundamentally, the dummy variables for the incidence of a problem are less precise measures of problem severity than are the costs of repairing them. Even with 90 categories, not all problems in a given category are equally serious.

There are additional problems arising because of my choice of the CHC sample for the analysis. With 90 categories and 170 observations, there would be problems with degrees of freedom, and with categories for which only one or two homeowmers reported problems (as for instance with the "no heat" variable in Table 1). These problems might be circumvented by using the larger telephone survey sample. But others would remain. One involves the structure of the survey. The various defects are not reported in the same manner for each household. Instead, the respondent is asked to list up to six problems, essentially in order of importance, and each is then the subject of 60 questions. A given problem may be the first for one household, the third for another, and so on. The information about that problem will therefore appear in a different location on the computer record for each household. In order to assemble that information in the same place for each of the 90 problems, a substantial recoding of the computer tape would be required. This could be done, but the end result would still be inferior to the repair cost data.

The main reason for conducting any analysis of the incidence of defects is for purposes of comparison with the results of my earlier work. This is perhaps of some interest, but not worth a major effort. Fortunately, the MPR survey does have questions on two specific defects—wet basements and roof leaks—which are phrased in exactly the same way as the corresponding questions on the AHS, and indeed were included for comparability. These do not require recoding. I therefore estimated regressions using them. One is reported in Table 2. It does not suggest that defects have a strong relationship with house price. A wet basement is slightly more valuable than a dry one, but neither variable is significantly different from zero (no basement) or from each other. A leaky roof is associated with a 10 percent reduction in price, but the coefficient, while large, is not significant at the 20 percent level. These results are not very different from those in my earlier work. They did not suggest that high priority should be given to estimating regressions with other defect measures.

Repair Costs and Home Prices

Before discussing the relationship between repair costs and home prices, I want to define the repair cost concepts. The telephone survey asks respondents to report both the cost of the repair and the payer. In the event that the owner has borne the cost, or the problem has not been resolved, the respondent is asked if the builder was contacted, and if not, why not. The measures of repair cost used in this study include only the costs of repairing problems concerning which the builder was contacted, but did not resolve the problem. These are the actual costs paid by the homeowner for those problems which he or she resolved, and the estimated costs of repairing those problems which were unresolved at the time of the survey.[17] If the builder took care of the problem at his own expense, or the homeowner did not contact the builder, the costs of repair are not included.

A second cost measure is based on the home inspections. It includes the owner's costs for those problems verified by the CHC inspectors and considered to be problems which are the builder's responsibility, and also those additional problems identified by the inspector which were not noticed by the homeowner. It happens that the average total costs reported by the owners and the inspector are quite similar ($895 and $944, respectively), but the inspectors disallowed about half the costs reported by the owners, while finding many new problems. The distribution of costs across the sample of houses is quite different. The simple correlation between the two cost estimates is only +.23, and there are some perfectly extraordinary differences for specific houses—in one case, more than $14,000!

Table 3 reports the regression results for these cost measures. For convenience, I have converted the regression coefficients to price changes associated with each $100 increase in repair costs. The semi-logarithmic functional form forces each additional $100 to have a smaller impact on cost, but the effect is negligible—less than $1 for each of the coefficients reported in the table. For all practical purposes, it is appropriate to consider the relationship to be linear, and I do so.

The results are quite different and much stronger than those for the defect measures. Each $100 in actual or prospective repair costs incurred by the owner is associated with a $163 reduction in the price of the house. Each $100 verified or anticipated by the inspector is associated with a $137 price reduction. Both coefficients are significant at the two-tail, 10 percent level. Put differently, a buyer who paid $69,000 for the typical house (measured by the attributes in Table 1) could expect to incur no repair costs. A buyer who paid $67,500 for the same house could expect to incur about $900 in repair costs during the first 2½ years. These are reasonable magnitudes; they indicate that buyers incur non-pecuniary costs of repairs, such as time spent in finding a repair firm and

[17] The survey includes a category of "disputed" problems as well as "unresolved" ones; none of the homeowners in my sample reported any such problems.

the annoyance of putting up with the problem before the repairs are made.

Table 3 also reports the time pattern of repair costs reported by the owner. Problems occurring in the first and last six months of the 2½ year period are unrelated to value; those during the second year have a large and significant relationship. This is interesting because homes included in the survey could have been occupied for as short a period as one year. It is likely that some of the difference between the owner's and inspector's cost estimate occurs because the latter is anticipating problems that the former has not yet experienced. The CHC cost variable may therefore be the better measure.

Information is also available on the repair costs incurred by the builder, as reported by the owner and also as estimated by the CHC inspector for the same problem. This variable proved to be unrelated to home prices; the owner's estimate has a very large positive coefficient and the inspector's estimate a small negative one, but neither is as large as its standard error.

Quality, Repair Costs, and Home Prices

Besides the specific characteristics reported in Table 1 and the specific problems and costs encountered, the MPR survey also includes the owner's subjective rating of the quality of his or her home, on a scale of one to five. This measure, which was also available for the AHS, is worth some separate attention. I first included it as an additional independent variable in a regression similar to Table 1. It was moderately important; each one-step rise in quality was associated with about a 2.5 percent rise in price, but the coefficient was not quite significant at the 10 percent level; the t-ratio was 1.5. The other variables were essentially unaffected. I then re-estimated the regression in Table 1 with the quality rating as the dependent variable. The results were very poor; the overall regression did not pass the F-test for significance at the 5 percent level, the coefficient of determination was .19, and the only significant variable was air conditioning, with the wrong sign. Further investigation showed that the quality ranking is much more closely correlated with the repair cost variables than with the attributes of the house; it has a simple correlation of -.46 with the owner's estimate, for example. As might be expected, the correlation with the CHC estimate is smaller, -.18, but this is statistically significant; it is also larger than for almost any of the basic attributes of the house. The correlations with the two defects variables are -.12 with wet basements and -.15 with leaky roofs. Both are just below the conventional significance level. It appears that homebuyers are satisfied with their houses if they do not incur problems after moving in. That seems reasonable; the buyer knows the size, structure type and amenities before buying, and if dissatisfied with these features, is after all not required to buy.

The preliminary investigation of the quality rating suggested that it might serve as a proxy for problems. I therefore re-estimated the regression in Table 2, including the quality rating. The results, shown in Table 4, were not very different, either for quality or defects. Wet basements were still worth more than dry ones. A greater difference occurred when I re-estimated the repair cost regressions with the quality rating. The coefficients for both quality and total cost were smaller and less significant, as appears in Table 5.

In my earlier work, I used a similar quality rating and found that it was frequently related to the owner's subjective estimate of the value of the house; but it was not often related to the reported incidence of specific defects. Nor was it related to actual or planned repairs, which were also measured as dummy variables. These with other results led me to conclude that the value was unaffected by defects. The present results suggest that I might have had a stronger relationship had I been able to use actual repair costs instead of merely the fact that repairs occurred or were planned.

The Stochastic Nature of Problems

So far, this paper has assumed that the actual occurrence of a defect, or the actual repair cost, is what affects the price of the house. Stated sequentially, the buyer of a "cheap" house is buying a lower-quality house, and will have problems and incur repair costs in the future, which the builder will not pay for.

But this oversimplifies the relationships among defects, home quality, and prices. Defects occur—and fail to occur—in both high and low quality houses. Not all low quality homes develop expensive problems; some develop no problems at all. Similarly, some high quality homes turn out to have defects. The relationship is stochastic rather than deterministic. The buyer of a low-priced house is accepting a higher probability of problems, and higher expected repair cost, not a certainty that something will go wrong. The buyer of a high-priced home is purchasing some assurance that he is not likely to have problems, or that the builder will fix them if they arise, not a guarantee that none will occur. Buyers and probably builders cannot know in advance exactly which homes will turn out to have problems, but the quality of materials and workmanship can be observed and should be relevant. Thus it is the probability that a defect will occur, or the expected cost or repairs, which will affect the market price of a home. The use of the actual incidence of defects or repair costs, as in this and previous research, will therefore result in a downward bias in the regression coefficients for defects and costs, and an upward bias in their standard errors. This is certainly consistent with the weak relationships between defects and prices or rents in most studies using the AHS.

This phenomenon was first pointed out and investigated by Kenneth Wieand (1983). He also developed a technique for calculating the "quality" of a house and the corresponding probability that it will experience a given defect. This study uses his methodology.

Computing the Probability of Defects

Wieand begins by assuming that the price of a house is a function of its characteristics (size, amenities, location) and the expected probability of a defect or expected cost. Symbolically,

$$\ln P = b_o + \sum_{i=1}^{n} b_i X_i + \sum_{j=1}^{m} b_j E(D_j) \qquad (3)$$

where $E(D_j)$ is the probability of a defect. This assumes that buyers judge the quality of various attributes in terms of the expected costs of repairing them; the b_j measure total repair costs, including non-pecuniary costs, incurred by the owner. It also assumes that the various defects are independent, and that defects and attributes are also unrelated. Buyers are assumed to be risk-neutral; the change in price is linear with respect to the probability that problems will occur and the expected cost of repairs.

To determine $E(D_j)$, Wieand first estimates a hedonic regression with no quality, defect, or repair variables, such as equation (2) and Table 1. But this now has a somewhat more complicated interpretation:

$$\ln P = b_o^* + \sum_{i=1}^{n} b_i X_i + e^* \qquad (4)$$

where the constant term b_o^* is in reality

$$b_o + \sum_{j=1}^{m} b_j \overline{D}_j,$$

$\overline{D}_j$ being the average probability of defects or the average level of repair costs, and thus the constant term for average quality units in the sample; and e^* is

$$\sum_{j=1}^{m} b_j (D_j - \overline{D}_j) + e.$$

Thus the error term in equation (4) is a function of the difference between the actual and the expected incidence of defects or repair costs. The expected value of the error term will be negative if the probability of defects is above average, and positive if it is below average. A negative error term in equation (4)—and therefore in equation (2) and Table 1—is likely to indicate a lower quality house, meaning one with a higher probability of defects or higher expected repair costs.[18]

[18] At the conference Richard Duke of the FTC raised the possibility that building codes may af-
(footnote cont'd)

This means that the residuals from Table 1 are themselves a measure of home quality, and can be used to construct probabilities of defect or expected repair costs. To do so, the residuals are grouped into intervals, ranging from large negative to large positive. I created 10 such categories, each with about 17 observations, but with the endpoints of the intervals being set where natural breaks in the residual distribution occurred as much as possible.[19] The number of units in each class with specific defects—wet basements or leaky roofs—is then calculated. The incidence of each defect becomes the measure of $E(D_j)$ for that defect. Similarly, the average repair cost incurred by owners of units in each interval, and the average estimated by the CHC inspector, are also calculated and used to measure expected repair costs. These averages are then added to the variables for each observation in the sample, with the $E(D_j)$ for the observation being the average for that observation's residual class. Finally, equation (3) is estimated with the new expectational variables included instead of the actual occurrence of a problem or the actual repair costs.

Empirical Results

The results of this procedure are shown in Tables 6 and 7, corresponding to Tables 2 and 3. The differences are dramatic confirmation of Wieand's insight. In Table 6, the coefficients are large and very significant indeed. For example, a 10 percentage point increase in the probability of a leaky roof is associated with a 42 percent reduction in house price, or $28,000 for the typical house. Since a leaky roof is highly unlikely to cost $28,000 to repair, I hypothesize that this variable is correlated with other quality problems not identified separately in the MPR survey. A leaky roof is a rare phenomenon in a new home; the incidence in my sample was 6.5 percent (11 observations) and the probabilities ranged only from zero to .11. A 10 percentage point increase in the probability of a wet basement is less serious; wet basements are more common, occurring in 34 percent of all sample homes with basement, and having a probability range of zero to .75 for residual classes. However, in this formulation wet basements do have the expected negative sign and are significant.

A similar pattern occurs for the measures of owner's repair costs in Table 7. Whereas neither the owner's or the CHC inspector's estimate was quite conventionally significant when the actual costs for the home were used, both are now highly significant. The coefficients are also somewhat larger than Table 3. Buyers who paid $100 less for the typical house could expect to incur $38 to $50 in repair costs. Since the cost estimates are quite a bit less than $100,

---

fect e*. Houses in jurisdictions with stringent building codes may have negative error terms, because builders are forced to build higher quality houses than would be preferred by buyers in an unregulated market. My view is that the effect of building codes on house prices is slight, but I have not attempted to investigate their possible impact in this paper.

[19] I actually created two residual distributions, one giving more weight to equality of observations by residual category, the other to "natural" boundaries between categories. Results for both were similar. The paper reports the latter.

it seems likely that they could also expect some aggravation, and perhaps some further repair costs later on, after 2½ years.

It is perhaps also worth mentioning that the coefficients of determination in these regressions increase sharply. The difference between Table 2 and Table 6 is 16 percentage points, from .73 to .89. The increases for the repair cost regressions are six percentage points (.73 to .79) for the owner's estimate, and 11 (.71 to .82) for the inspectors. Many of the other independent variables—the X_i—also become more significant. In all three regressions, this is true for rooms, bathrooms, air conditioning, electric heat (without heat pump) and most of the regional dummies, and for FHA financing. It is also true in one case for the HOW variable. In the CHC cost regression of Table 7, the HOW variable has a coefficient of -.09 and a t-ratio of 1.8.

Conclusion

Although many economists have expected the opposite, the housing market seems to work. Buyers may not be able to evaluate the quality of new homes, but the market apparently does. Homes which turn out to be lower in quality do initially sell for less. And the magnitudes of the relationships between initial price and later repair costs are quite plausible; future repair costs are less than the initial price difference, allowing for non-pecuniary costs of arranging for repairs.

The market may work because home buyers and builders are not the only actors in it. There are at least two others with technical expertise, one on each side of the market. Construction lenders verify each stage in the building process by inspection, before advancing further funds. Appraisers evaluate homes for the purpose of advising mortgage lenders, who then determine how much and one what terms they will loan money to buyers. Their expert knowledge should contribute to the effective operation of the new home market. In addition, builders do acquire reputations for quality, even though entry and exit is easy and frequent in the homebuilding industry; bankruptcies are high in every recession. By whatever process, consumer protection for new home buyers appears to be provided by the market.

It is still possible for a buyer to get a "lemon," because quality problems are stochastic phenomena. Judging by the CHC inspections, some of the buyers in this sample will get stung, or perhaps they already have (since the inspections occurred in 1980). Thus warranties would appear to be desirable; the market price is a guide to the expected quality of the home, but it is not an infallible predictor. Buyers, however, don't seem to be too worried about "lemons." None of the warranty variables were significant in any of the regressions; most were negative. Buyers were apparently willing to take their chances.

These findings are of course rather tentative. They are limited by the nature of the data on which they are based. The MPR survey omits some characteristics which have been found to be important in other studies, especially the presence of a garage. On the other hand, it includes the square feet of floor space within the house, which is acknowledged to be important but is infrequently available in housing data. I do not see any reason why the omitted variables should bias the results for the defect and repair cost variables, but it might have happened.

More important is the fact that the sample is small, and covers metropolitan areas in many parts of the nation. Housing markets in fact are local rather than national, and my attemps to insure comparability within the sample may be inadequate. Also, the sample was designed to overrepresent houses where the owner reported problems. It may be worthwhile to replicate this study with the much larger telephone sample, despite the lack of CHC inspection data and the need to create the SMSA size and distance variables.

Table 1

HEDONIC REGRESSION OF NEW HOME PRICES, EXCLUDING DEFECTS
(dependent variable is natural log of real purchase price, measured in 1979 dollars)

INDEPENDENT VARIABLE	COEFFICIENT	STD. ERROR	T-RATIO
CONSTANT	10. 4697	0.1767	59. 24
BATHS	0. 0645407	0.04082	1. 581
ROOMS	0. 0197372	0.01980	0. 9970
SEWERS	0. 0196898	0.05248	0. 3752
BASEMENT	0. 0348645	0.05829	0. 5981
ELECTRIC	-0. 110443	0.06403	-1. 725
OILHEAT	-0. 0272517	0.09051	-0. 3011
NOHEAT	0. 232760	0.2525	0. 9217
OTHERHEAT	-0. 291775	0.1841	-1. 585
AC	0. 0872141	0.05370	1. 624
GUTTERS	0. 00376099	0.05175	0. 07268
ONEFLOOR	-0. 0375809	0.06196	-0. 6066
TWOFLOOR	-0. 0388662	0.06587	-0. 5900
HEATPUMP	0. 147055	0.06869	2. 141
BRICK	-0. 0480017	0.07160	-0. 6704
CONCRETE	-0. 0145247	0.1160	-0. 1253
WOOD	-0. 0404680	0.06427	-0. 6297
ALUMINUM	0. 0174916	0.07513	0. 2328
DATE	0. 000285692	0.004311	0. 06627
HOUSIZE	0. 000303211	4.115E-05	7. 368
DISTANCE	-0. 000162270	0.0006671	-0. 2432
SQDISTANCE	2. 34735E-07	1.587E-06	0. 1479
POPCORR77	1. 57189E-05	8.763E-06	1. 794
NEWENG	-0. 0825823	0.1379	-0. 5989
MIDATL	-0. 337707	0.1024	-3. 298
ENORCEN	-0. 169749	0.08784	-1. 932
WNORCEN	-0. 222571	0.1233	-1. 805
SOUATL	-0. 235334	0.08175	-2. 879
ESOUCEN	-0. 191945	0.09300	-2. 064
WSOUCEN	-0. 297285	0.1207	-2. 463
ROCKIES	-0. 144414	0.1035	-1. 395
CASHSALE	0. 0786091	0.08029	0. 9791
FHA	-0. 232838	0.08607	-2. 705
VA	-0. 128420	0.06666	-1. 927
OTHERMOR	-0. 142212	0.2564	-0. 5547
TRACT	-0. 165061	0.05010	-3. 295
HOW	-0. 0499141	0.06097	-0. 8187
BUILDWAR	0. 00135233	0.05017	0. 02695
OTHERWAR	0. 0940759	0.1340	0. 7018

$R^2 = .725$

Table 1 (cont'd)

Description of Variables:

Baths	Number of bathrooms
Rooms	Number of rooms, excluding bathrooms, halls, porches, unfinished attics or basements
Sewers	Dummy variable for connection with public sewer system
Basement	Dummy variable for presence of basement
Electric	Dummy variable for electric heat
Oilheat	Dummy variable for oil heat
Noheat	Dummy variable for absence of any heating system
Otherheat	Dummy variable for other heat (omitted category is gas heat)
Ac	Dummy variable for presence of central air conditioning
Gutters	Dummy variable for presence of gutters and downspouts
Onefloor	Dummy variable for houses having only one floor (excluding basements and attics)
Twofloor	Dummy variable for houses having two or more floors (omitted category is split level)
Heatpump	Dummy variable for houses with electric heat and heat pump
Brick	Dummy variable for houses with brick as primary material of exterior walls
Concrete	Dummy variable for houses with concrete as primary material of exterior walls
Wood	Dummy variable for houses with wood as primary material of exterior walls
Aluminum	Dummy variable for houses with aluminum siding as primary material of exterior walls (omitted category is "other," apparently mostly stucco)
Date	Month in which closing occurred (March 1977 = 0; October 1978 = 19)
Housize	Number of square feet of living space within the house
Distance	Distance of the centroid of the zip code in which house is located from the Central Business District of the Standard Metropolitan Statistical Area in which it is located
Sqdistance	Square of "distance" variable
Popcorr77	Population of SMSA in 1977 (as corrected by author)
Neweng	Dummy variable for location in New England Census Division
Midatl	Dummy variable for location in Mid-Atlantic Census Division
Enorcen	Dummy variable for location in East North Central Census Division
Wnorcen	Dummy variable for location in West North Central Census Division
Souatl	Dummy variable for location in South Atlantic Census Division
Esoucen	Dummy variable for location in East South Central Census Division
Wsoucen	Dummy variable for location in West South Central Census Division
Rockies	Dummy variable for location in Mountain Census Division (omitted category is Pacific Census Division—Alaska and Hawaii excluded from sample)
Cashsale	Dummy variable for house purchased without a mortgage
FHA	Dummy variable for house purchased with FHA-insured mortgage
VA	Dummy variable for house purchased with VA-guaranteed mortgage
Othermor	Dummy variable for house purchased with mortgage other than FHA, VA or Farmers Home Administration (omitted category is conventional mortgage)
Tract	Dummy variable for house located in tract development
HOW	Dummy variable for house with Home Owners Warranty
Buildwar	Dummy variable for house with builder's warranty
Otherwar	Dummy variable for house with sales agent's or other warranty (omitted category is no warranty)

Table 2

RELATIONSHIP BETWEEN DEFECTS AND HOME PRICES

DEFECT	COEFFICIENT	T-RATIO
Dry basement	+.026	(0.4)
Wet basement	+.048	(0.7)
Leaky roof	-.100	(1.2)
R^2	.728	

Note: Other variables in the regression are those shown in Table 1

Table 3

REPAIR COSTS AND HOME PRICES

	REGRESSION				
COST MEASURE	(1)	(2)	(3)	(4)	(5)
Ownertotal	-$163 (1.7)	...	...	...	...
Owner-6	...	-$56 (0.5)	...	...	...
Owner-12	...	-$297 (1.3)	...	...	...
Owner-24	...	-$1126 (2.7)	...	...	...
Owner-30	...	+$ 7 (0.0)	...	...	...
CHCtotal	...	...	-$137 (1.6)	...	...
Owner/builder	...	...	...	+$1031 (0.4)	...
CHC/builder	...	...	...	...	-$61 (0.3)
R^2	.731	.755	.730	.737	.730

Note: Dollar figures are changes in new home price for each $100 of repair costs. Numbers in parentheses are t-ratios of regression coefficients. Other variables in the regressions are those shown in Table 1.

Description of variables:

Ownertotal Telephone survey: total costs incurred by owner on problems not resolved by builder, or if builder was not contacted (in 1978 dollars)

Table 3 (cont'd)

Owner-6	Telephone survey: amount of "ownertotal" for problems incurred in first six months of occupancy
Owner-12	Telephone survey: amount of "ownertotal" for problems incurred in second six months of occupancy
Owner-24	Telephone survey: amount of "ownertotal" for problems incurred in second year of occupancy
Owner-30	Telephone survey: amount of "ownertotal" for problems incurred in 25th through 30th month of occupancy
CHCtotal	CHC inspection: total costs estimated by inspector for problems included in "ownertotal" and verified by inspector, and additional problems discovered by inspector
Owner/builder	Telephone survey: owner's estimate of total cost incurred by builder in resolving problems (in 1978 dollars)
CHC/builder	CHC inspection: inspector's estimate of total costs of repair builder-resolved problems listed by owner

Table 4

RELATIONSHIP BETWEEN QUALITY RATING, DEFECTS, AND HOME PRICES

VARIABLE	COEFFICIENT	T-RATIO
Dry basement	+.012	(0.2)
Wet basement	+.053	(0.8)
Leaky roof	-.088	(1.1)
Quality rating	+.025	(1.4)
R^2	.810	

Note: Other variables in the regression are those shown in Table 1

Table 5

QUALITY RATING, REPAIR COSTS, AND HOME PRICES

VARIABLE	REGRESSION (1)	(2)	(3)
Quality rating	+.014	+.006	+.020
	(0.7)	(0.3)	(1.1)
Ownertotal	-$121	...	...
	(1.1)		
Owner-6	...	-$39	...
		(0.3)	
Owner-12	...	-$278	...
		(1.1)	
Owner-24	...	-$1089	...
		(2.5)	
Owner-30	...	+$48	...
		(0.4)	
CHCtotal	...	...	-$111
			(1.3)
R^2	.732	.745	.732

Note: Dollar figures are changes in new home price for each $100 of repair costs. Numbers in parentheses are t-ratios of regression coefficients. Other variables in the regressions are those shown in Table 1. Description of variables is the same as in Table 3.

Table 6

THE RELATIONSHIP BETWEEN EXPECTED DEFECTS AND HOME PRICES

DEFECT	COEFFICIENT	T-RATIO
Wet basement	-.029	(4.2)
Leaky roof	-.549	(13.9)
R^2	.891	

Note: Coefficients represent changes in the natural log of price associated with a 10 percentage point increase in the probability of the defect. Other variables in the regression are the same as those in Table 1.

Table 7

EXPECTED REPAIR COSTS AND HOME PRICES

	Regression	
Cost measure	(1)	(2)
E(Ownertotal)	-$202 (6.4)	...
E(CHCtotal)	...	-$263 (8.4)
R^2	.791	.821

Note: Dollar figures represent changes in price associated with each $100 of expected repair costs. Numbers in parentheses are t-ratios of the regression coefficients. Other variables in the regressions are the same as those in Table 1.

References

Brewster, J. Alan; Crespi, Irving; Kaluzny, Richard; Ohls, James; and Thomas, Cynthia, *A Study of Home Inspection and Warranty Programs*, Princeton, New Jersey: Mathematica Policy Research, 1977.

Dole, Elizabeth Hanford, "Homebuyer Warranties: Constructing a Good Foundation for Builder/Consumer Relations," Speech to the National Association of Home Builders convention, Las Vegas, Nevada, January 21, 1979.

Follain, James R., Jr., and Malpezzi, Stephen, *Dissecting Housing Value and Rent: Estimates of Hedonic Indexes for Thirty-Nine Large SMSAs*, URI 28700, Washington, D.C.: The Urban Institute, 1980.

Halvorsen, Robert, and Palmquist, Raymond, "The Interpretation of Dummy Variables in Semilogarithmic Equations," *American Economic Review*, June 1980, 70, 474-475.

Kaluzny, Richard L., *A Survey of Homeowner Experience with New Residential Housing Construction*, Princeton, New Jersey: Mathematica Policy Research, 1980.

Leland, Hayne E., "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," *Journal of Political Economy*, December 1979, 87, 1328-1346.

Malpezzi, Stephen; Ozanne, Larry; and Thibodeau, Thomas, "Characteristic Prices of Housing in Fifty-Nine Metropolitan Areas," Working Paper 1367-1, Washington, D.C.: The Urban Institute, n.d.

McKean, Roland N., "Products Liability: Implications of Some Changing Property Rights," *Quarterly Journal of Economics*, November 1970, 84, 611-626.

Miller, Norman G., "Residential Property Hedonic Pricing Models: A Review," in C.F. Sirmans, editor, Research in Real Estate, vol. 2, Greenwich, Conn.: JAI Press, 1982, 31-56.

Mills, Edwin S., *Studies in the Structure of the Urban Economy*, Baltimore: The Johns Hopkins Press, 1972.

Muth, Richard F., *Cities and Housing*, Chicago: University of Chicago Press, 1969.

Ozanne, Larry; Andrews, Marcellus; and Malpezzi, Stephen, "An Assessment of Annual Housing Survey Hedonic Indexes Using Demand Experiment Data," Working Paper 249-21, Washington, D.C.: The Urban Institute, May 1979.

Snow, Jane, "Five Years Later—The HOW Story," *Challenge!* September 1979, 10, 2-5. (This magazine is published by the U.S. Department of Housing and Urban Development.)

Stanton, Thomas H., "Consumer Protection and National Housing Policy: The Problem of New-Home Defects," *Case Western Reserve Law Review*, Spring 1979, 20, 527-549.

U.S.Bureau of the Census, *Annual Housing Survey: Housing Characteristics for Selected Metropolitan Areas*, Series H-170, Washington, D.C.: Government Printing Office.

Weicher, John C., "Product Quality and Value in the New Home Market: Implications for Consumer Protection Regulation," *Journal of Law and Economics*, December 1981, 24, 365-397.

Weicher, John C. and Hartzell, David, "Hedonic Analysis of Home Prices: Results for 59 Metropolitan Areas," in C.F. Sirmans, editor, *Research in Real Estate*, vol. 2, Greenwich, Conn.: JAI Press, 1982.

Wieand, Kenneth F., "The Performance of Annual Housing Survey Quality Measures in Explaining Dwelling Rentals in 20 Metropolitan Areas," *AREUEA Journal*, Spring 1983, 11, 45-68.

The Effect of Warranty Provisions on Used Car Prices

James E. Anderson and *Frank M. Gollop**
Boston College

A 1974 House Subcommittee report concluded that warranties generally did more to limit sellers' obligations than to protect or enhance buyers' rights.[1] If true, a warranty would be expected to reduce a seller's expected liability and consequently the marginal cost of the warranted good.

Ten years have passed since that report was released. During those years the Magnuson-Moss Warranty Act was passed and many states have enacted or embellished Uniform Commercial Code warranty provisions and/or Unfair and Deceptive Acts and Practices statutes. Consumer protection divisions within state attorney generals' offices are now commonplace.

It thus seems reasonable to ask whether warranty legislation and enforcement have made a difference. The economic criterion is clear. If warranties meaningfully protect consumers, then the marginal cost and consequently the price of a warranted good should be positively associated with legally protected warranty provisions. If warranty laws and legal action are responsible, then the price of an otherwise homogeneous good should vary systematically with each state's commitment to warranty protection both through legislation and enforcement.

This study tests this hypothesis for the used car market and estimates the dollar cost of legally imposed warranties on the sale of used cars. The method is essentially to compare prices of identical used cars sold in states with differing warranty provisions while controlling for other variables influencing car prices over states. Two factors make this approach promising. First, there are good reasons for believing that the structure of production and demand conditions in the used car market are quite simple. An economic model of dealer behavior need not be complicated. Second, the National Automobile Dealers Association has made available a transaction-specific price data base covering more than 120,000 retail car sales during the June to August 1983 period. Thus we can propose a model where identification is reasonably straightforward and the data extremely well-suited to the analysis.

A short run model of producer behavior is derived in section 1. Used car dealers are assumed to exist in monopolistically competitive markets where spatial differentiation is the presumed principal dimension of product heterogeneity. Marginal cost is assumed constant out to a short run capacity constraint that is assumed to be non-binding. Variable production costs include the wholesale

* We wish to thank Shawn Duff and Jim Platteter for able research assistance. We especially acknowledge the contribution of Joshua Zissman who provided invaluable assistance both in our understanding of the warranty laws and in the development of the warranty variables.

[1] Staff of House Interstate and Foreign Commerce Committee, Subcommittee on Commerce and Finance, 93d Congress, 2d Session, ***Report on Consumer Product Warranties*** (1974).

price of the used car to the dealer, labor and fuel/electricity expenses, and the dealer's expected warranty costs. This last term equals the product of the probability of the car's "failure," the probability that the buyer will bring a successful warranty claim, and auto repair costs. A cross section model across states and car makes is developed. All necessary data are described in section 2.

Warranty provisions mandated by state law and state enforcement efforts vary substantively by state. The Appendix details our characterization of the Magnuson-Moss Warranty Act (MM), Article 2 of the Uniform Commercial Code (UCC), and the Unfair and Deceptive Acts and Practices statutes (UDAP) as applied in each state to used car transactions. Nine indicator variables summarize the interstate variability in the economic severity of warranty legislation and enforcement under the consumer protection laws. These dummy variables proxy the central probability affecting the expected cost of warranty: the probability of a successfully enforced claim.

Dealers voluntarily offer warranties and voluntarily comply but the model developed below suggests that dealers only offer the minimum required warranty and voluntarily comply so as to bring costs of compliance and noncompliance into equality. Both are driven by the above legislative and enforcement variables.

The econometric results are presented in section 3. Some warranty provisions have statistically significant effects while others do not, but all statistically significant legislative coefficients have the expected signs. State enforcement efforts, in addition, appear to be quite important. The results indicate that, taken as a whole, enforced consumer protection legislation adds slightly more than $200 to the retail price of a representative late-model used car. This amounts to about 3.5 percent of the average purchase price of the used car models examined in this study. The magnitude of the effect may be sensitive to further modeling refinements. However, the basic finding that the overall effect of warranties is statistically significant and has the anticipated positive sign appears robust. We can conclude (1) that buyers and sellers take the warranty statutes seriously and (2) that used car dealers charge a non-trivial sum for the insurance policy they are required to extend to buyers.

1. The Model of Used Car Prices

We build a simple model of the price of used cars based on profit maximizing behavior of dealers in a monopolistically competitive environment. "Good will" is a fixed factor in the short run and a representative dealer faces a downward sloping demand curve for his product based on location. The cost function is linear (there are fixed coefficients in production) in the relevant range of factor prices, which allows a particularly simple form for the short run profit maximizing price. In the long run, with constant returns, this model will produce an equilibrium identical to the competitive equilibrium.

Assume the demand curve facing the representative dealer is a linear function of price (P), income (Y), and the number of licensed drivers per dealership (LPD). Short run marginal (average) cost is a linear function of the price of used cars at wholesale (w_M), energy costs (w_E), salaries per employee in dealerships (w_L), and warranty costs (w_W). The profit maximizing price is

$$P = -P_Q Q + C_Q(w_M, w_E, w_L, w_W). \qquad (1)$$

Using the linear demand ($P = a_1 Y + a_2 LPD - a_3 Q = a - a_3 Q$) and cost functions in parametric form, (1) is solved for P. Note that $-P_Q Q = -P + a$. Then (1) is replaced by

$$2P = D_0 + D_1 Y + D_2 LPD + C_1 w_M + C_2 w_E + C_3 w_L + C_4 w_W. \qquad (2)$$

Trivially, $C_4 = 1$ (a warranty is required on each car sold). It might at first blush also be reasonable to require $C_1 = 1$ (a car out requires a car in). In fact, the assumption $C_1 = 1$ can be rejected. Ideally, w_M measures the opportunity cost of used cars at wholesale. The NADA wholesale price data, however, add only some new information to the dealer's forecast of each model's opportunity cost. The constant term D_0 thus picks up part of the influence of w_M.

We should note that (2) approximates (1) even in the nonlinear demand case where $-P_Q Q$ remains a function of endogenous variables. The right-hand side of (2) uses instruments for the endogenous variables. The interpretation of the coefficients is of course blurred, but it can be shown that the signs will be preserved.

A more serious specification problem with (2) is that the warranty term may affect the demand curve. We expect that the better the warranty, the greater the consumers' willingness to pay. Thus $C_4 = 1 + \varrho$, $\varrho > 0$, where ϱ represents the consumers' valuation of the warranty. This possibility may be rather remote when one considers that it requires consumers to be sensitive to the nuances of law as they vary across states in the manner detailed in the Appendix. Nevertheless, if $\varrho > 0$, it means that not all the "cost of warranty" identified below is in fact a cost to dealers.

The warranty cost variable w_W is the most interesting term. All dealers are assumed to offer voluntarily only the minimum required warranty. This in fact appears to be the case based on a very casual sampling. Disregarding risk aversion for the moment, a fair premium for a warranty is the probability of paying out times the amount which must be paid. The amount to be paid depends on what fails and on the expense of fixing it. Suppose the expense varies by state, proxied by a labor cost variable for repair shops, w_R. The probability of a failure necessitating an expense of w_R varies by make/vintage, but not by state. The probability of payment equals the probability of failure times the probability of a successful claim. The probability of an effective claim depends on state-specific warranty legislation and enforcement efforts. Profit-maximizing behavior implies that, in equilibrium, the value of the dealers' strategies of voluntary compliance and resistance are the same. At the margin, the enforceability of the claim determines the value, and this varies by state but not by make/vintage. Thus, our model of warranty cost is

$$w_W(v,s) = \Pi(d(s))\pi(v)a(v)w_R(s), \tag{3}$$

where s denotes state s, v identifies make/style/vintage v, Π represents the probability of an enforceable claim, π denotes the probability of failure of model v, and d(s) is a vector of dummy variables representing the strength of warranty provisions. Note that d(s) varies by state. a(v) is the labor requirement to fix the failure. In practice, it is not possible to separately identify π and a.

Equation (3) is made parametric by imposing the logistic cumulative density function on Π and π. This function closely approximates the normal but is far easier to use. We justify the use of the normal by appeal to the law of large numbers. The probability of an enforceable claim, $\Pi(d(s))$, using the logistic function is $L(X(s)) = 1/(1+\exp(-X(s))$ where $X(s) = \Sigma B_i d_i(s) + B_0$, a linear function of state characteristics. While we eventually plan to use the true logit function, the results below are based on the first-order approximation to it expanded about $X(s) = 0$; *i.e.*, $\hat{L}(X(s)) = \frac{1}{2} + \frac{1}{4}X(s)$. For the probability of failure, a longer range research objective is to make π a logistic function of the Consumer Union reliability ratings. There are five Consumer Union categories, which translate to five dummy variables, $r_j(v)$. Thus the probability of failure, $\pi(v) = L(y(v))$, where $y(v) = \Sigma D_j r_j(v)$. For the results reported below, the sample is stratified by reliability class as described in Part 2. No effort is made to identify $\pi(v)$. $\pi(v)$ and a(v) are set using external information with considerable arbitrariness. This affects the scale of the terms in $\hat{L}(X(s))$.

The model above is certainty equivalent, which is appropriate for small variance cases. We now examine this assumption. For a large enough volume of insurance transactions, each with independent risk drawn from an identical distribution, the variance vanishes. An effectively certain proportion of the insurance policies result in claims. Formally, let Z be the variance associated with the individual risk. For N policies insuring identical independent risks, the variance for the pool is Z/N. For the representative used car dealer, the issue is whether N is large enough so that Z/N is too small to matter. For finite N, the standard analysis of the behavior of risk averse agents tells us that the formula for W must be amended to

$$w_W(v,s) = \Pi(s)\pi(v)a(v)w_R(s) + (\tfrac{1}{2})Az(v,s)/N(s), \tag{3'}$$

where A is the coefficient of absolute risk aversion of the representative dealer, assumed the same across states. N(s) is the average volume of sales, and z(v,s) is the variance of the payout:

$$z(v,s) = \pi(s)\pi(v)[1-\Pi(s)\pi(v)]\{a(v)w(s)\}^2. \tag{4}$$

It would be useful to test (3′).

The main eccentricity of the model is that no allowance is made for adverse selection phenomena. There are two justifications. First an empirical study of the used pickup truck market (Eric Bond, *American Economic Review*, Sept., 1982) found no evidence of adverse selection. Second, despite the prestige of Akerlof's seminal paper on adverse selection, a paper which relied on the used car market as a motivating example, we believe asymmetric information between ultimate buyer and ultimate seller is a trivial problem in that market. "Lemons" are quite rare, and the main uncertainty is over failure of systems whose likelihood of failure is known by a number of informed buyers and sellers. If adverse selection is in fact at work, it will mean that weaker warranty states will tend to sell worse cars. Consumers will discover this, the price will drop, and the price difference between the strong and weak warranty states, *ceteris paribus*, will not give a true measure of the cost of the warranty. In a limiting case, the price difference will reflect adverse selection alone with warrantied cars being failure-free, and the cost of the warranty being zero. Thus our assumption on adverse selection is critical. Unfortunately, since there are no data for cars like the pickup truck data used by Bond, there is no way to check this assumption.

2. The Data

The central feature of this paper's data set is the transaction-specific used car data collected by the National Automobile Dealers' Association (NADA). The Association surveys hundreds of car dealers and auction companies monthly to obtain information for its publication, *The NADA Official Used Car Guide*. Each respondent is asked to identify itself by name and geographic location, to describe each sold vehicle's physical characteristics (make, year, series, body style, non-standard equipment, and condition), to report the transaction price gross of any trade-in allowance, and to identify whether the sale was retail or wholesale. NADA constructs a record for each transaction, but only after deleting the respondent's name and any description of the vehicle's condition and non-standard accessories.

NADA kindly made available the monthly data tapes for June, July, and August 1983. The tapes collectively contain more than 300,000 observations for retail and wholesale transactions in used cars and trucks. The 126,000 retail car records form the basis for the retail price variable P.

Observations were grouped by make, year, series, and body style and then allocated to states on the basis of the transaction state codes. The variable P(v,s) defines the mean retail price of the vth model in state s. It is important to note that the NADA price data measure net prices. NADA asks its responding dealers to submit sale price data net of any sales and excise taxes, license fees, and dealer preparation charges.

A sample of 33 car models was selected for this study. They are identified in Table 1. Each model is represented on the NADA tapes by at least 50 transactions spanning no fewer than 25 states. The actual number of retail sale NADA

Table 1

Model Summary

Model	Number of Observations: Retail	Number of Observations: Wholesale	Frequency of Repairs*
1977 Chevrolet Monte Carlo-V8 Coupe (2d.) "S"	194	1119	A
1978 Chevrolet Monte Carlo-V8 Sport Coupe (2d.)	236	1012	A
1978 Chrysler Cordoba-V8 Hardtop (2d.) Specialty	77	110	A
1978 Ford Thunderbird-V8 Hardtop (2d.)	294	1025	A
1978 Pontiac Grand Prix-V8 Coupe (2d.) SJ	265	695	A
1978 Honda Accord Hatchback (3d.) FWD	538	216	A
1978 Honda Civic Sedan (2d.) FWD	268	205	A
1979 Cadillac Deville Sedan (4d.)	209	646	A
1979 Chevrolet Camaro-V8 Spt. Coupe (2d.)	99	321	W
1979 Chevrolet Monte Carlo-V8 Sport Coupe (2d.)	276	1043	A
1979 Chevrolet Monza-L4 Coupe (2d.) H'back Spt. 2*2	76	132	W
1979 Dodge Omni-4 cyl. Hatchback (4d.)	90	109	W
1979 Oldsmobile Cutlass Supreme-V8 Coupe (2d.)	234	812	A
1979 Plymouth Horizon-4 cyl. Hatchback (4d.)	232	320	W
1979 Pontiac Grand Prix-V8 Coupe (2d.) SJ	273	160	A
1979 Honda Accord Hatchback (3d.) 5 speed	485	207	A
1979 Honda Prelude Coupe (2d.) 5 speed	368	96	A
1979 Mazda Rotary Engine RX7 Coupe (2d.) "S"	563	153	B
1980 Buick Skylark-V6 Sedan (4d.)	66	147	W

(Continued)

Table 1 (cont'd)

MODEL SUMMARY

Model	Number of Observations: Retail	Number of Observations: Wholesale	Frequency of Repairs*
1980 Cadillac Deville Coupe (2d.)	73	156	A
1980 Chevrolet Monza-L4 Coupe (2d.) H'back, Spt. 2+2	116	125	W
1980 Ford Thunderbird-V8 Hardtop (2d.)	276	870	A
1980 Plymouth Horizon-4 cyl. Hatchback (4d.)	92	100	W
1980 Pontiac Phoenix-L4 Coupe (2d.)	79	50	W
1980 Pontiac Sunbird-L4 Spt. Coupe (2d.)	395	320	W
1980 Honda Accord Hatchback (3d.) 5 speed	416	171	B
1980 Mazda Rotary Engine RX7 Coupe (2d.) "S"	404	109	B
1980 Toyota Celica Sport Coupe (2d.) SJ	318	342	B
1980 Toyota Corolla Sedan (2d.) Deluxe	183	43	B
1980 Toyota Pickup	262	177	B
1981 Honda Accord Hatchback (3d.)	377	209	B
1981 Toyota Corolla Sedan (2d.) Deluxe	191	59	B
1981 Toyota Pickup 4 speed	253	158	B

* A = average
B = better than average
W = worse than average

records corresponding to each model is presented in Table 1.

The short run unit cost function is defined in terms of the wholesale prices of used cars (w_M), the prices of labor (w_L) and energy (w_E) inputs, and warranty costs (w_W). State-specific energy and labor prices are measured, respectively, as electricity plus fuel costs per MBTUs and annual salaries per employee in car dealerships selling both new and used cars (SIC 551) or used cars exclusively (SIC 552). The former is reported in the *Annual Survey of Manufactures*[2]; the latter is found in *County Business Patterns*[3]. Both documents are published by the Census Bureau. The wholesale price $w_M(v,s)$ measures the material cost of the vth model car to the representative dealer in state s. Mean wholesale prices across models and states are derived from the NADA tapes in exactly the same manner as is P(v,s). The number of wholesale transactions on the NADA tape for each model is given in Table 1.

The $w_M(v,s)$ variable not only provides the necessary wholesale price data but also provides a partial control for variation across states in used car condition and accessories, physical properties that are not available from the NADA tapes. Cars sold in either retail or wholesale markets in, for example, Arizona or Georgia are unlikely to have been undercoated or to have decayed through exposure to road salt but are likely to be equipped with air-conditioning. The opposite most probably holds in Montana and Maine.

Per capita personal income (Y) and licensed drivers per dealership (LPD) together with P form the arguments of the demand function facing the representative used car dealer. Values for Y by state are taken directly from the *Statistical Abstract*[4]. This variable serves the familiar role of any income variable in a demand function. The LPD variable identifies the extent of the representative dealer's geograhic market area. Spatial differentiation, recall, is the presumed principal dimension of the product heterogeneity present in the monopolistically competitive industry. It is measured as the ratio of licensed drivers reported in the *Statistical Abstract* to the number of dealer establishments reported in *County Business Patterns*[5].

The probability that an abused consumer will bring a successful warranty claim against a used car dealer is a function of the state's commitment to warranty protection through both legislation and enforcement as well as the consumers' awareness of the state's commitment. The latter is proxied by d_{ED}, the fraction of each state's adult population with 1-3 years of college as reported in the *Statistical Abstract*[6]. The former is measured by a vector of nine indicator variables. Eight characterize the stringency of warranty laws across states. The d_{UCC}, d_{UCCOP}, d_{NEGD}, and d_{NOD} variables summarize the scope of each state's UCC

[2] Bureau of the Census (1982a), Table D, page 10. The most recently available data refer to 1980.

[3] Bureau of the Census (1983), Table 3 of each state report. The most recent surveys report 1981 data.

[4] Bureau of the Census (1982b), p. 427. The data refer to 1981.

[5] *Ibid.*, p. 614 and Bureau of the Census (1983), Table 3 of each state report. The data refer to 1981.

[6] Bureau of the Census (1982b), p. 144. The data refer to 1981.

provisions; d_{UP}, d_{PA}, d_{TD} and d_{MD} describe the states' UDAP statutes:

d_{UCC} = UCC express and implied warranty provisions applicable to transactions in used goods;

d_{UCCOP} = UCC implied warranty of fitness for a particular purpose interpreted to mean fitness for an ordinary purpose;

d_{NEGD} = UCC implied warranty disclaimer allowed but only if pre-negotiated;

d_{NOD} = no UCC implied warranty disclaimer allowed;

d_{UUP} = UDAP prohibits unfair and/or unconscionable as well as deceptive acts and practices;

d_{PA} = private action is allowed under UDAP;

d_{TD} = UDAP provides for treble damage awards; and

d_{MD} = statutory minimum damages may be recovered under UDAP.

The variables take unit values for those states having UCC or UDAP legislation exhibiting the corresponding property. The final warranty variable d_{SE} (strong enforcement) takes a unit value for each state having consumer protection agencies with above median per capita budget and manpower resources and used car complaint activity. The appendix presents a formal discussion of these nine variables and argues why they collectively summarize the critical dimensions of warranty protection across states. The values of the nine variables are presented in the table to the appendix.

The independent probability that an individual car model will experience some defect or failure, $\pi(v)$, is set at .25. This estimate was derived from a telephone survey of dealers who reported that, on average, two to three cars out of every ten sold are returned for some repair work with costs borne by the dealer. The probability of failure undoubtedly varies across models. The model developed in section 1 is presently not rich enough to parametrically control for variation in π. Consequently, the sample of cars is divided into three groups based on the auto reliability ratings published by Consumer Union (1984).[7] Subsamples of cars with average (A), much worse than average (W), and much better than average (B) repair histories are formed. The subsample assignment of each car model is indicated in the last column of Table 1.

The joint probability of car failure and a successfully pursued claim under state

[7] Consumer Union (1984), pp. 340-56. Consumer Union differentiates five "frequency-of-repair" groups: average, better than average, much better than average, worse than average, and much worse than average. The "average" (A) group constructed for this study combines all cars assigned by Consumer Union to the average and above and below average groups.

warranty laws premultiplies the auto repair cost variable $w_R(s)$. The base national estimate of auto repair cost is the "typical auto repair bill" reported in the 1972-73 Consumer Expenditure Survey (CES) published by the Bureau of Labor Statistics, brought to its August 1983 level ($347) by the seasonally adjusted "maintenance and repair" CPI.[8] State variation in auto repair costs is generated by multiplying the base CES 1983 estimate by an auto repair labor cost index formed from a payroll per employee ratio for laborers in auto repair shops (SIC 753) as reported in *County Business Patterns*.[9]

3. Results

Tables of results are presented for three major divisions, one for an average range frequency-of-repair group of models, and one each for the two extremes of much better and much worse than average frequency-of-repair model categories. Each car make is allowed a model-specific intercept, but model effects are otherwise ruled out. Model slope dummies were tested and found insignificant once the three-part grouping was imposed.

The model tested is the linear approximation to the logit, expanded about the point $x(s) = 0$ (probability of enforced claim = ½). The results below indicate the approximation is poor in that the estimated probability for the strongest possible warranty state ($.5 + \frac{\text{SUM2}}{4}$ in the tables)[10] is outside the zero-one interval, though it is not statistically different from unity.

Our inability to separately identify $\pi(v)$ and $a(v)$ by model and any mismeasurement of the repair cost variable w_R are potential problem sources. To the extent the certainty equivalence assumption is inappropriate, some part of the effect of risk aversion may well be appearing in the warranty coefficients. Considerable work remains to be done. The eventual solution is the estimation of the logit rather than the approximation, but this involves time- and budget-consuming nonlinear estimation.

An additional complication in the econometric model is a correction for heteroskedasticity. The dependent variable is a retail transactions average price for model v in state s. The number of retail transactions in the sample differs by state, hence the error variance associated with the average price differs by state according to the formula $Z/N(s,v)$, where Z is the population variance for model v retail price (assumed identical for all v) and $N(v, s)$ is the number of observations of retail sales of model v recorded in state s. The appropriate cor-

[8] Bureau of Labor Statistics (1978). Table 1, p. 40 and Bureau of Labor Statistics (1983). Table 2, p. 11. The CES "typical auto repair bill" is defined as out-of-pocket expenses. Because it excludes all repair costs covered by insurance, it provides a good measure of both the financial incentive to a consumer considering a warranty claim against a used car dealer and the cost to be incurred by a dealer if the consumer's claim is successful.

[9] Bureau of the Census (1983), Table 3 of each state report. The most recent reports present data for 1981.

[10] *See* the derivation of the approximation to the logit on page 70.

rection for heteroskedasticity of this type is to multiply both sides of equation (2) by the square root of N(v,s). The results below incorporate this correction, effectively giving more weight to observations based on larger samples.

It is also quite possible that Z is not constant across models and, even more importantly, that there is correlation between the error terms for various models. The appropriate correction for this type of disturbance structure is Zellner's seemingly unrelated regression procedure. We have not thus far used it since it can make unrelated regression procedure. We have not thus far used since it can make uncovering specification problems very difficult. On this ground, our estimates are probably inefficient, but consistent.

The results reported in the tables are reasonably in accord with expectations and show that warranty variables do make some difference statistically. The magnitudes of the warranty effects are rather suspect at this stage of our investigation, for reasons given above. The results are more reasonable for the average and much above average range groups than for the much below average group. This and related grouping difficulties teach us that more work needs to be done to explain model effects. Also, only in that context can a successful attack be made on modeling the probability of failure.

For the average group, as for the others, much of the variation is explained by the model specific intercepts. The magnitude of these numbers indicates this largely reflects a uniform national wholesale price expectation. Local variation in wholesale price is insignificant, save for the much worse than average group. None of the other dealer cost terms is significant. The income terms are all insignificantly different from zero, but their negative sign is in accord with intuition suggesting that used cars are an inferior good. The competition term LPD is never significant.

The group of warranty terms are more interesting. For the average range group, the warranty slope terms as a group are statistically significant. The key elements in this are UCC, treble damages (TD) under UDAP, and the strong enforcement variable SE. In addition the education variable ED is marginally significant. Each of these coefficients has the expected positive sign.

The numerically largest and by far the most significant estimates are $\hat{B}_{UCC}$ and $\hat{B}_{SE}$. The importance of a state's enforcement efforts is not surprising and is reassuring. The same holds for B_{UCC}, a coefficient associated with a variable having a particularly wide interpretation. As explained in the appendix, a unit value for the UCC variable (46 states plus Washington, D.C.) suggests that buyers are protected by all three UCC warranties (its express warranty and two implied warranties) as well as the MM prohibition against limiting the duration of any UCC implied warranty to less than the term of the product's written warranty. The estimates suggest, however, that the marginal effect of UCCOP, NEGD, and NOD embellishments is inconsequential.

The importance of the basic UDAP statute prohibiting deceptive practices cannot be determined unambiguously from the results. A UDAP statute exists in every jurisdiction and therefore has its effect included, along with other unspecified ubiquitous effects, in the warranty constant B_0. What we can say is that the extension of basic UDAP statutes to cover unfair and/or unconscionable acts is too subtle an enhancement to make any statistically significant difference. Not so, however, for the treble damage penalty provision.

If the magnitudes of the warranty terms were taken seriously, the overall effect of going from the weakest possible state to the strongest possible state adds about \$237 to the price of a representative used car in a state with average repair costs. This is the product of the sum (SUM1 in Table 2) of the slopes (5.48), the assigned probability of failure (.25), the average repair cost (\$347), and ½, the scaling term required by equation (2). This figure is at least of plausible order of magnitude. It amounts to nearly 3.5 percent of the retail price of the representative used car in the average model group.

Neither the individual warranty coefficients nor their sums are significant in either the much above or much below average groups. This result is consistent with our expectation for the much above average group. The group's history of infrequent and trivial repairs indicates that the presence of warranties should have little effect on a dealer's marginal cost and therefore on a used car's retail price. The opposite, however, was expected for cars in the much below average group. We suspect that model-specific effects are creating specification bias. Note, this latter group was the only one for which the wholesale price coefficient was significant, suggesting that model-specific effects are quite important in this reliability class.

In future work several major refinements are advisable. First, it is important to do a better job of controlling for model-specific effects. The variation in price of gasoline over jurisdictions clearly alters the desirability of guzzlers vs. econoboxes in a way not yet controlled for in our model. Other features will no doubt also be important. Second, risk aversion should be incorporated into the model. Its importance should not be underestimated. If used car dealers are risk averse, the measured marginal cost of warranty provisions reported above may well exceed the dealers' true outlay for warrantied repairs. Third, it would be worthwhile to attempt to model the probability of failure parametrically. Fourth, both the probability of failure and the probability of enforced claim should be estimated for the true logit form. Fifth, we may finally reach a stage of confidence where application of the Zellner-efficient technique seems reasonable. Finally, if additional data could be found (on quantity demanded or possibly on private market prices), it might be possible to address the issue of the value consumers place on warranties. Given identification of the demand structure, it would be possible to undertake incidence analysis to answer the question whether mandated warranties are welfare improving.

Table 2

REGRESSION STATISTICS FOR AVERAGE FREQUENCY OF REPAIR SET

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
MODEL SPECIFIC INTERCEPTS			
F2	8691.12	472.586	18.3906
F7	3603.41	452.314	7.96660
F8	4839.09	451.991	10.7062
F12	4754.49	451.265	10.5359
F13	6477.32	459.611	14.0931
F14	6072.99	462.982	13.1171
F20	5078.66	452.651	11.2198
F26	3140.37	452.692	6.93710
F27	4165.29	452.008	9.21508
F3	10520.1	500.547	21.0172
F9	5565.15	455.208	12.2255
F10	3941.51	479.221	8.22483
F21	5942.07	457.681	12.9830
F28	4748.88	452.728	10.4895
F29	5767.46	458.247	12.5859
WHOLESALE PRICE	0.507262E-02	0.272112E-01	0.186417
ENERGY PRICE	-58.0303	42.8517	-1.35421
LABOR PRICE	-0.456973E-01	0.413736E-01	-1.10450
INC	-0.506711E-01	0.409077E-01	-1.23867
LPD	0.541304E-01	0.709742E-01	0.762677
WARRANTY PARAMETERS			
B0	-1.60984	6.29161	-0.255871
NOD	-0.855200	0.973471	-0.878506
NEGD	-0.371493	0.792148	-0.468969
UCC	3.60679	1.09641	3.28964
UCCOP	0.614640	0.739942	0.830659
UUP	0.146371	0.810009	0.180702
PA	-1.08411	1.92561	-0.562998
TD	1.32988	0.719597	1.84809
MD	-0.266768	0.655034	-0.407258
SE	2.36477	0.697598	3.38988
ED	0.121894	0.699469E-01	1.74267

R-squared = 0.969625 NOBS = 476

R-bar-squared (adj for DF) = 0.967504 Mean of Dependent Variable = 11395.2

Analysis of Warranty Effects: SUM1 = all but ED and constant
SUM2 = all

Right-hand Variable	Estimated Coefficient	Standard Error	T-Statistic
SUM1	5.48488	2.92863	1.87285
SUM2	3.99693	6.28597	0.635850

Table 3

Regression Statistics for Much Better Than Average Frequency of Repair Set

Right-Hand Variable	Estimated Coefficient	Standard Error	T-Statistic
Model Specific Intercepts			
F24	7937.94	1118.65	7.09603
F25	8929.25	1126.91	7.92369
F30	6448.15	1088.34	5.92473
F31	7482.57	1107.04	6.75905
F32	5960.64	1080.04	5.51890
F33	6424.79	1075.12	5.97587
F34	7275.55	1098.15	6.62525
F35	6722.12	1094.48	6.14186
F36	7595.93	1088.25	6.97995
Wholesale Price	-0.564616E-01	0.505174E-01	-1.11767
Energy Price	27.6546	110.270	0.250790
Labor Price	0.432324E-01	0.102507	0.421750
INC	-0.375551E-01	0.981201E-01	-0.382746
LPD	-0.251984E-01	0.184725	-0.136410
Warranty Parameters			
B0	-16.1059	14.2239	-1.13231
NOD	4.08582	2.38345	1.71425
NEGD	1.55998	2.13118	0.731980
UCC	0.913526	2.10263	0.434469
UCCOP	0.803027	1.79457	0.447475
UUP	0.575938	2.09904	0.274381
PA	-1.77821	6.61867	-0.268665
TD	-0.856147	1.52884	-0.559997
MD	-0.836791E-01	1.57316	-0.531917E-01
SE	0.767001E-01	1.47273	0.520803E-01
ED	0.370707E-02	0.158879	0.233327E-01

R-squared = 0.977769 NOBS = 231

R-bar-squared (adj for DF) = 0.975058 Mean of Dependent Variable = 16504.7

Analysis of Warranty Effects: SUM1 = all but ED and constant
SUM2 = all

Right-hand Variable	Estimated Coefficient	Standard Error	T-Statistic
SUM1	5.29696	9.10265	0.581914
SUM2	-10.8052	14.1984	-0.761017

Table 4

REGRESSION STATISTICS FOR MUCH WORSE THAN AVERAGE FREQUENCY OF REPAIR SET

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
MODEL SPECIFIC INTERCEPTS			
F1	4372.60	644.253	6.78709
F4	3264.72	622.974	5.24054
F5	3631.49	621.815	5.84014
F6	4846.05	669.676	7.23642
F11	3233.56	619.040	5.22351
F15	3208.43	607.372	5.28247
F18	4142.86	656.228	6.31314
F16	3759.00	624.854	6.01581
F17	3695.49	623.375	5.92820
WHOLESALE PRICE	0.194705	0.728284E-01	2.67348
ENERGY PRICE	69.0252	56.9529	1.21197
LABOR PRICE	-0.164304E-01	0.568330E-01	-0.289099
INC	-0.943782E-01	0.561015E-01	-1.68227
LPD	-0.111966	0.836717E-01	-1.33816
WARRANTY PARAMETERS			
B0	7.67495	7.70139	0.996567
NOD	1.32952	1.10175	1.20673
NEGD	0.554139	0.861152	0.643486
UCC	-0.484589	1.45057	-0.334069
UCCOP	-1.11982	0.879125	-1.27379
UUP	-0.196267	0.931884	-0.210614
PA	1.66271	3.20440	0.518885
TD	-1.08911	0.915465	-1.18968
MD	0.966949	0.842706	1.14743
SE	-0.756409	0.970684	-0.779254
ED	0.108018	0.960886E-01	1.12415

R-Squared = 0.968873 NOBS = 193

R-Bar-Squared (adj for DF) = 0.964213 Mean of Dependent Variable = 7984.82

Analysis of Warranty Effects: SUM1 = all but ED and constant
SUM2 = all

Right-hand Variable	Estimated Coefficient	Standard Error	T-Statistic
SUM1	0.867124	4.18984	0.206959
SUM2	8.65009	7.68338	1.12582

References

Bond, Eric (1982). "A Direct Test of the 'Lemons' Model: The Market for Used Pickup Trucks," *American Economic Review*, 72 (September), pp. 836-40.

Bureau of the Census, U.S. Department of Commerce (1982a). *Annual Survey of Manufactures, Fuels and Energy Consumed*. Washington, D.C.: U.S. Government Printing Office, October.

________________ (1982b). *Statistical Abstract of the United States, 1982-83*. Washington, D.C.: U.S. Government Printing Office, December.

________________ (1983). *County Business Patterns, 1981*. Washington, D.C.: U.S. Government Printing Office.

Bureau of Labor Statistics, U.S. Department of Labor (1978). "Consumer Expenditure Survey: Interview Survey, 1972-73," *BLS Bulletin*, 1997.

________________ (1983). *CPI Detailed Report, September 1983*. Washington, D.C.: U.S. Government Printing Office, November.

Consumer Union (1984). "1984 Buying Guide Issue," *Consumer Reports*, 48, No. 12 (December).

Appendix

Warranty Protection Under the Law

Consumers can seek relief for used car warranty violations under any of three consumer protection laws: the Magnuson-Moss Warranty Act (MM), Article 2 of the Uniform Commercial Code (UCC), and state unfair and deceptive acts or practices statutes (UDAP). Other legislation applies to used car transactions but only these three laws have direct implications for warranties.[1]

Consumer rights under these acts are not uniform across states. Not all states, for example, have enacted article 2 of the UCC and those that have adopted it have drafted laws that differ in important dimensions, especially as to the type of goods covered under the law and the conditions, if any, under which the seller may disclaim the buyer's warranty rights. Similarly, while all states have UDAP statutes, variability is substantial in the scope of covered practices, procedural matters, and the treatment of damages. Even MM, a federal statute, does not apply uniformly across jurisdictions. Consumers within a state are protected under the implied warranty liability provision of MM only to the extent that a UCC implied warranty exists under state law.

The objective of this abstract is to highlight those provisions of the consumer protection laws which create interstate variability in the economic severity of the statutes as applied to used cars. These differences, if truly important, should result in different economic behavior among sellers and ultimately in different transaction prices.

The following discussion is thus an economic description of the laws and is intended to be neither an exhaustive nor a legal treatment. Complete, legal descriptions of MM, UCC, and UDAP statutes appear in Rigg and Alpert (1982) and Sheldon (1982a), two texts published by the National Consumer Law Center. The material presented below draws heavily from these volumes.

MAGNUSON-MOSS. The Magnuson-Moss Warranty Act was passed by Congress on July 4, 1975.[2] The law applies to transactions in consumer goods, not services. The transaction can involve new or used products but the product must have been initially manufactured or substantially modified after July 4, 1975. All used cars examined in this research project were manufactured after 1976 and therefore qualify for MM protection. The Act does not require that a product

---

[1] Motor vehicle disclosure, safety inspection, and odometer disclosure laws also protect buyers of used cars. Disclosure regulations regarding the vehicle's prior use, accident record, flood damage, repair history, and odometer reading are designed to ensure that the buyer and the seller possess the same information at the point of sale. These regulations, however, do not enhance the buyer's warranty protection. State safety inspection rules determine the vehicle's conformity with minimum safety requirements but, as Rigg and Alpert (1982) argue, with the possible exception of Wisconsin's very stringent inspection law which sets minimum merchantability standards, "state standards are not intended to be standards of minimum quality" (p. 292). Also, *see* note 64 below.

[2] 15 U.S.C. sec. 2301-3213 (Supp. 1975).

be warranted, that a warranty be in writing, or that it have a prescribed duration. It only provides that if the sale of a product is accompanied by a "written" warranty, then (1) the warrantor must "fully and conspicuously disclose in simple and readily understood language the terms and conditions of such warranty"[3] and (2) certain prohibitions will apply. The FTC is specifically designated as the source for interpretation of the Act and as its enforcement agency. The FTC may bring action as can consumers who, individually or as members of a class, can sue for damages and attorney fees.

The Act establishes two important prohibitions. First, MM disallows the disclaimer of UCC implied warranties when a "written" warranty is given.[4] Second, the seller may not tie the consumer's purchase of another product or service to the issuance of its written or implied warranty. Only the former is of relevance to this study and it, as Rigg and Alpert (1982) point out, "may be the most significant provision of the Act":[5]

> Anyone who warrants a product...within the Act's scope is prohibited from disclaiming or modifying the implied warranties of merchantability and fitness for a particular purpose when they arise as a matter of state law. The Act does not, however, impose implied warranty liability on a party who gives a written warranty but who does not, under state law, make an implied warranty.[6]

The important implication is that, although MM is a federal statute, the application of its disclaimer prohibition is not uniform across states. While its "simple and readily understood" disclosure requirements apply to transactions in all jurisdictions, its implied warranty disclaimer prohibition is operative for transactions in used cars only in those states both having adopted the UCC and interpreting it as applicable to sales of used goods.[7]

Another condition limiting the applicability of the MM disclaimer prohibition is that MM covers transactions only when a "written" warranty is given, where the word "written" has a particularly narrow interpretation. In brief, a written warranty under MM means

(a) "any written affirmation of fact or written promise that...such material or workmanship is defect free or will meet a specified level of performance over a specified period of time, or

(b) "any undertaking in writing in connection with the sale by a supplier of a consumer product to refund, repair, replace, or take other remedial action with respect to such product in the event that such product fails to meet the specifications set forth in the undertaking."[8]

A warranty promising a defect-free product seldom accompanies the sale of

[3] *Ibid.*, sec. 2302(a).
[4] *Ibid.*, sec. 2308(a).
[5] Rigg and Alpert (1982), p. 227.
[6] *Ibid.*
[7] This has important implications for the interpretation of the "UCC" variable described below.
[8] 15 U.S.C. sec. 2301(6) (Supp. 1975).

a used car. More typically, the buyer's bill of sale includes, if anything, a statement that the car is warranted "for 30 days" or "for 30 days or 1000 miles whichever comes first" or perhaps is covered by a "30-day 50-50 warranty on parts and labor." Rigg and Alpert (1982) suggest that these phrases arguably fall within the section 2301(6b) definition of "written warranty."[9] Any lesser statement, however, most likely removes the transaction from MM protection.[10]

For purposes of this study, MM is assumed to apply to all used car transactions. The dealerships responding to the NADA data requests are generally medium to large size dealerships which sell both new and used cars. These dealers typically offer one of the above variants of the 30-day warranty which, it appears, can be interpreted as a "written" warranty under MM.[11]

The MM disclaimer prohibition, however, has only a brief duration in used car transactions. For limited (as opposed to 'full") warranties, MM section 2308(b) permits "implied warranties (to) be limited in duration to the duration of a written warranty of reasonable duration."[12] In effect, the MM disclaimer prohibition on a used car transaction accompanied by a limited written 30-day warranty is operative only within that 30-day window.[13] Presuming the seller has taken advantage of the MM provision that (s)he may restrict the duration of the implied warranty, the MM no disclaimer provision is not available to the buyer beginning on the 31st day after the sale. At that point, the buyer will have to rely on the specific implied warranty disclaimer provisions of his/her state's UCC statute.

UNIFORM COMMERCIAL CODE. The current version of the UCC is dated 1967. It creates certain buyers' rights and sellers' obligations. Like MM, Acticle 2 of the UCC applies only to "transactions in goods." Commercial transactions such as real estate, insurance, and services are excluded. Importantly, the definition of "goods" does not limit the term to new goods.[14] However, the presence of the word "transactions" in section 2-102 has created some confusion. Other

[9] Rigg and Albert (1982), p. 230.

[10] It should be noted that damages and attorney fees are not limited to written warranties under MM.

[11] The authors verified by telephone survey that some form of a 30-day warranty is the typical warranty given by medium to large size used car dealers.

[12] 15 U.S.C. sec. 2308(b) (Supp. 1975). Rigg and Alpert (1982) point out that the section 2308(b) language is confusing since there is no duration associated with an implied warranty. Implied warranties apply only at the time of sale. "The only logical interpretation of this reference to duration then is that duration refers to the time during which the buyer must notify the seller of the defect" (Rigg and Alpert (1982), p. 232).

[13] That the typical written used car warranty is a "limited" (as opposed to "full") warranty is important. MM allows sellers to restrict the duration of an implied warranty associated with a limited warranty, but section 2304(a)(2) expressly disallows any limitation on the duration of an implied warranty when a full warranty is given. If used auto warranties were considered full warranties, as defined by Rigg and Alpert (1982) p. 234, the UCC and NOD variables defined below would be perfectly collinear.

[14] UCC sec. 2-105(1) defines "goods" as "all things which are made at the time of identification to the contract for sale."

sections of Article 2 refer not to transactions but to "sales," "sellers," and "buyers." Clearly all sales of goods are covered by the UCC, but "transactions" is a broader concept than "sales."

No confusion arises, however, regarding the Code's distinction between the notions of "seller" and "merchant." The former includes private parties while the latter does not.[15] The distinction is crucial since important UCC sections, including the implied warranty of merchantability, apply only to sales by merchants. Fortunately, there is no ambiguity regarding the full applicability of the UCC to the used car transactions examined in this study. All are "transactions in goods" as required by Article 2. Furthermore, all recorded auto transactions used here are sales by merchants.

Not all states have adopted Article 2 and its applicability to used car purchases is questionable in some. Court cases in Alabama (1976), Georgia (1971), and Texas (1973) suggest that used goods are not covered by Article 2.[16] The state of Louisiana, moreover, has not adopted Article 2 as statutory law. Consequently, the only recourse available to consumers in these jurisdictions may be to sue for negligence which, of course, requires proof of seller misconduct. In UCC jurisdictions, in contrast, the seller may be liable regardless of fault. The extent of liability, however, is limited to contract damages. Neither punitive damages nor attorney fees are recoverable under Article 2.[17]

The UCC establishes three major types of warranties: express warranties, the implied warranty of merchantability, and the implied warranty of fitness for a particular purpose. All apply to used car transactions.

Article 2 establishes a number of types of express warranties. The form important to transactions in used cars is the "express warranty by affirmation of fact or promise." According to Rigg and Alpert (1982), "the four elements of an express warranty by affirmation of fact or promise are: (1) any affirmation of fact or promise, (2) made by the seller to the buyer, (3) which relates to the goods, and (4) becomes part of the basis of the bargain."[18] The warranty can be written, oral, or inferred from advertising, pictures, or product labels. The typical 30-day warranty accompanying a used car sale is such an express warranty.[19]

UCC section 2-316(1) prohibits the subsequent disclaimer of express warranties:

[15] UCC sec. 1-204(1) states: "Merchant means a person who deals in goods of the kind or otherwise by his occupation holds himself out as having knowledge or skill peculiar to the practices or goods involved in the transaction."

[16] *Trak, Inc. v. Tidmore*, 531 So. 2d 275, 19 UCC REP 92 (Ala 1976); *General Motoral Corp. v. Halco Instruments, Inc.*, 124 Ga. App. 630, 184 S.E. 2d 619, 9 UCC REP, 1193 (1971); *Chaq Oil Co. v. Gardner Mach. Corp.*, 500 S.W. 2d 877, 13 UCC REP. 806 (Tex. Civ. App. 1973).

[17] In addition, "Maryland, Massachusetts, and West Virginia have added provisions to their versions of the UCC specifically restricting the ability of warrantors to limit or modify remedies for breach of warranty in sales of consumer goods" (Rigg and Alpert (1982), p. 223). This latter dimension of state variability is judged to be too small to warrant an additional indicator variable.

[18] Rigg and Alpert (1982), p. 29.

[19] *Ibid.*

Words or conduct relevant to the creation of an express warranty and words or conduct tending to negate or limit warranty shall be construed wherever reasonable as consistent with each other; but subject to the provisions of this Article on parol or extrinsic evidence, negation or limitation is inoperative to the extent that such construction is unreasonable.[20]

The section refutes the standing of "as is" sales.

The important implication for this project is that the 30-day warranty on used car sales is considered an express warranty under the UCC and therefore cannot be disclaimed in those jurisdictions in which the UCC is applicable to transactions in used goods—*i.e.*, in all jurisdictions but Alabama, Georgia, Louisiana, and Texas.

The implied warranty of merchantability arises under section 2-314 and is perhaps "the most important warranty in the UCC."[21] As Rigg and Alpert (1982) note:

It...guarantees a basic standard of quality. The premise is that a buyer has a reasonable expectation that professional sellers will sell goods of at least minimally adequate quality. The law implies a warranty to protect that expectation interest. The warranty arises without any language or conduct by the seller relating to warranties, the goods, or promises; reliance by the buyer is unnecessary....

Merchantability does not depend on the seller's personal fault. The implied warranty of merchantability is a "no-fault" or strict liability concept; the warranty is breached, when the goods are of insufficient quality, regardless of the interest, knowledge, or thoroughness of the seller....(A) merchant is not absolved of liability if the defect was undiscoverable or unpreventable.[22]

Unless disclaimers are allowed in a jurisdiction, it is difficult not to establish an implied warranty of merchantability. All that is required is that a sale of a good occur and that the seller be a merchant: "Unless excluded or modified (disclaimed), a warranty that the goods shall be merchantable is implied in a contract for their sale if the seller is a merchant with respect to goods of that kind."[23] It follows that, unless disclaimed, an implied warranty of merchantability is certainly created in any used car transaction between a consumer and an automobile dealer.[24]

This latter statement holds in all jurisdictions except Alabama, Georgia, Louisiana, and Texas. The indicator variable UCC in the Appendix Table is assigned a unit value in all but these four states. The unit value signifies that the UCC express warranty and implied warranty of merchantability hold in these jurisdictions and therefore that MM provisions regarding the disclaimer of implied war-

[20] Since the used car transactions examined in this study typically occur under written 30-day warranties, the parol or extrinsic evidence provisions in section 2-316(1) are relatively unimportant.

[21] Rigg and Alpert (1982), p. 45.

[22] *Ibid.* Also *see* UCC comment 1 to section 2-313.

[23] UCC sec. 2-314(1).

[24] UCC sec. 2-314(2) lists the six criteria defining merchantability. The criterion most applicable to auto transactions is that the "goods to be merchantable must be at least such as are fit for the ordinary purposes for which such goods are used."

ranties apply as well.[25]

Article 2 of the Code also establishes a second wholly independent implied warranty, the implied warranty of fitness for a particular purpose. This warranty takes effect when the seller has reason both to know that the buyer is purchasing the good for a particular purpose and to know that the buyer is relying on the seller to provide a good that will meet that purpose.[26] The sale and therefore the implied warranty promise that the goods at time of sale are fit and suitable to serve the buyer's particular purpose. The warranty, like its companion implied warranty of merchantability, promises that the goods are suitable at delivery. Defects occurring later are not covered.

The implied warranty of fitness for a particular purpose, however, has some marked advantages over the implied warranty of merchantability:

> First, the implied particular purpose warranty is not limited to transactions with merchants. Second, a higher standard is placed on the goods; the implied particular purpose warranty may be breached if the goods do not satisfy the buyer's particular purpose, even though the goods are fit for their ordinary purpose. Third, breach of the particular purpose warranty is often easier to prove than breach of the implied warranty of merchantability.[27]

Since the used car transactions investigated in this study reflect purchases made from merchants, the first advantage is moot. The second and third, however, are substantively important for differentiating the stringency of UCC provisions across jurisdictions.

Rigg and Alpert (1982) elaborate the particulars of these latter advantages:

> Although proof of breach is in some respects similar to that required for the implied warranty of merchantability, the implied particular purpose warranty has important proof advantages over the implied warranty of merchantability. Goods may be unsuitable for the buyer's particular job even though they are not defective and operate exactly as designed....
>
> A second proof advantage...is that it can be a performance warranty. A buyer is protected by an implied particular purpose warranty if the goods do not perform, that is, do the buyer's job, as impliedly promised by the seller. The buyer need not show that the problem existed at delivery, other than by testifying to proper use. Importantly, facts giving rise to an implied warranty of fitness also create a future performance express warranty, which cannot be disclaimed.[28]

Given these important advantages, the principal issue is the determination of the boundary between a good's "ordinary" and "particular" purposes. Fitness of goods for ordinary purposes is the domain of section 2-314, the implied warranty of merchantability. Fitness for a particular purpose is covered by section 2-315, the implied warranty of fitness for a particular purpose. Those jurisdictions interpreting "particular purposes" to mean general or ordinary purposes

[25] *See* the above discussion of MM provisions limiting the seller's ability to disclaim an implied warranty when a written warranty is present.

[26] UCC sec. 2-315.

[27] Rigg and Alpert (1982), p. 53.

[28] *Ibid.*, p. 126.

are effectively providing buyers with a much stronger set of UCC arguments. For used cars the relevant question is whether purchasing an auto for travel on normal roads and highways is a particular purpose under section 2-315 as well as an ordinary purpose under section 2-314.

Rigg and Alpert (1982) list a number of case law instances where purchases for unambiguously ordinary purposes (including a used car for normal transportation) have been interpreted by state courts as falling within section 2-315 protection.[29] The represented states include Arkansas, Colorado, Illinois, Indiana, Maryland, Minnesota, New Hampshire, North Carolina, Oklahoma, and West Virginia.[30] As Rigg and Alpert (1982) note:

> This line of cases is based on the theory that the key to section 2-315 is the seller's affirmative conduct of selecting or furnishing goods to satisfy the buyer, with reason to know the buyer's needs and the buyer's reliance. This affirmative conduct to induce a sale and the buyer's reliance exist regardless of whether the buyer's need is characterized as particular or ordinary.[31]

The important conclusion is that sellers in these states have less protection under the UCC than do merchants in other UCC jurisdictions. As a result, the indicator variable UCCOP ("UCC ordinary purpose") takes a unit value in the Appendix Table for these ten states. All other states (except Alabama, Georgia, Louisiana, and Texas) are assumed to interpret "particular purpose" more narrowly. Whatever impact the more limited application of section 2-315 has on used car transactions will be captured in the more general UCC variable. The coefficient associated with UCCOP measures the marginal effect of broadening the application of the implied warranty of fitness for a particular purpose to fitness for ordinary purposes.

Not only does the interpretation of "particular purpose" vary across states but UCC disclaimer provisions vary as well. Though an express warranty under the UCC cannot be disclaimed, section 2-316 allows disclaimers and modifications of UCC implied warranties. Because of the basic UCC policy disfavoring disclaimers, however, a disclaimer typically will be successful only if a number of important conditions are met. A partial list of the more important restrictions applying to used car transactions includes the following: (1) the implied warran-

---

[29] The following case references are found in Rigg and Alpert (1982), p. 55, notes 21 and 28 as supplemented in Rigg and Alpert (1983). *Stickney v. Fairfield's Motors, Inc.*, 9 UCC REP. 236 (N.H. Super. Ct. 1970). *See also Karczewski v. Ford Motor Co.*, 382 F.Supp. 1346, 15 UCC REP. 605 (N.D. Ind. 1974) ("particular purpose of a passenger automobile is to drive on the public streets and highways safely without uncontrolled and unsafe behavior"). *Nelson v. Wilkins Dodge, Inc.*, 256 N.W. 2d 472, 21 UCC REP. 1001 (Minn. 1977). *Thomas v. Ford Motor Credit Co.*, 48 Md. App. 617, 429 A. 2d 277, 31 UCC REP. 1265 (1981).

[30] Cases reaching similar conclusions can also be found in the case law in Massachusetts, Mississippi, New York, and Tennessee but subsequent cases by higher state courts interpreted "particular purpose" as meaning a purpose that is special in some way. *See* the case law referenced in Rigg and Alpert (1982), pp. 55-57, n. 19-41 and Rigg and Alpert (1983), pp. 15-16, n. 21, 28.1, 29.1, and 30.

[31] Rigg and Alpert (1982), p. 55.

ty disclaimer must be conspicuous,[32] (2) the disclaimer must be made available to the buyer before the contract is signed, (3) a disclaimer of the implied warranty of merchantability typically must use the word "merchantability,"[33] (4) an "as is" disclaimer will apply only if the buyer can be shown to be unmistakably aware that no implied warranty is being offered, (5) the disclaimer cannot be unconscionable,[34] and (6) the disclaimer cannot shield the seller from liability for lack of good faith.

Most states simply adopt section 2-316 with its list of conditions limiting the scope of implied warranty disclaimers. Others, however, impose a more stringent condition requiring that the contract disclaimer either be actually negotiated with the buyer prior to the sale or be specifically brought to the buyer's attention (rather than simply made available to the buyer) prior to the sale. Rigg and Alpert (1982) point to the case law or statutes revealing this tightening of disclaimer provisions in sixteen states: Arizona, California, Colorado, Connecticut, Florida, Illinois, Indiana, Iowa, New Hampshire, New Jersey, North Dakota, Ohio, Oregon, Utah, Washington, and Wisconsin.[35] In addition, six jurisdictions—Kansas, Maryland, Massachusetts, West Virginia, New York City, and Washington, D.C.—prohibit disclaimer altogether.[36] Interestingly, Mississippi has chosen not to enact section 2-316, intending, according to Rigg and Alpert (1982), "to prohibit the exclusion of implied warranty authority by section 2-316."[37]

The interstate variability in disclaimer provisions is modeled through the NEGD ("negotiated disclaimer") and NOD ("no disclaimer") variables in the Appendix Table. The former variable takes a unit value for the sixteen states requiring negotiated disclaimer. Unit values are assigned to the latter for Kansas, Maryland, Massachusetts, Mississippi, New York, West Virginia, and Washington, D.C.[38] Zero values appear elsewhere.

The four UCC related columns in the Appendix Table provide a useful summary of warranty protection under Magnuson-Moss and the Uniform Commercial Code. Those jurisdictions (4) with zero entries in all columns have neither MM nor UCC warranty protection for used car transactions. Buyers in jurisdic-

[32] Comment 1 to UCC sec. 2-316 states that this requirement "seeks to protect a buyer from unexpected and unbargained language of disclaimer by...permitting the exclusion of implied warranties only by conspicuous language or other circumstances which protect the buyer from surprise."

[33] Exceptions are noted in Rigg and Alpert (1982), p. 76.

[34] A disclaimer will be held to be unconscionable if, *e.g.*, the buyer has only a limited understanding of English. *See* Rigg and Alpert (1982), pp. 77-78.

[35] *Ibid.*, p. 75.

[36] Kansas statute sec. 50-639(1) (1976); Maryland Com. Law Code Ann. sec. 2.316.1 (1975); Massachusetts Gen. Laws Ann. Ch. 106, sec. 2.316A (west 1973); West Virginia Code sec. 46A-6-107 (1980); New York City Ad. Code ch. 32, art. 19, sec. 8(a)(3) (1962); and Washington, D.C., Code Ann. sec. 28:2-316.1 (1982). Maine has a similar law though it applies only to new cars. Used car sales in Maine are exempt under the Used Car Information Act.

[37] Rigg and Alpert (1982), p. 80.

[38] Though New York state does not prohibit disclaimer, the state is assigned a unit value for the NOD variable since it is assumed that most used cars in the state are purchased in New York City, a jurisdiction that disallows any disclaimer.

tions (47) with "ones" in the UCC column have available a number of protective statutes. MM requires that a 30-day written warranty cannot subsequently be disclaimed and prohibits a seller from limiting the duration of a UCC implied warranty to less than the term of the limited written warranty. Furthermore, all three UCC warranties (the express warranty and the two implied warranties) are applicable, though "particular purpose" is not equated with "ordinary purpose." Consumers in those jurisdictions (10) with a unit value in the UCCOP column additionally have the advantage of having the implied warranty of fitness for particular purpose applicable to goods purchased for purely ordinary purposes. Unit values in the NEGD (16) and NOD (7) columns identify jurisdictions that, respectively, require prenegotiation of implied warranty disclaimers or disallow disclaimers altogether.

It is clear from the table that UCC and MM protection varies considerably across states. Alabama, Georgia, Louisiana, and Texas represent one extreme; Maryland and West Virginia identify the other. Used car consumers in the former set of states have neither MM nor UCC warranty protection. The latter two states not only disallow any disclaimer of express or implied warranties but further permit consumers purchasing used cars for general transportation to sue sellers violating implied warranties under either the UCC implied warranty of merchantability or its implied warranty of particular purpose.

UNFAIR AND DECEPTIVE ACTS AND PRACTICES. State UDAP laws provide an additional array of possible remedies for "marketplace misconduct and consumer abuse."[39] Most statutes contain the FTC Act section 5(a)(1) language prohibiting "unfair or deceptive acts or practices," explaining why UDAPs are commonly referred to as "mini-FTC Acts."[40]

Individual state laws have been constructed so as to allow broad and flexible interpretations in response to new forms of seller misconduct. This not only leads to substantial variation in UDAP provisions across states but also makes UDAP a unique tool for consumer protection. Where a practice has not been previously banned, regulated, or otherwise recognized as abusive, UDAP can be used to bring that practice under such control. "Almost any abusive business practice aimed at consumers is at least arguably a UDAP violation, unless the trade practice falls clearly outside the scope of the statute."[41]

Many UDAPs identify particular trade practices as per se violations though other unfair, unconscionable, and/or deceptive practices are typically proscribed in more general terms. What differentiates the UDAP from common law fraud is that most UDAPs do not require "consumer reliance, damage, or even actual deception" and none requires the "proof of seller's fraudulent intent or

[39] Sheldon (1982a), p. 3.

[40] Sheldon (1982a), p. 3 notes that the "mini-FTC label is only precise for those statutes that parallel the FTC Act and prohibit 'unfair methods of competition and unfair or deceptive acts or practices.' "

[41] *Ibid.*

knowledge."[42]

Sheldon (1982a) describes the set of model UDAP Acts on which individual state laws are based.[43] The FTC Act forms the primary model. Three states, for example, have laws modeled directly on the FTC Act. Twenty-four others follow the 1967 Unfair Trade Practices and Consumer Protection Law (amended in 1970) which itself was drafted by the FTC and is based largely on the FTC Act. The most popular version (15 states) is patterned directly on the FTC Act prohibition of unfair methods of competition and unfair or deceptive acts and practices. An alternative form (9 states) enumerates thirteen specific offensive practices but typically includes a general prohibition against any other unfair or deceptive practice.

The Uniform Deceptive Trade Practices Act, first adopted in 1964 by the National Conference of Commissioners on Uniform State Laws is "presently law in seven states and forms the basis for UDAP statutes in six other states."[44] The act originally was intended for applications in commercial transactions, but its language has been interpreted so broadly that it has had application in consumer transactions. The act identifies eleven per se deceptive practices and includes a general prohibition against engaging in conduct which "create(s) a likelihood of confusion or of misunderstanding."[45]

Three states have adopted the Uniform Consumer Sales Practices Act issued in 1971 by the National Conference on Uniform Laws and the American Bar Association. It applies only to consumer transactions and prohibits unconscionable and deceptive acts or practices, including a particular list of practices.

Seven states have their own Consumer Fraud Acts. They typically proscribe "deception, fraud, false pretense, false promise, misrepresentation or knowing concealment, suppression or omission of any material fact with intent that others rely."[46]

All states and the District of Columbia have enacted one or more of these UDAP models though amendments to the basic models are not uncommon. Most laws were adopted in the period spanning the mid-1960s to the mid-1970s, though the range of years over which individual states enacted such laws is substantial. Utah appears to have been the first state proscribing "unfair methods of competition" in an Act passed in 1937, thus predating even the 1938 Wheeler Lea Amendments to the FTC Act.[47] Alabama was the last, enacting a UDAP statute in April 1981.[48]

The language of the state laws differs in important respects. Some preserve the most general FTC language, prohibiting all "unfair or deceptive acts and

[42] *Ibid.*
[43] *See* Sheldon (1982a), pp. 4-5 for a full description.
[44] *Ibid.*, p. 5.
[45] *Ibid.*
[46] *Ibid.*
[47] Utah Code Ann. sec. 13-2-11.
[48] Alabama Code 8-19-5

practices." Others itemize specific prohibited practices. Some exempt specific practices including unknowing acts in debt-collection activities or in credit, realty, and landlord-tenant transactions. Some exempt particular trades, most usually utilities, printers, publishers, and broadcasters. Others exempt transactions not involving final consumers. Importantly, however, all apply to consumer transactions. None exempts used car sales.

The state laws vary in substantive dimensions to be discussed below, but all are designed to combat abuses in the marketplace. Deceptive practices are prohibited in every state.[49] Most practices of concern in used car transactions typically fall within the definition of deception. In particular, failure to disclose material facts is generally deceptive. State courts, for example, have found deceptive "a warranty failing to disclose that significant costs for parts would be charged,...the failure to disclose a defective car engine at the time of the car's sale, the failure to disclose...that the vehicle was a racing car,....a car dealer misrepresent(ing) a car's condition, mileage, and warranty,....misrepresent(ing) an express warranty, such as claiming a warranty is unconditional when there are conditions specified,...disclaim(ing) an oral warranty with an 'as is' warranty,....and fail(ing) to comply with offered or implied warranties."[50] More generally, a number of FTC cases concluded: "To tell less than the whole truth is a well known method of deception."[51] State laws typically follow suit.[52]

Thirty-four states go much further, prohibiting unfair and/or unconscionable conduct as well as deceptive practices:

> An unfair practice need not be deceptive, but can apply to contexts where the consumer is taken advantage of through the seller's superior position or oppressive sales methods. Unfairness is a dynamic concept and offers opportunities to prohibit various abuses never touched by common law fraud....
>
> Unconscionability is a broader concept than deception, somewhat akin to unfairness.[53]

Unfair practices are proscribed in 29 states. Twelve prohibit unconscionable seller conduct.

The standards of unfairness or unconscionability have particularly wide application to used car transactions. The standards can become operative simply because of the "imbalances of knowledge" between the buyer and seller regard-

[49] Sheldon (1982a) notes that Indiana, Oklahoma, and Washington, D.C., statutes "enumerate specific deceptive practices but do not prohibit deception more generally" (p. 6, n. 25). Reviewing the statutes, however, reveals that the itemized practices include all those typically relevant to auto transactions.

[50] Sheldon (1982a), pp. 32, 40, and 56.

[51] The following FTC case law references are found in Sheldon (1982a), p. 31, n. 62: *P. Lorillard Co. v. FTC*, 186 F.2d 42, 58 (4th Cir. 1950); *Benrus Watch Co. v. FTC*, 352 F.2d 313 (8th Cir. 1965), cert. denied, 384 U.S. 939 (1966); *Bennett v. FTC*, 200 F.2d 362 (D.C. Cir. 1952); *Bockenstette v. FTC*, 134 F.2d 369 (10th Cir. 1943).

[52] *See* Sheldon (1982a), p. 31 for examples.

[53] *Ibid.*, p. 6.

ing product attributes or contract terms. Deception, intended or not, is not required. This broad standard of unfairness follows from the Supreme Court decision in *FTC v. Sperry and Hutchinson Company* in which the Court affirmed the FTC criteria for determining whether a practice is "unfair." The FTC criteria especially relevant to used car sales require determining:

(a) whether the practice is immoral, unethical, oppressive, or unscrupulous and

(b) whether the practice offends public policy. Is it within at least the penumbra of some common law, statutory, or other established concept of fairness?[54]

State courts, according to Sheldon (1982a), give content to the Supreme Court directive by looking to section 2-302 of the UCC. The official comment to the section announces the following test for unfairness:

The basic test is whether, in light of the general commercial background and the commercial needs of the particular trade or case, the clauses involved are so one-sided as to be unconscionable at the time of making of the contract....The principle is one of the prevention of oppression and unfair surprise... and not of disturbance of allocation of risks because of superior bargaining power.

This UCC authority has been relied upon in a number of UDAP cases to effectively prevent warranty disclaimer or modification. Examples include "unfair surprise resulting from consumer ignorance and seller guile, such as the consumer's lack of education and seller's use of virtually incomprehensible fine-print standard-form contract provisions,...the seller binding the buyer to additional written terms after the contract is signed, switching contract documents at the last moment to include non-negotiated, one-sided terms, (and) sellers' use of one-sided contract terms that are standardized throughout the industry giving the consumer no opportunity to negotiate."[55] It is important to emphasize that this broad application of UDAP holds only in those states prohibiting unfair or unconscionable, not just deceptive, practices.

The list of 34 states having UDAPs with such expansive language is identified in the Appendix Table in the column headed UUP ("unfair or unconscionable practices"). The variable UUP is assigned a unit value for these 34 states, indicating that the scope of UDAP applicability is much broader in these states than in the remaining 17 jurisdictions. A zero value for UUP, note, does not imply the absence of a UDAP statute but does identify the corresponding jurisdiction as one limiting the application of its UDAP statute to deceptive practices.

UDAP statutes are state laws enforced by state agencies. The state attorney general's office typically has enforcement authority, though it is not uncommon to find local prosecutors possessing this authority as well.[56] State agencies often

[54] *Ibid.*, p. 32.

[55] *Ibid.*, p. 35.

[56] "In 20 states, local prosecutors, such as county or district attorneys, also have enforcement authority" (Sheldon (1982a), p. 7).

mediate complaints but can seek judicial remedies as well as injunctions and cease-and-desist orders. Thus, consumers can rely on state actions on their behalf.

What most differentiates consumer actions across states, however, is the right to private action and, subsequently, to injunctions and/or to remedies beyond actual damages. Most states authorize a private UDAP right of action. The only exceptions are Arkansas, Iowa, Nevada, and North Dakota.[57] Sellers in private action states face an effectively more stringent UDAP statute for three reasons. The seller not only faces the very real possibility of private damage action but also, should its defense fail, will at a minimum be liable for the buyer's actual damages and attorney fees.[58] In addition, there is the non-trivial threat of private injunctive action:

> One of the potentially most effective UDAP remedies against widespread marketplace misconduct is for a private individual to seek a court ordered injunction preventing the seller from engaging in specified conduct in the future. A merchant may treat occasional damage awards, even if trebled or increased with punitive damages, as an acceptable cost of business, not deterring future misconduct. But a properly framed injunction can eliminate the seller's use of the challenged practice against all future customers.[59]

It appears then that the right to private action substantially strengthens a state's UDAP legislation. Consequently, the PA ("private action") variable in the Appendix Table takes a unit value for the District of Columbia and for the 46 states providing consumers with the private action option. It is interesting to note that those four states without a right to private action also exhibit weak legislative provisions in other consumer protection areas as well.

Remedies in private action jurisdictions often go beyond the award of contract damages and attorney fees. Treble damages are available to successful plaintiffs in nineteen jurisdictions. Statutory minimum damages ranging from $25 to $2000 are authorized in nineteen states, with the most common statutory award falling in the $100 to $200 range.[60] The two sets of jurisdictions providing access to these respective private damage remedies are identified in the Appendix Table with unit values for either the TD ("treble damage") or MD

[57] Arizona's UDAP statute does not explicitly authorize a private UDAP right of action. However, the state's Supreme Court ruled in *Sellinger v. Freeway Mobile Home Sales, Inc.*, that while the state's UDAP does not contain explicit language granting a private right of action, such a right is granted inferentially (Sheldon (1982a), pp. 73-74).

[58] According to Sheldon (1982a) p. 7, "virtually every state UDAP statute with a private right of action authorizes the award of attorney fees."

[59] Sheldon (1982a), p. 92. Private injunctive actions are explicitly authorized in 30 states and, according to Sheldon (1982a), p. 93, may also be available in other states granting private damages.

[60] *Ibid.*, p. 7 and Sheldon (1983), p. 1. Sheldon (1982a), p. 7 also notes that "nine UDAP statutes explicitly authorize punitive damages." Interstate differences in allowances for punitive damages are not considered in this study because, as Sheldon (1982a), p. 90 later notes: "Punitive damages are appropriate when an act involves a particularly aggravated disregard for the rights of the victim or a grievous violation of societal interests." Should this definition apply to a used car transaction, the offending practice arguably could be reached under the common law, an ubiquitous consumer protection.

("minimum damage") variables.

The importance of the minimum damage provision should not be underestimated. As Sheldon (1982a) points out:

> Statutory provisions for minimum damages allow consumers to recover a specified damage amount based solely on evidence that the seller has violated the statute. Since a tendency or capacity to deceive, and not actual deception, may be sufficient to prove deception, actual damage may not otherwise be an element of a UDAP action....[61]

Consequently, not only can a plaintiff recover minimum damage that in fact exceeds his/her actual damage but the burden of proof required to obtain the damage award is considerably lower in minimum damage states relative to other private action jurisdictions.

Some states, however, do require that claims for statutory minimum damages first demonstrate that the plaintiff suffered some damage, however minimal. Sheldon (1982a), however, notes:

> Statutory requirements that the consumer suffer loss of money or property do not specify the amount of damage that must be proved. Consumer litigants should be able to meet this requirement by proving minimal injury. Thus litigants have successfully alleged that a long distance telephone call was sufficient damage to allow an action to proceed that was based on a creditor's initiating collection action in an inconvenient forum.[62]

Application to unfair practices, availability of the right to private action, and provisions for treble and/or minimum statutory damages summarize the important dimensions differentiating UDAPs across states. The base case, however, is not unimportant. All states and the District of Columbia have UDAPs. All UDAP statutes prohibit deceptive practices and authorize some state agency to proceed against offending merchants. Whatever effect these ubiquitous UDAP provisions have on seller behavior is captured in the constant term B_0 defined in the text.

The placement of zeros and ones in the UUP, PA, TD, and MD columns in the Appendix Table identifies the combination of UDAP embellishments, if any, available to consumers in each jurisdiction. The interstate variation is substantial. Most jurisdictions have enriched the basic UDAP provisions in at least one dimension, but a handful of states deserve note. Consumers in seven states—Alabama, Alaska, Hawaii, Massachusetts, Montana, Ohio, and Pennsylvania—have the broadest UDAP protection and the greatest incentive for UDAP action. Unit values are displayed in all four columns. In contrast, all UDAP indicator variables equal zero in Arkansas, Iowa, Nevada, and North Dakota. UDAP options for consumers in these states are quite limited.

It is important to note, however, that state-specific conditions change frequently, as evidenced by the National Consumer Law Center publication *Un-*

[61] Sheldon (1982a), p. 83.

[62] *Ibid.*

fair and Deceptive Acts and Practices and its now two annual supplements. Prior to 1981, for example, Alabama was the only state without any UDAP statute. It now appears in the select list of seven jurisdictions promising its consumers the widest possible UDAP consumer protection and procedural incentives.

STATE ENFORCEMENT. State adaptations of the UCC, state UDAP laws, and MM authorize state agencies, usually the state attorney general's office, to investigate and enforce violations. The preceding sections describe the legislative provisions of each statute. This section focuses on enforcement.

There are, to begin with, many dimensions to effective enforcement. The dollar and manpower resources available to the attorney general's office to handle consumer complaints is one. The extent of complaint activity is another. The frequency of successfully mediated or litigated complaints is yet a third.

The consumer protection office in each state was contacted. In most instances the office is a branch of the attorney general's office. Information regarding the number of successfully mediated or litigated complaints typically was unavailable. However, wherever available, states willingly supplied information regarding the size of the office's consumer protection budget, the number of professional staff allocated to consumer protection, the number of overall consumer complaints, and, given the particular interest of this study, the number of complaints related to used car transactions.[63] Information in all four areas was supplied by forty jurisdictions. Eleven states were unable to provide the requested budget data, most often because consumer protection activity was not a separate item in the attorney general's budget. Fiscal 1983 data were requested. Most states complied. Seven forwarded fiscal 1982 or 1984 data.[64]

Responses in each subject area were converted to a per (adult) capita basis and then ranked in ascending order. Within each of the four response categories, states providing information were divided into above and below median groups. Any state occurring in the above median group in at least three response areas was considered a strong enforcement jurisdiction.[65] Twenty-one states and the District of Columbia met this criterion.[66] For these jurisdictions, a unit value

[63] Nearly all states define complaints only as those submitted as formal, written statements. Only Alabama and Delaware do not distinguish between phone and written complaints.

[64] The state agencies were extremely cooperative. Generally, information was provided in writing or over the course of two or three telephone conversations. The average lag between initial contact and a full response was about one week. The notable exception was Wisconsin. Our research assistant was provided answers to his questions by a staff person who simply coded his questions into a terminal located at his desk.

[65] A mandatory auto inspection system intended to establish a standard of minimum quality rather than merely conformity with minimum safety requirements is an additional dimension defining enforcement. However, Wisconsin is the only state with such a role for auto inspections and even without considering its inspection statute, already is among the strong enforcement states as defined here. No separate indicator variable for Wisconsin's inspection statute is included in the model. Whatever effect this inspection system has is captured in the state's enforcement variable.

[66] The eleven states providing information in only three of the four areas typically had below me-
(footnote continued on next page)

was assigned to SE ("strong enforcement"), the final indicator variable in the Appendix Table. The variable takes a zero value for all other states.[67]

The SE variable is admittedly imprecise. Nonetheless, given its discrete form, the variable is probably a reasonably accurate qualitative indicator of the relative intensity across states of enforcement activity in consumer protection areas in general and in used car transactions in particular.

SUMMARY. The three sets of laws discussed above do not displace one another. Indeed, they augment each other, providing consumers with cumulative rights and remedies. Their legislative language assures this.[68]

Cumulation of rights and remedies is important for three reasons. At a minimum, it first means buyers are generally protected by two or more statutory schemes. Therefore, even if one law does not provide a mechanism to bring suit, another may. For example, most MM provisions apply only when a written warranty is given. It does not cover oral warranties or express warranties based on samples or models. All these, however, are covered by the UCC. Similarly, a state UDAP statute may apply when a UCC action is not available because, say, the UCC statute of limitations has expired or product rejection was accomplished improperly. Second, cumulation implies that when there is an apparent conflict between any two statutes, the law with more specific provisions will apply. For example, though a state's adaptation of the UCC may not specifically disallow important warranty disclaimers, MM prohibits any such disclaimer for the term of a limited written warranty. Third, cumulation of remedies means that abused consumers have a variety of remedy options. For example, the UCC permits cancellation of the sale and recovery of actual damages while MM and UDAP statutes provide for attorney fees and UDAPs often further offer statutory minimum or treble damage awards. Consumers, therefore, can pursue remedies under all three statutes, if applicable.

The important implication for this study is that a state's warranty protection statutes are not to be viewed as a set of discrete options such that a consumer, selecting one, forecloses proceeding under another. A state's warranty protection statutes are best considered, as in this study, as an interdependent whole, possessing a constellation of attributes that together define the stringency of the state's warranty-related laws.

---

dian per capita measures in those three categories. In short, these states' consumer protection budget information, had it been supplied, would not have changed the states' enforcement rating.

[67] It might be argued that our survey data should be used directly to construct a continuous measure of enforcement across states. Our procedure to convert it to a dummy variable effectively throws away information. This is a deliberate strategy on our part to emphasize the crudity of our information relative to the attribute we wish to measure. First, enforcement is multi-faceted. It is not obvious which dimensions to choose to model enforcement efficiently. Second, while most would agree that each state's per capita consumer protection expenditure is perhaps the best single candidate for a continuous enforcement variable, the necessary information, as explained in the text, is unavailable for eleven states.

[68] *See* Rigg and Alpert (1982), p. 14.

Viewed this way, the following states appear to offer buyers the least legislative protection: Arkansas, Delaware, Iowa, Missouri, Nevada, North Dakota, South Dakota, and Wyoming. Each has no more than two indicator variables with unit values. Nevada stands out. It has only the most basic UCC and UDAP statutes. In contrast, fifteen states—California, Illinois, Kansas, Maryland, Massachusetts, New Hampshire, New Jersey, New York, North Carolina, Ohio, Oregon, Utah, Washington, West Virginia, and Wisconsin—promise strong legislative protection. Each has unit values for at least five of its eight indicator variables, with no fewer than two in each UDAP and UCC subset. Massachusetts, Ohio, and West Virginia are notable. Each has unit values in six of eight columns. Once the important role of enforcement is recognized, however, the list of fifteen collapses to six—California, Maryland, Massachusetts, New Hampshire, Washington, and Wisconsin.

It is important to conclude this appendix with a disclaimer of its own. The discussion presented in this appendix is not an exhaustive description of consumer protection legislation across states. Its scope is limited in two important ways. It focuses only on the consumer protection statutes' warranty provisions and, moreover, only on those that are relevant to merchant-buyer transactions in used cars. Nonetheless, the appendix description and especially its summary table provide a concise overview of the variability across states in the commitment to used car warranty protection both through legislation and enforcement.

Appendix Table

ATTRIBUTES OF WARRANTY PROTECTION LAWS RELEVANT TO USED CAR TRANSACTIONS, BY STATE

	UNIFORM COMMERCIAL CODE				UNFAIR AND DECEPTIVE ACTS AND PRACTICES STATUTES				
	BASIC PROVISIONS (UCC)	ORDINARY PURPOSE (UCCOP)	NEGOTIATED DISCLAIMER (NEGD)	NO DISCLAIMER (NOD)	UNFAIR PRACTICES (UUP)	PRIVATE ACTION (PA)	TREBLE DAMAGES (TD)	MINIMUM DAMAGES (MD)	STRONG ENFORCE-MENT (SE)
ALABAMA	0	0	0	0	1	1	1	1	0
ALASKA	1	0	0	0	1	1	1	1	1
ARIZONA	1	0	1	0	0	1	0	0	0
ARKANSAS	1	1	0	0	0	0	0	0	0
CALIFORNIA	1	0	1	0	1	1	0	1	1
COLORADO	1	1	1	0	0	1	0	0	0
CONNECTICUT	1	0	1	0	1	1	0	0	1
DELAWARE	1	0	0	0	0	1	0	0	1
FLORIDA	1	0	1	0	1	1	0	0	0
GEORGIA	0	0	0	0	1	1	1	0	1
HAWAII	1	0	0	0	1	1	1	1	1
IDAHO	1	0	0	0	1	1	0	1	1
ILLINOIS	1	1	1	0	1	1	0	0	0
INDIANA	1	1	1	0	0	1	0	0	0
IOWA	1	0	1	0	0	0	0	0	1
KANSAS	1	0	0	1	1	1	0	1	0
KENTUCKTY	1	0	0	0	1	1	0	0	0
LOUISIANA	0	0	0	0	1	1	1	0	0

MAINE	1	0	0	0	1	1	0	0	1
MARYLAND	1	1	0	1	1	1	0	0	1
MASSACHUSETTS	1	0	0	1	1	1	1	1	1
MICHIGAN	1	0	0	0	1	1	0	1	0
MINNESOTA	1	1	0	0	0	1	0	0	1
MISSISSIPPI	1	0	0	1	0	1	0	0	0
MISSOURI	1	0	0	0	0	1	0	0	0
MONTANA	1	0	0	0	1	1	1	1	0
NEBRASKA	1	0	0	0	1	1	0	0	0
NEVADA	1	0	0	0	0	0	0	0	1
NEW HAMPSHIRE	1	1	1	0	1	1	0	1	1
NEW JERSEY	1	0	1	0	1	1	1	0	0
NEW MEXICO	1	0	0	0	1	1	0	0	0
NEW YORK	1	0	0	1	0	1	1	1	0
NORTH CAROLINA	1	1	0	0	1	1	1	0	0
NORTH DAKOTA	1	0	1	0	0	0	0	0	1
OHIO	1	0	1	0	1	1	1	1	0
OKLAHOMA	1	1	0	0	0	1	0	1	0
OREGON	1	0	1	0	1	1	0	1	0
PENNSYLVANIA	1	0	0	0	1	1	1	1	0
RHODE ISLAND	1	0	0	0	1	1	0	1	1
SOUTH CAROLINA	1	0	0	0	1	1	1	0	1
SOUTH DAKOTA	1	0	0	0	0	1	0	0	1
TENNESSEE	1	0	0	0	1	1	1	0	0
TEXAS	0	0	0	0	1	1	1	0	0
UTAH	1	0	1	0	1	1	0	1	0
VERMONT	1	0	0	0	1	1	1	0	1
VIRGINIA	1	0	0	0	0	1	0	1	0
WASHINGTON	1	0	1	0	1	1	1	0	1
WEST VIRGINIA	1	1	0	1	1	1	0	1	0
WISCONSIN	1	0	1	0	1	1	1	0	1
WYOMING	1	0	0	0	0	1	0	0	0
WASHINGTON, D.C.	1	0	0	1	0	1	1	0	1

References

Rigg, Margaret S. and Richard H. Alpert (1982). *Sales of Goods and Services*. Boston: National Consumer Law Center.

____________________ (1983). *Sales of Goods and Services, 1983 Supplement*. Boston: National Consumer Law Center.

Sheldon, Jonathan (1982a). *Unfair and Deceptive Acts and Practices*. Boston: National Consumer Law Center.

____________________ (1982b). *Unfair and Deceptive Acts and Practices, 1982 Supplement*. Boston: National Consumer Law Center.

____________________ (1983). *Unfair and Deceptive Acts and Practices, 1983 Cumulative Supplement*. Boston: National Consumer Law Center.

Comments

Robert W. Crandall
Brookings Institute

John Weicher has provided us with a very nice empirical paper on the operation of the housing market, concluding that the market works to discount the value of houses with high prospective repair costs. In addition, he finds that buyers appear unwilling to pay for new-home warranties, presumably because they do not need them to protect against a risk that they or others can identify with some precision before buying the houses.

The housing market provides a number of interesting "consumer protection" issues for an economist:

i. Is there asymmetrical information in the market between buyers and sellers, and if so, is full seller liability the appropriate policy choice to assure efficient operation of the market?
ii. Does seller liability for defects in a long-lived asset, such as a house, require third-party insurance of the seller's obligation?
iii. Do the potential moral hazard problems in the operation of a market with full seller liability negate the benefits of warranties or other mechanisms for enforcing seller liability for defects?
iv. Do seller warranties present further problems of adverse selection, with the higher risk houses being insured while houses with a lower probability of defects go uninsured?
v. Can full seller liability allow the operation of a market with a sufficient range of price-quality trade-offs to yield a reasonably efficient allocation of resources?

Obviously, the possible questions are endless, but we do not have even the barest hints of empirical answers to most to them. Weicher has now provided us with a provocative answer to (i), and his results may provide some insights into some of the other issues.

Defects and Repair Costs. It is Weicher's conclusion that once one acknowledges the essentially stochastic nature of product-quality risks in houses, the market appears to discount future excess repair costs very well. His analysis of a national sample of 170 houses suggests that the discount in the price of a new home occasioned by anticipated excess future repair costs exceeds the expected value of these repair costs by a substantial amount — an amount he believes consistent with the transactions costs and frustrations required in seeking repairs. This result is almost too good to be true, particularly given the difficulty in standardizing new-home prices for location — both across states and within individual metropolitan areas. One can only ask that Weicher's test be repeated on additional samples of home sales to see if this sample has not uncovered an unusually large number of shrewd or prescient home buyers.

Warranties. Weicher cannot identify a positive value for new-home warranties, but he admits that there are problems with this result. If warranties serve as a reassurance of product quality for smaller, less experienced builders who lack a reputation with prospective home buyers, the warranties will only restore the value of these builders' homes to equality with those homes offered by builders with a more long lasting reputation for reliable homes. Thus, warranties may be offered by a group of builders with a limited record while the established builders choose to self insure. In fact, warranties may be required to assure that the builder has not disappeared or gone bankrupt before defects begin appearing in the house. Home buyers may not require much assurances from established builders.

In addition, there is clearly a problem of adverse selection in home warranties just as in any insurance contract. In the early years of the HOW Corporation, this adverse selection was surely present in the state of Colorado in areas of expansive soils where enrollment rates were among the highest in the country. Under this situation, even the most insensitive buyer might discount the value of a house for potential foundation problems despite the presence of a warranty. Pursuing a warranty claim is not completely costless nor riskless to the prospective buyer.

In summary, I think that Weicher has given us a very useful set of results on the operation of the housing market. His analysis of new-home warranties, however, would be improved if he could develop a simultaneous equations model of both the demand for and supply of home warranties.

Comments

*James Langenfeld**
General Motors Corporation

Both of these papers are extremely useful first steps in understanding how warranty regulations affect the housing and used car markets. Although I focus my comments on Anderson and Gollop's analysis, the strengths and weaknesses of this paper exemplify those of Weicher's and many of the papers presented in this volume.

Anderson and Gollop begin their analysis with a simple profit maximization model of used car dealers. From this economic model, the authors derive econometric tests to estimate the impact of state warranty laws on the price of used cars. Their estimator is a variant of the hedonic technique, which is a standard tool of economic research in the automobile industry. Unlike most hedonic regressions, however, Anderson and Gollop's approach does not use individual automobile characteristics such as horsepower or weight,[1] but instead tries to measure the impact of dealer costs and gross demand characteristics on price. This particular approach has a great deal of potential, and similar approaches have been used successfully in other research.[2]

Despite my favorable impression of the paper, I have some reservations. First, assuming the authors are correct that enforcement of consumer protection legislation adds about $200 (or 3.5 percent) to the price of a used car, interpreting this result is difficult and offers limited guidance for policy makers. Second, the estimation of $200 figure has some technical problems, so the figure should be viewed as an approximate upper limit.

To illustrate the interpretation problem, consider the central hypothesis of the paper: Have warranty laws affected the price of used cars? The answer, subject to some caveats discussed below, is yes for used automobiles with average frequency of repair records and no for used cars with good or bad frequency of repair records.[3] The authors hypothesize that warranty laws may raise the dealer cost of selling used cars, thus forcing up the price of used cars. However this explanation tells us nothing about the value consumers place on the warranties. For example, consider for a moment that some states pass a

*Formerly an Economic Advisor to Commissioner, Federal Trade Commission.

[1] Dummy variables that represent particular automobile models are used to capture the net impact of these physical characteristics.

[2] For a similar approach that measures the impact of environmental and safety regulations on new cars, see Langenfeld (1983).

[3] The only significant results regarding warranty proxies are in the average frequency of repair groups. This limits the generality of the results and raises questions about biased estimates, especially considering that every car in the low frequency of repair group was made in Japan and every car in the high frequency of repair group was made in the U.S.

law requiring "Hell on Wheels" be painted in large block letters on the side of all used cars. This would increase the cost of dealers preparing cars for sale, and therefore increase price. However it is doubtful that many consumers would feel they benefited from such a law, so the price increase would reflect only an increase in cost to dealers.

Now consider the opposite extreme. Assume some states pass laws mandating that new tires be put on all used cars and that the cost of these tires be completely underwritten by the state. In this case, used cars would become more desirable relative to new cars, and we would expect the price of used cars to be bid up in response to increased consumer demand.[4] There would be clear benefits to used car buyers from such a law (although tax payers in general may have reason to object). Dealers would experience no increase in costs, except scarcity rents. Any price increase, therefore, would exclusively reflect the benefits of the law.

With warranty laws, we expect both of these effects to be present. However, it is impossible to determine whether the price increase estimated in Anderson and Gollop's work reflects increased costs of warranties to dealers and no associated benefits (the "Hell on Wheels" case), increased consumer demand for cars with a minimum cost to dealers (the "free wheels" case), or some combination of both. If we believe the authors' simple theoretical model,[5] then they measure increased dealer cost. But even under these circumstances, Anderson and Gollop do not answer the larger question of whether warranty rules are desirable. This can lead to some strange policy recommendations.

Consider the impact on price of Anderson and Gollop's variable for the enforcement activity of state governments. The authors find that the price of used cars increases when consumer protection agencies spend more than average on used car complaint activity. Given this result, the best way to increase consumer welfare would be to cut consumer protection enforcement budgets. After all, absent any evidence of benefits from warranties, cutting the budget would at least benefit consumers by the amount that used car prices drop.

Moreover, the only highly significant dummy variable[6] of the eight used by Anderson and Gollop to measure state Uniform Commerical Code provisions or state uniform deceptive practices acts provisions is zero in only four of the 50 states and the District of Columbia. These states are Alabama, Georgia, Louisiana and Texas, so the dummy may be measuring the impact of warranty provisions or, alternatively, may be picking up some other common regional attribute of these states. This problem is highlighted by Anderson and Gollop's use of the total sale price rather than net price. The real price of cars to consumers is usually affected by the trade-in price of cars, and this may bias Ander-

[4] The bidding up of prices is likely in the used car market where the fixed number of each make, model, and year of car ensures that at least a portion of the supply curve is vertical.

[5] The authors assume constant marginal cost in a competitive market for used cars. But see Footnote 4 on the marginal cost assumption.

[6] That is, statistically different from zero at the 95 percent (or greater) confidence level.

son and Gollop's results. In particular, dealer trade-in policies in the four southern states may systematically differ from policies in other states, or trade-ins may simply be worth less in these states due to (say) a greater preference for new cars relative to the rest of the country. Moreover, there may be a problem of adverse selection, *i.e.*, poor quality cars may migrate to the four states that have more lax warranty laws.[7] If these or any other relevant phenomena are unique to the four southern states, then we would expect the authors to find lower prices, without warranties having any real affect on the price of used cars.[8]

Some of these problems can be solved if data on the quantity sold of each model used car can be obtained, so a simultaneous equations model of supply and demand could be used to separate the cost of warranties (*i.e.*, the supply effect) from the value consumers place on these laws (*i.e.*, the demand effect). Absent these data, supplemental research is necessary to measure the impact of regulations on cost (rather than price) or on benefits of these regulations.[9] The problem of imperfect measures of warranty laws will be more difficult to correct. More institutional knowledge of the four southern states, however, will help to determine how serious the problem is.

In sum, the paper makes significant progress in answering the question "Do warranty laws have any effect?" However, it leaves completely unanswered the question whether warranty laws are desirable. This is indicative of the relatively undeveloped state of most empirical research in consumer protection, and highlights a key question that needs to be answered in future research.

References

Judge, George G.; Griffiths, William E.; Hill, R. Carter; and Lee, Tsoung-Chao, *The Theory and Practice of Econometrics*, New York: John Wiley and Sons, 1980.

Langenfeld, James, *Federal Automobile Regulations*, Ph.D. dissertation, Washington University in St. Louis, 1983.

[7] One way to check whether this is a significant problem is to obtain data on interstate used car wholesaling. If the flow of used cars is not unusually high between stricter warranty states and more lax warranty states, then the problem is unlikely to be significant.

[8] To ascertain whether not using net price has biased the estimates, additional information about used car prices is needed.

[9] In fact, the authors suggest that seemingly unrelated regression techniques may be superior to the simpler pooling procedure they use. There are tests that indicate whether pooling is appropriate, and the authors should employ those tests. See Judge et al (1980), pp. 323-373.

Advertising and Consumer Learning

Y. Kotowitz and *F. Mathewson*
University of Toronto

Abstract

This paper analyzes the impact of advertising on consumer demand. Extant theories postulate that higher advertising expenditures signal either best buys or more experience quality embodied in the product or, alternatively, that persuasive advertising can be a substitute for experience quality. Based on the equilibrium relationships in these theories between price premia, advertising and marginal costs, regressions are run for two industries, automobiles and life insurance. The results offer little support for the signal proponents but are consistent with the substitutability notion.

These results suggest that, counter to the pure signaling theory, advertising content matters in understanding the economics of advertising. We proceed to analyze content in a stylized model of consumer learning with two brands. What limits the claims of firms for their products? This is the question we address; we identify several important margins. Ancillary empirical evidence supports our contention that advertising may mislead consumers, at least in the short run, and that the pure signaling theory alone does not explain adequately the empirical evidence.

1. *Introduction*

An understanding of the effect of advertising on consumer demand is essential for any analysis of the economics of advertising. Some economists proceed by permitting advertising to alter consumer tastes (*e.g.*, Dixit and Norman, 1978); most analysts assume that advertising affects demand either directly by supplying information about product characteristics or indirectly by signaling quality through the committed (sunk) nature of brand advertising. When direct information is provided, the consumer learning process (frequently one period) disciplines firms to reveal only truthful information. When advertising expenditures are brand specific and sunk, then a quasi-rent in the form of a price premium (margin of price over marginal cost) for honesty is a sufficiently large incentive to deliver the perceived quality: dishonesty never pays. Furthermore, the sunk advertising expenditures are sufficiently large that the price premium in equilibrium permits firms to earn normal rates of return. Advertising expenditures are an accurate signal of the embodied (experience or credence) quality. Chiseling by advertisers is unprofitable. Rational expectations characterize the equilibrium.

Casual observations such as court convictions for false advertising suggest that, at least in the short run, chiselers do exist: advertising need not be a signal of quality; some firms do mislead and exaggerate product claims. One difficulty with the message that sunk advertising expenditures alone signal a price premium for honesty is that it offers no explanation for advertising content: any sunk expenditure that consumers (correctly) interpret as a quality signal could potentially yield an equilibrium. If advertising content as opposed to expenditure is unimportant, why do firms frequently resist so vigorously FTC content rules?

Is the advertising commitment theory consistent with the empirical facts? If not, is it possible to reconcile rational expectations with observations of misleading advertising? How do consumers incorporate possibly suspect information into their purchase decisions? What are the market forces that discipline advertising content? Do the answers to these questions lead to testable implications? These are the issues that motivate this paper.

In the next section of this paper, Section 2, we test the competing explanations of advertising for two industries characterized by experience (or even possibly credence) quality. These competing hypotheses for experience quality may be succinctly stated: (i) greater advertising expenditures as a signal indicate "best buys"—those products with a lower price premium (Nelson 1970, 1974); (ii) greater advertising expenditures as a signal indicate higher experience quality whose delivery is guaranteed in equilibrium through a larger price premium (Klein and Leffler 1981; Shapiro 1983; Kihlstrom and Riordan 1984); (iii) greater advertising expenditures can be a substitute for lower product quality conditional on the rate of learning by consumers (Kotowitz and Mathewson 1979; Shapiro 1982). Our tests indicate the need for a further explanation of advertising content with explicit consumer learning. This we tackle or at least begin to tackle in Sections 3 and 4. Section 5 examines ancillary evidence to see if our results find additional empirical support. Finally, Section 6 contains our summary and conclusions.

2. *Empirical Evidence on the Extant Theories*

For empirical purposes, the extant theories present a logical grouping. Nelson (1974) argued that those firms whose products have a smaller price/marginal cost differential—best buys—will advertise the most.[1] Klein and Leffler (1981) argue that those firms whose products have larger embodied experience quality will command a larger price premium and advertise more, so that, in contrast, larger price premia are associated with larger advertising

[1] Nelson defines P^* as the price per unit of utility of the brand. He argues that "...firms vary in their efficiency in producing the utility that consumers seek. Some firms produce brands that yield more utility to the consumer for a dollar of production cost than do other brands. In general, a firm that has lower costs relative to the utility of its brand than other firms will find that it pays to expand its output by both increasing advertising expenditures and decreasing P^*. This behavior of firms by efficiency generates a negative association between advertising and P^* by brands." (Nelson 1974: 732).

expenditures—advertising signals quality. This would also hold if advertising were a substitute for experience quality—persuasive advertising.

If q represents experience quality, c(q) represents the per unit costs of quality (with $c' > 0$), P represents the equilibrium price and A represents the level of advertising exepnditures, then the first critical relationship for these theories may be written as:

$$P - c(q) = f(A) \tag{1}$$

where $f' < 0$ in the "best-buys" argument,[2]
while $f' > 0$ in the "signal" and "persuasive" arguments.

There are two additional tests to differentiate between the "signal" and "persuasive" theories of advertising in experience quality markets. According to the "signal" proponents, in equilibrium, higher levels of experience quality are associated with higher levels of advertising; the "persuasive" proponents argue the converse. In symbols, this may be written as:

$$c(q) = g(A) \tag{2}$$

where $g' > 0$ in the "signal" theory and $g' < 0$ in the "persuasive" theory.

As well, the "signal" proponents argue that the price premia are larger when the experience quality is larger; for "persuasive" proponents, the converse holds. In symbols, this may be written as

$$P - c(q) = h(q) \tag{3}$$

or in terms of observable per unit costs

$$P - c = h(q^{-1}(c)) \tag{3'}$$

where $h' > 0$ for the "signal" theory and $h' < 0$ for the "persuasive" theory.

Two comments are in order at this point. First, by substituting (2) into (3) and considering the sign restrictions, both the "signal" and "persuasive" theories would predict that $f' = h' \cdot [\omega^{-1})' \cdot \gamma' > 0$ (as $(q^{-1})' > 0$). Therefore, (1) cannot discriminate between these hypotheses.

Second, these structural relationships hold in equilibrium; tests however do

[2] If U(q) defines consumer utility as a function of product quality, then $P^* = P/U(q)$; Nelson's hypothesis (cited in fn. 1) directly implies that $(P-c(q))/U(q) = f(A)$ where $f' < 0$ or $ln(P-c(q)) - ln\, U(q) = ln\, f(A)$. We assume that $ln\, U(q)$ is linear in product quality characteristics; this expression appears subsequently in our test of the "best buys" hypothesis (equations (5) and (6)).

not constitute anything beyond correlations. (There are no exogenous variables driving variation in these relationships.) Nevertheless, available data for a restricted sample should offer some discriminating power to assess these theories.

We propose to test the critical signs for these competing theories over two industries where products have experience quality—automobiles and whole-life insurance. First, we consider only (1) in an attempt to differentiate between the "best-buys" hypothesis on the one hand and the "signal" or "persuasive" hypotheses on the other; we begin with our automobile data set.

The passenger car industry supplies a good test case during the 1960's prior to the changes that occurred because of changing energy prices. A very extensive study by Ohta and Griliches (1974) estimated hedonic price equations for 13 U.S. car makes[3] and many models for selected years between 1955-71. The basic form was:

$$\log P = \alpha + \sum_i \beta_i X_i + \sum_j \gamma_j D_j \tag{4}$$

where P is the list price of a given model; X_i are product characteristics and D_j are brand dummies for each of 12 brands.

The estimated $\hat{P}$ thus measures the hedonic price of any model or brand due to its measured attributes and make. The brand coefficients γ_js may measure systematic quality aspects which are present for the brand in question relative to the base brand, or a higher cost per unit of quality for the brand because of higher costs or higher markup.

To the extent that γ_j measures higher cost per unit quality rather than omitted quality attributes, Nelson's hypothesis requires that γ_j be negatively correlated with its corresponding brand advertising (A_j), *i.e.*, higher advertising implies a lower brand price per unit of quality relative to other brands. To the extent that γ_j measures unaccounted for systematic quality elements, however, the negative relation between γ_j and advertising is weakened and may even be reversed, if the brand advertising is highly positively correlated with the omitted brand quality attributes and, in turn, if these are uncorrelated with the X_is.

There is an indirect test for this bias. If γ_js primarily measure omitted quality attributes—deviations of a brand price from its estimate price by attributes—(γ_j + errors),[4] they should be positively related to market share in the relevant class. A study by Cowling and Raynor (1970), however, suggests that brand deviations from the hedonic price are negatively correlated with market share.

[3] The brands of automobiles analyzed by Ohta and Griliches (1974) are: AMC, Buick, Cadillac, Chevrolet, Chrysler, Dodge, Ford, Imperial, Lincoln, Mercury, Oldsmobile, Plymouth and Pontiac. Chrysler and Imperial (2 models manufactured by Chrysler Motors) were dropped because disaggregated advertising data for their brands are unavailable.

[4] As the estimated R^2s are very high (around .98), the errors are small relative to the γ_js.

Thus, a negative γ_j indicates "good value per unit quality" within a product class rather than unobserved superior quality. Another possible bias may arise if γ_js measure unobserved quality attributes between classes. In order to account for this possibility, two dummy variables were added to allow γ_j to vary among standard, medium quality and luxury cars.

The Nelson hypothesis, in particular, may then be tested with the following regression:

$$\gamma = \delta + \epsilon A + \eta S + \sum_i \xi_i D_i \tag{5}$$

where γ is the vector of estimated γ_j's through time, A is the real advertising expenditure by brand and year, separated by total media expenditures (A) and electronic media expenditures (A_e);[5] S is real dollars of sales for each brand (to correct for size effects) (Nominal values are deflated by the CPI); $D_1 = 1$ if the brand is Buick, Mercury, Oldsmobile or Pontiac (medium quality) or $= 0$ otherwise; $D_2 = 1$ if the brand is Cadillac or Lincoln (luxury) or 0 otherwise.

This regression was estimated using the combined time-series, cross-section values of γ_j estimated by Ohta and Griliches for the years 1962-63, 1964-65, 1966-67, 1968-69, and 1970-71. Advertising expenditures by brand come from *Advertising Age*. The regression contains 11 brands over 5 year pairs. As the constant term, which measures the deviation of the base brand price from its hedonic valuation, may vary between model years, we added time dummies for each year to eliminate possible bias. The results turned out to be relatively insensitive to this addition.

We next consider an estimate of (1) for participating whole-life insurance. The data are participating whole-life policies with a face value in 1983 dollars of \$100,000 (constant in real terms over the sample) with a duration of 20 years, issued to a male age 35, for a sample of 15 of the largest Canadian life insurance firms from 1968 to 1977. The price data were taken from *Stone and Cox Life Insurance Tables*. Actuarial costs are calculated using mortality tables (1957-60 ultimate basic mortality for males) and policy cancellation rates (Moorhead's S Rates) taken from the *Federal Trade Commission Report* (1979). Administrative costs per policy are inferred under the assumption that life insurance firms in the sample earned competitive rates of return on these policies in 1960, a period of nominal interest rate stability. These data permit us to calculate [P - c(q)] which is deflated by the CPI to give a real retention index by firm. (These calculations are described more fully in Mathewson and Winter 1985.)

[5] In the measurement of advertising effects for automobiles, we separate electronic from print media. The assertion is that electronic advertising contains less objective information than print advertising; for this reason, separate regressions were estimated using only expenditures on electronic advertising.

Advertising expenses are measured as sales expenses, the sum of explicit advertising expenditures and salesmen's commissions. These data are taken from the *Report of the Superintendent of Insurance for Canada, Volume III, Annual Statements—Life Insurance Companies and Fraternal Benefit Societies (1968 to 1977)*. In general, the media mix selected by life insurance firms is relatively light on traditional advertising and relatively heavy on direct personal selling. We argue that the relevant "advertising" measurement is the sum of these selling expenses. If "best buys" advertise more heavily, then sales personnel are a form of information to consumers. If advertising is a sunk "signal" of quality, then so are these sales expenditures. Participating whole-life insurance, a complex contingency contract, is only one component of estate planning. If life insurance firms offer estate planning, insurance information and insurance contracts as their service, higher sales commissions signal a more personalized individual service to potential consumers. If advertising is a potential substitute for quality under some conditions, then sales personnel glean information on the characteristics of individual consumers and facilitate price discrimination across consumer classes—a "persuasive" role. (This role is developed more fully in an open-entry equilibrium model of the life insurance industry in Kellner and Mathewson 1983 and Mathewson 1983.) As sales personnel are contract specific and do not have the public-good effect of advertising, sales expenses for each firm are measured per dollar of premium revenue.

In particular, the regression equation for the life insurance sample takes the form

$$\log R = \Phi + \mu \log A/S + \Gamma \log S + \sum_i \xi_i D_i \tag{6}$$

where R, the retention figure, equals P - c(q); A is total real sales expenses; S is total real premium revenue in the first year of the new contracts; D_i is a set of time dummies (to correct for the impact of changing nominal interest rates on regulatorily imposed solvency constraints).

The regression results are reported in Table 1. As the coefficients themselves hold no interest, we present only signs and the t-statistics. (For automobiles, the dependent variables are themselves regression coefficients. In this case t-statistics need to be interpreted with caution; they offer some indication of the strength of the measured relations.) These estimates reveal that with some exposure to a Type-1 error, we may reject the "best-buys" hypothesis.

Can these data discriminate between the "signal" and "persuasive" hypotheses? The answer is yes. The test for automobiles is indirect; the test for life insurance is direct. We begin with automobiles.

The "signal" hypothesis states that higher experience quality is associated with greater advertising commitment as an ex-ante signal of quality. The difference between the depreciation of an individual brand of automobiles in the first year of a model's existence from the group average is treated as a relative measure of the consumer's disappointment with the product after use. If advertising is a signal of experience quality, then the depreciation in the first year should be negatively correlated with the advertising intensity of the model in the previous year. If, however, persuasive advertising leads consumers to believe that the product advertised is superior, but these beliefs are not reinforced by experience in the first year, then the depreciation in this year should be positively correlated with the advertising intensity of the model in the previous year.[6]

To test this hypothesis, we regress the deviations of first-year depreciation rates of each brand from the mean first-year depreciation rate for all brands against advertising in the preceding year for the years 1962, 1965, 1967 and 1971. Class dummies were included to allow for systematic variation in discount or durability. (Time dummies were insignificant and so they were dropped.)

The tests for life insurance are direct. To test (2), we regress logs of estimated real marginal cost for each firm against logs of real total sales expenses per dollar of first-year premium revenues, logs of firm size measured as real first-year premium revenues and time dummies to correct for nominal interest rate changes on regulatorily imposed solvency constraints. According to the "signal" theory, the coefficient on sales expense should be positive; according to the "persuasive" theory, the coefficient on sales expense should be negative.

To test (3′), we regress the logs of insurance firm real retention figures against the logs of estimated real marginal cost for each firm, logs of firm size measured as real first-year premium revenues and time dummies as before. According to the "signal" theory, the coefficient on marginal cost should be positive; according to the "persuasive" theory, the coefficient on marginal cost should be negative. Again, only signs and t-statistics are reported; the results appear in Table 2.

The signs of the point estimates are uniformly consistent with the "persuasive" theory and counter to the "signal" hypothesis. Some effects are highly significant, others less so. In our view, the set of tests for these two industries offers little comfort to the "best buys" or "signal" hypotheses, although the evidence is stronger against the former.

Previous analyses of persuasive advertising where such advertising has the potential to substitute for experience quality use simple linear learning rules for consumers (Mathewson and Kotowitz 1979, Shapiro 1982). Is there a more detailed model of consumer learning consistent with rational expectations and the absence of "long-run" chiseling but consistent with (i) the above empirical results and (ii) ancillary evidence from marketing? We turn our attention to this task in the next section.

[6] Depreciation includes the effects of new car discounts and therefore may vary systematically across brands. For example, luxury cars tend to depreciate less in the first year than other types. Such systematic variation across class depreciation rates are captured by the class dummies (D_1 and D_2). There does not appear to be, however, a significant pattern over time for relative depreciation by make (Ohta and Griliches 1974: Table 13 and 14).

3. *A Model with Consumer Learning*

(1) Experience Only, No Advertising

We first set out a simple model of a consumer learning through experience about an unknown brand. Consider the problem facing a single consumer at each point of time in determining a budget allocation between the consumption of two brands of a specific product (Brand 1 and Brand 2) and a numeraire good Z composed of a bundle of all other commodities, where the consumer has a constant stream of income per period. In order to focus solely on the informational aspects of the problem, we assume that the marginal utility of the numeraire good is fixed and independent of the consumption of the specific product, so that income effects are ignored. As well we assume that the two brands of the product are perfect substitutes, and have the same price per unit, and that the consumer purchases only one unit of this product per period. The qualities of good Z and of Brand 1 are known with certainty and the marginal utility index of Brand 1 is set arbitrarily at 1.

The quality of Brand 2 is however uncertain. The consumer's expected utility in period t is defined by

$$EU_t^i = EU\,(p+\sigma(i)q_t(1\text{-}p)) + Z_t \tag{7}$$

where $\sigma(i)$ measures the intensity of preference for Brand 2 for consumer i, p = 1 if the consumer buys Brand 1 and 0 otherwise and $U' > 0$, $U'' < 0$. Quality (q_t) depends on the product's characteristics, where these characteristics are scaled so that the consumer production technology is linear in characteristics. The quality index is measured to reflect the subjective evaluation of the product by the consumer if all product characteristics were known. In most of the following we assume that q is comprised of only one characteristic in order to simplify the analysis. The results are insensitive to this assumption.

Define a consumer's actual experience with Brand 2 as y. It may vary due to intrinsic variation in the quality of the product (assumed normal with mean zero and known variance) or due to errors of measurement on the part of consumers (again assumed normal with mean zero and unknown variance). This experience quality is then normally distributed with mean defined as Q and unknown variance.

Subjective errors may arise from variability in the conditions of usage or complementaries in use with other commodities and are likely to be large for complex products. Consumers are assumed to be aware of such errors and incorporate their variance in their learning. Define s as the consumer's estimate of var(y). This estimate may not be accurate because it cannot be refuted by experience as the error is not observable. Thus s, even if it is erroneous, is fixed for each individual.

In order to make a decision about the product, the individual must form an opinion about the average quality of the uncertain product. This opinion may be gleaned by word-of-mouth from other consumers, from the news media, or from personal search. We assume that the individual formulates an opinion in the form of a Bayesian prior probability distribution of the mean quality (Q). We assume this prior is normally distributed with mean μ and variance v, and that the consumer uses a Bayesian update rule to modify beliefs (reflected in μ and v) through time, in light of accumulated experience from the consumption of Brand 2. A time (t+1), therefore, the posterior is also normally distributed with the rate of change of the mean and variance given by equations (8) and (9) respectively:

$$\mu_{t+1} = [(s\mu + vy)/(v+s)]_t \tag{8}$$

$$v_{t+1} = [(sv/(v+s)]_t \tag{9}$$

or, in terms of changes in the prior mean and variance (ignoring time subscripts),

$$\triangle\mu = m(y-\mu) \tag{8'}$$

$$\triangle v = -mv \tag{9'}$$

where $m = v/v+s$ is the speed of learning.

Equations (8) and (9) assume that information does not depreciate. To incorporate depreciation of information in the model, we would need to modify equations (9) and (9′) slightly to allow v to change over time. The fact that information becomes dated or is forgotten does not change the consumer's prior conception of the average quality of Brand 2; it does change the consumer's estimate of the reliability of prior beliefs and modifies the estimate of the prior's variance. As the reliability of prior beliefs is likely to decline with the passage of time, the variance of the prior must rise with time. We shall assume this process occurs in a linear fashion so that equation (9') would be modified to include a "coefficient of information depreciation" in the form:

$$\triangle v = (-m + \delta)v \tag{9''}$$

The effect of "forgetting" is to counteract the effect of experience on the prior variance and to decrease the rate at which the prior variance is reduced with experience. Note that, as $t \rightarrow \infty$, v approaches a finite positive limit. This is because $\triangle v$ approaches zero before v approaches zero. As the rate of reduction in uncertainty due to accumulated experience falls to the level of the rate of information depreciation, the level of uncertainty stabilizes at the steady state level $v^* = s/(1-\delta)$. In the steady state for v, equation (8′) degenerates into the standard adaptive expectation equation $\triangle\mu = \delta(y-\mu)$. The "forgetting" coefficient is then a lower bound of the speed of learning when experience is accumulated.

If the consumer attempts to maximize discounted expected utility over a given horizon, an optimal policy consists of choosing a reservation quality at each point of time R_t for the unknown brand, such that if $\mu_t > R_t$, the unknown brand is purchased at time t. That is, if the consumer believes that the mean quality of Brand 2 at time t exceeds this reservation level at time t, Brand 2 is purchased. Otherwise Brand 1 is bought. Note that once the consumer buys Brand 1, the consumer always buys Brand 1 unless some additional information from outside sources leads to a change in the prior in favor of Brand 2 or unless the time horizon changes. If the individual buys Brand 2, the individual revises the estimate about the prior mean and variance in the light of experience with the product, compares the new prior with the reservation value and proceeds as before.

(2) Experience and Advertising

We are now in a position to incorporate persuasive advertising into the model. We first analyze the *content* of advertising and then the advertising expenditure. Advertising expenditures determine the number of consumers and their attentiveness to advertising messages, while content is assumed to affect their perception and their behavior.

Advertising content has three functions—first to capture the consumer's attention, second to insure message retention and finally to transmit information about the product. In the language of communication experts, the first and second functions are to create an "interrupt" (Simon 1957) in order to induce conscious consideration of the information supplied. (This is the role of the pretty girl/boy in the bikini, which may not have anything to do with the information transmitted.) We ignore this component and concentrate here on the information content of advertising.

Consider the message content of advertising. Advertising is assumed to make claims about the quality of the characteristic relevant to the determination of quality. Advertising then acts as another piece of information. Define the advertising claim of quality as c (measured in units of quality). The consumer incorporates this information in the prior according to Bayes rule, weighing its value in inverse relation to the estimate of its reliability defined as r, which may be interpreted as the consumer's estimate of the proportion of the variance of the true mean quality (q) explained by the advertising message; equivalently, r may be interpreted as the consumer's unconditional estimate of the coefficient of variation between c and q . Upon receipt of the advertising message and assuming no passage of time to process the information, the revised prior mean (μ') and variance (v') become (again ignoring time subscripts):

$$\mu' = [(r\mu + vc)/(r + v)] \tag{10}$$

$$v' = [rv/(r + v)] \leq v \tag{11}$$

Once the advertising message is received and incorporated in the consumer's posterior, there is no role for further advertising. There is no reason to believe that a consumer revises the prior again upon receipt of the same message from the same advertiser. Repetition in this context may be necessary to insure that the information is noticed by the consumer. But once noticed, there is no further effect unless the message changes, or the identity of the advertiser is disguised.

There is nothing in the Bayesian learning process to determine the credence of the advertising message. There are two ways to proceed. First, we may assume rational expectations. That is, we may define a value of r which is consistent with equilibrium levels of c assuming that such equilibrium levels exist. Because, in the absence of government regulations, there is no dollar cost to making exaggerated claims, there is no limit on such claims. Firms will thus make unlimited claims which are unrelated to the true quality of their product *unless* there is a market mechanism to discipline exaggerated claims, so that c is correlated with true product quality.

In the absence of government regulations, the consumer model as it stands here does not supply a penalty for false or exaggerated advertising. While the benefits to firms from exaggerated advertising rise with true product quality, they are always positive: if it pays to advertise at all, it pays to exaggerate without limit because the marginal cost of exaggeration is zero.

As a consequence, in the absence of truth in advertising regulations, no correlation between advertising claims and true quality can be established and hence advertising claims contain no useful information. The content of advertising is therefore irrelevant and should not affect rational consumers. Notice that this is also the position of those who cast advertising *exclusively* as a signal of quality.

A finite exaggeration level requires the imposition of additional structure. This defines the second way to proceed. The psychology and marketing literature dealing with cognitive processes suggests that the credibility of information depends on its conformity with both general and particular preconceptions. Conformity of general preconceptions may be related to persuasiveness in the sense that the information supplied does not contradict general knowledge. For example, the washing power of detergents may be reasonably related to the presence of enzymes, automobile horsepower is related to acceleration, etc. Conformity to *particular* preconceptions must be related to the prior. The theory of cognitive dissonance states that people are uncomfortable in the presence of conflict. When information is received which contradicts accepted beliefs, a dissonance is created. To resolve the conflict, consumers may acquire additional information or re-evaluate the source of the conflicting information, possibly discarding some of it. In particular, where new information conflicts with prior decisions, it may be ignored. In terms of our model, the theory suggests that the more likely is the information to conflict with the prior—cognitive dissonance—the less credible it is.

A convenient way of measuring such conformity, in terms of the probability that the information came from the prior, is the distance of the information from the prior mean. We may specify the following relationship for reliability:

$$r = r(|c - \mu|/c), \text{ where } r' > 0 \text{ and } r'' > 0.$$

Returning to equations (10) and (11), we see that the effect of advertising on μ' and v' rises with the consumer's ignorance because when the prior is diffuse, consumers are more receptive to advertising ($\partial r/\partial v < 0$) and give lower weight to their uncertain prejudices. As r rises, advertising is discounted, so that $r \to \infty$ leads to $\mu' \to \mu$ and $v' \to v$: The advertising is ignored and priors are not revised.

This specification of r yields a limit on the ability of advertising to affect consumer perceptions; as advertising claims are exaggerated, consumer confidence in them falls, reducing the weight given to the advertising. This negative indirect effect, represented by the first term of equation (13), may offset the direct effect of increased c, represented by the second term.

$$d\mu'/dc = r'(\mu\text{-}c)/(v+r)^2 + v/(v+r) \tag{12}$$

As $d^2\mu'/dc^2 = r''(\mu\text{-}c)/v < 0$, the marginal effects of advertising claims on consumer perceptions declines as c rises, $d\mu'/dc = 0$ therefore defines the maximum influence which advertising may have on a consumer with initial perceptions of μ and v.

The effects of advertising on experimentation is somewhat complex. Positive advertising ($c > \mu$) tends to raise the prior mean and encourage the relevant consumer to purchase the uncertain brand. However, advertising also reduces v' affecting the reservation quality (R). The effect of such a reduction in v' on R is uncertain and depends on the degree of risk aversion. In particular, if the consumer is risk neutral, $\partial R/\partial v < 0$, so that an increase in c reduces v', inducing in turn an increase in the reservation quality which partly offsets the favorable effect on μ'. If, however, the consumer is sufficiently risk averse, $\partial R/\partial c < 0$, so that increased confidence reduces the subjective risk evaluation encouraging experimentation.

In order to investigate whether such consumer decisions are consistent with rational expectations (*i.e.*, whether they are rational in the context of market equilibrium), we must analyze the firm's decisions.

4. *The Firm's Advertising Decision*

These decisions are extremely complex when consumers are diverse, because they must take into account the reactions of different consumers to the common information supplied. To simplify the problem, we analyze the firm's decisions under two sets of assumptions, first, that consumer tastes are diverse but quality priors are uniform, and second that consumer tastes are identical but quality priors are diverse. The results vary significantly between these two sets of assumptions.

We assume that firm 2 (producing Brand 2) operates as a Nash competitor in determining its advertising content and expenditure vis-a-vis firm 1 (producing Brand 1) and that it accepts the price determined by firm 1 which we treat as exogenously determined. (This separation of advertising and price competition is similar to Schmalensee 1982.) As well we assume that the market can sustain only the two firms in question.

(1) Consumers with Identical Quality Expectations and Varying Tastes

Assume all consumers hold the same initial values of μ_o, v_o and s. But their preferences (σ(i)) differ. Consumer demand for Brand 2 is then determined by the condition $\mu_o \geq R(\sigma)$. Define the advertising reach (N) as the proportion of consumers exposed to advertising; N is assumed to be an increasing function of advertising expenditure (A) so that $N = N(A)$ with $N' > 0$, $N'' < 0$. The proportion of consumers who buy Brand 2 out of those reached (n_t) is determined by the condition $\mu'_t \geq R_t(\sigma(i))$ (*i.e.*, n_t is the proportion of consumers whose reservation quality falls short of their mean quality perception). If consumers are arranged in declining order of preference, we may write $R_t = R_t(n_t)$. These relationships at $t=0$ are illustrated in Figure 1.

If advertising only informs consumers of the firm's existence, $N \cdot n_0$ consumers purchase the product initially. Of these, some continue to buy the product and others do not, depending on their experience, *i.e.*, on the evolution of μ'_t and $R_t(n)$ over time. Define the advertising of a specific quality level as c; this advertising improves consumer evaluation of the product leading to an upward shift of μ_o to $\mu'_o(c)$ in the first period and therefore in all subsequent periods, and may raise or lower $R_t(n)$ to $R'_t(n)$ depending on the degree of consumer risk aversion. This is illustrated at $t=0$ in Figure 1.

The firm's discounted profits are then:

$$E\pi = (1\text{-}C)N(A) \cdot \sum_{0}^{T} \beta(t)En_t - A \equiv (1\text{-}C)N(A)\Phi - A \qquad (13)$$

where $n_t = n(\mu'_t, R'_t)$ and μ'_t and R'_t evolve according to rules established in Section 3; $\beta(t)$ is the discount rate and $C = C(q)$ is the marginal cost assumed to be independent of output but increasing in true quality. True quality is assumed to be exogenous.

The firm's first-order conditions are:

$$\partial E\pi/\partial A = (1\text{-}C)\Phi \cdot \partial N/\partial A = 1 \qquad (C < 1) \qquad (14)$$

$$\partial E\pi/\partial c = (1\text{-}C)N(A) \cdot \partial\Phi/\partial c = 0 \qquad (15)$$

Equation (14) determines the advertising budget or reach for a given level of advertising content (c). There is insufficient structure in this model to investigate whether this advertising expenditure yields a valid signal of the true product quality.

We can however investigate the relationship between advertising content and true product quality. Equation (15) determining the optimal advertising content (c) may be alternatively written as:

$$\partial\Phi/\partial c \cong E\{(\partial n_t / \partial n_0)[(\partial n_0 / \partial \mu_0')(\partial \mu_0' / \partial c) + (\partial n_0 / \partial R_0') \bullet (\partial R_0' / \partial c)]\} = 0 \quad (16)$$

The term $\partial n_t / \partial n_0$ measures the decrease in repeat purchases generated by experience with the product and depends mainly on the differences between perceptions and experience as described by equations (8) and (9). Because all consumers who buy the product have the same initial perceptions, the expected evolution of these expectations is the same for all. The term in the square bracket represents the change in the proportion of customers reached who buy because of a change in their initial perceptions and reservation prices[7] induced by advertising. It is illustrated by the shifts from μ_0 to μ'_0 and R_0 to R'_0 in Figure 1. Equating this term to zero defines the maximum increase in first period purchases obtainable by an increase in c . This level of c is independent of the true product quality. In this context, advertising content cannot supply a rational expectations signal of true quality. The reason is that there is no penalty for false advertising. Whatever the reinforcement effect of true product quality on exaggerated advertising, such effects never offset the profitability of the exaggerated claims. Unless we introduce a punishment related specifically to the difference between the advertised and true quality, advertising content cannot be a rational expectations signal.

In the context of advertising for a single product, there is no incentive for a rational consumer to punish false advertising, as the consumer's buying pattern should be independent of the *source* of any disappointed expectation. Such punishment would be a public good: Punishment for exaggerated claims would establish advertising as a truthful instrument benefitting all consumers yet no single consumer has the appropriate incentive to inflict the punishment.

It is clear that truth in advertising regulations has a role to play under these circumstances. The role of government regulation is clearly to increase the correlation between the advertised message and the true quality. The degree to which this can be achieved depends on the effectiveness of enforcement of truth in advertising regulations.

Consider a modification of our previous model so that there is a penalty for false advertising. Imagine the continuous introduction of new products or attributes by the firm, so that new advertising is required to induce consumers

[7] This is only approximate as the initial change in reservation prices is not the same for all consumers because of possible differences in the concavity of n at different levels of α.

to try the new product. In this setting, false advertising for one product will likely lead to greater skepticism about future advertising. A direct penalty is thus introduced which is specifically related to the false advertising messages (Golden 1977). Similar effects would be generated by growth in the number of potential customers for the product, coupled with word-of-mouth information about false advertising, a modification not pursued by us.

The preceding model leaves no room for advertising by Brand 1, whose quality is known with certainty. To take account of such advertising, it is necessary to allow some uncertainty about some characteristics of the known brand. Such uncertainty can be introduced without compromising the basic analysis of the preceding sections, if we assume that experience with the known brand has been sufficient to reduce the uncertainty of its prior to its steady state value $v^{\infty} = s\delta/(1-\delta)$. Continuous consumption of Brand 1 yields some information but this information has no value because it does not alter the consumer's behavior. Therefore, although no model is developed formally, we believe that the logic of consumer decisions of our model remains unchanged.

(2) Consumers with Varying Quality Expectations and Identical Tastes

An Extension

In this setting, individual consumers still make decisions according to the model described earlier in Section 3, except that in equation (7) $\sigma(i) = \sigma$ for all consumers. Furthermore, the expected profit function for the firm producing Brand 2 remains as equation (13). The decision on reach defined by the magnitude of the advertising expenditures is again defined by equation (14).

The critical effect that changes is the selection of the marginal consumer to define the content of the advertising message. (Recall that in our model firms can neither segment markets nor discriminate in their advertising messages across consumers.) At the outset, we may define two levels of product quality critical for the advertising campaign, an upper level q and a lower level q.

q is sufficiently high that all consumers who buy Brand 2 continue to purchase the brand regardless of their priors. (Again, recall that price is exogenously fixed at 1, the known quality level of Brand 1, the substitute product.) Then for all $q \geq q$, the value of c that maximizes the proportion of informed consumers who try the product will also maximize the present value of future revenues. For $q \geq q$ then, the prior of infra-marginal consumers is not relevant and the advertising content should be addressed to the marginal consumer who is least inclined to try Brand 2 in the first period.

The same decision is relevant for $q \leq q$ but for a different reason. For $q \leq q$, consumers who purchase the product do so for only the first period. In this quality interval, there are no future sales, so that the relevant consumer for advertising content is again the marginal consumer who is least inclined to try Brand 2 in the first period.

This is not the case for $q < q < q$. Here the effects of the choice of c on consumers inframarginal in the first period must be considered, for such effects alter the probability of repurchase and therefore the future flows of revenues. In terms of equation (15), these effects may be written as:

$$dn_0/dc + \sum_{1}^{T} \beta(t)\cdot dn_t/dn_0\cdot dn_0/dc = 0 \qquad (\text{with } C < 1) \qquad (15')$$

The relevant margins are most easily illustrated in a diagram; once these are clear, the relationship between c and q may be defined.

Consider Figure 2 which illustrates quality expectations for consumers and alternative advertising programs. Potential consumers are defined by declining expectations of quality for Brand 2. The decision rule for the consumer's purchase decision is unchanged from the model of Section 3, namely to buy if the relevant expectation exceeds the reservation value R_t.

Consider the first time period and the advertising policy $\hat{c}_0$ which the last consumer in this time period ($\hat{n}_0$) would find just sufficiently credible that this consumer would purchase the product. In other words, $\hat{c}_0$ is tailor-made for consumer $\hat{n}_0$. Consumers who hold higher or lower quality expectations have these expectations increased by advertising but less so than an advertising campaign tailor-made to their expectations. These expectations are represented by the curve in Figure 2. This holds for all feasible advertising campaigns. In particular, consider $\tilde{c}_0$ ($>\hat{c}_0$) which is designed to maximize the increase in the expectations for consumers $\tilde{n}_0$($<\hat{n}_0$). (A repetition of this for all consumers defines the upper envelope denoted as μ' in Figure 2.)

Consider the increase from $\hat{c}_0$ to $\tilde{c}_0$. As illustrated in Figure 2, the consequence of this increase is that consumers in the interval $(0,\pi_0]$ hold higher expectations as the increased campaign is closer to their optimal campaign; consumers in the interval $(0,\bar{n}_0)$ hold lower expectations as the increased campaign is further from their optimal campaign. Consumers in the interval $(\hat{\hat{n}}_0,\hat{n}_0)$ who formerly bought the product in period 1 no longer purchase at all.

In terms of (15′), these effects may be approximated as:

$$[\triangle n_0 / \triangle c]_{\hat{\hat{n}}_0}^{\hat{n}_0} + \sum_{1}^{T} \beta(t)[\triangle n_t / \triangle n_0\cdot\triangle n_0 / \triangle c)_{\bar{n}_0}^{\hat{\hat{n}}_0}$$

$$+ (\triangle n_t / \triangle n_0\cdot\triangle n_0 / \triangle c)_0^{\bar{n}_0}\ ] = 0 \qquad (15'')$$

where $[\triangle n_0 / \triangle c]_{\hat{\hat{n}}_0}^{\hat{n}_0} < 0$, $(\triangle n_t / \triangle n_0\cdot\triangle n_0 / \triangle c)_{\bar{n}}^{\hat{\hat{n}}_0} < 0$,

$(\triangle n_t / \triangle n_0\cdot\triangle n_0 / \triangle c)_0^{\bar{n}_0} > 0$.

In this model, therefore, an interior equilibrium defines a critical "inframarginal" consumer at $t=0$ who determines the nature of the advertising campaign.

The consequence is that where the true quality of the unknown product is sufficiently low or high, the marginal consumer at $t=0$ determines the appropriate level of c; in the middle range, an interior equilibrium focused on the appropriate inframarginal consumer at $t=0$ yields at first an increase in c with q and subsequently a decrease as illustrated in Figure 3. As a result, knowledge of c is not alone sufficient to infer q, *i.e.*, the relationship in Figure 2 is not invertible. Or c is not a valid rational expectations signal of q.

5. *Ancillary Evidence*

The consumer learning possibilities modeled in the preceding sections leave us with the proposition that, at least in the short-run where consumers are learning about the embodied experience quality, advertising and quality can be substitutes. In this case, price premia exist because of increased advertising as a substitute for quality, not as a signal of quality. This explanation is consistent with the empirical results of Section 2 of this paper.

The "signal" theory of advertising may gain some empirical support from studies of the information content of advertising on television (Resnic and Stern 1977) and in magazines (Healy and Kassarjian 1983). While many advertisements contain some information, the claim is that a large portion of television ads appear to contain none. For example, of 378 commercials analyzed by Resnic and Stern (1977), more than one half were found to contain no information. Even when television advertising contains information, such information frequently takes the form of claims that are difficult or impossible to verify, for example claims that products contain attributes or confer benefits on a consumer that depend on the consumer's social situation or state of mind. For example, Shimp (1975) found that 58.5% of representations in a large sample of television advertising consisted of such claims and only 41.5% asserted objective information.

These observations then might support a signal theory of advertising which contends that advertising content is unimportant. Only the sunk resource expenditure is the commitment to deliver. This evidence, however, is not entirely convincing. These studies take only verbal information into account. Television and magazine ads may include significant non-verbal information. The visual image of a car gives information about appearance, comfort and size; a car driven expertly on wet pavement supplies an implicit message of handling and safety features. Consequently, the informational content of television advertising is likely to be underestimated.[8] Further, marketing studies that evaluate alternative advertising campaigns suggest that, as some campaigns are more effective, content is important.

The final piece of ancillary evidence comes from R. Caves (1983). Using data on individual firms over the period 1973 to 1980, Caves was able to test the

[8] Our empirical results on automobiles suggest that electronic media advertising tends to be more "persuasive" and less objectively "informative" than print media.

influence of a set of explanatory variables on 4 jointly determined advertising variables—media advertising expenditures, sales force expenditures, promotional (catalogues, exhibits, coupons, etc.) expenditures and other marketing expenditures. Among the variables in the explanatory set are (i) a variable to capture the percentage of a given firm's output deemed higher quality than its rivals minus the percentage deemed lower quality and (ii) a variable to measure the amount of time elapsed between purchases by the final buyer of the product—infrequent buyers have more obsolete information so that advertised signals of quality should have greater consumer value. Contrary to the "signal" hypothesis, Caves (1983: 18) found that increases in both these variables *decreased* seller expenditures on advertising.

6. *Conclusions*

At the outset, this paper addressed a set of questions on the economics of advertising. For at least two industries whose products are logical candidates for embodied experience quality and therefore advertising as a signal of product quality, automobiles and whole-life insurance, the empirical evidence offers little support for proponents of hypotheses that greater advertising either indicates better buys or signals more experience quality. The results are, however, consistent with the hypothesis that advertising is substitutable for quality under certain conditions on consumer learning. These effects have previously surfaced in consumer linear learning, product life-cycle models of a product supplied by a monopolist (Kotowitz and Mathewson 1979; Shapiro 1982). These empirical results, however, warrant a more detailed examination of the role of advertising content as informational input into consumer learning and the profit-maximizing decisions of firms supplying both goods and information to such consumers.

This model of consumer learning yields a supply of exaggerated claims by firms in their products under certain conditions. Penalties for exaggeration limit the magnitude of puffery by firms; the FTC rules on advertising content play a role here. Firms choose optimal levels of exaggeration; advertising need not be a signal of product quality.

Ancillary evidence tends to corroborate our story. Marketing studies indicate low levels of information in some advertising and high levels of non-verifiable claims. Another study using individual firm data (Caves 1983) found little support for the signal hypothesis.

Table 1

ESTIMATES OF THE CORRELATION BETWEEN PRICE PREMIUM AND ADVERTISING (EQUATION (1))*

(i) *Automobiles*

Number	A	A/S	A_e	A_e/S	S	D_1	D_2	Time Dummies Presents	$\bar{R}^2$
1	.80				1.15	1.40	16.40**	Yes	.89
2		1.08			-.77	1.51	16.03**	Yes	.89
3			1.95**		-1.03	1.81**	16.60**	Yes	.90
4				1.89**	1.69**	1.98**	13.12**	Yes	.90

(ii) *Participating Whole-Life Insurance*

log A/S	log S	Time Dummies Present	$\bar{R}^2$
1.67**	-8.85	Yes	.90

* Entries in the table for each variable are signed t-statistics.
** Significant at the 5% confidence level.

Table 2

ESTIMATES OF THE CORRELATION BETWEEN EXPERIENCE QUALITY AND ADVERTISING (EQUATION (2)) AND BETWEEN PRICE PREMIUM AND EXPERIENCE QUALITY (EQUATION (3))*

(i) *Automobiles* (Equation (2))

Number	A	A/S	A_e	A_e/S	S	D_1	D_2	$\bar{R}^2$	T+
1	1.43				-1.62	-2.02**	-6.53**	.74	
2		1.89**			.31	1.61	-7.22**	.72	
3		2.10**				-2.26**	-6.27**	.73	
4			1.38		-1.60	-1.73**	-5.91**	.70	
5				2.54**	.39	-1.14	-4.04**	.74	
6				2.53**		-1.91**	-6.27**	.74	
7		.93		1.83**	.71	-.91	-4.12**	.74	

(ii) *Participating Whole-Life Insurance*

(a) Equation (2)

log A/S	log S	Time Dummies Present	$\bar{R}^2$
-.22	-.49	Yes	.63

(b) Equation (3)

log c	log S	Time Dummies Present	$\bar{R}^2$
-3.65**	-9.72**	Yes	.91

* Entries in the table for each variable are signed t-statistics.
** Significant at the 5% confidence level.
\+ Time dummies were uniformly insignificant and so were omitted.

Figure 1

CONSUMER PURCHASES AND THEIR RESPONSE TO ADVERTISED MESSAGES

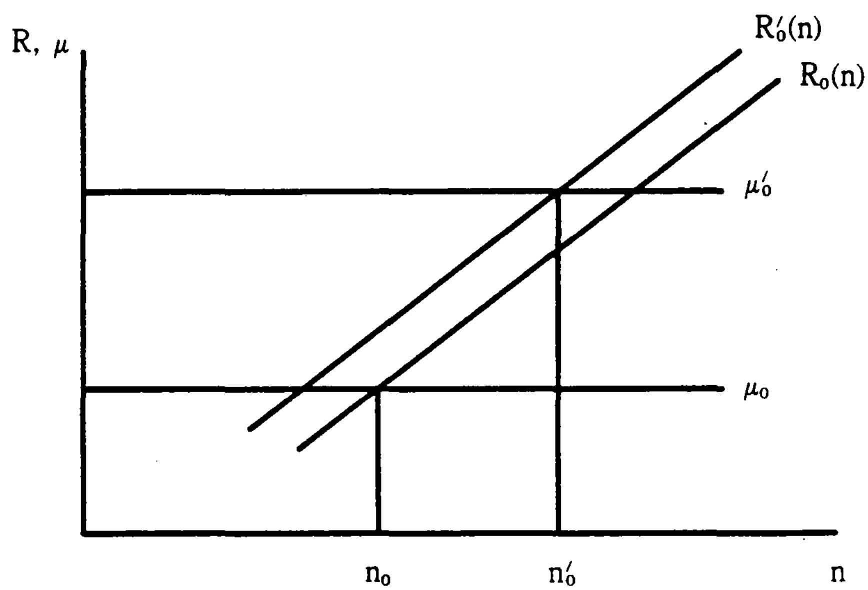


Figure 2

THE DETERMINATION OF THE OPTIMAL ADVERTISING CAMPAIGN FOR CONSUMERS WITH IDENTICAL PREFERENCES BUT DIFFERENT QUALITY EXPECTATIONS ($\underline{q} < q < \bar{q}$)

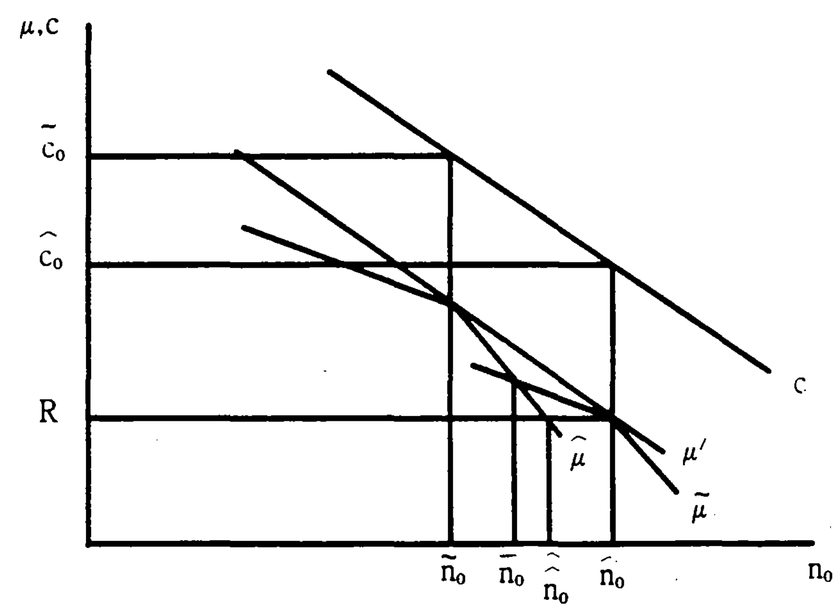
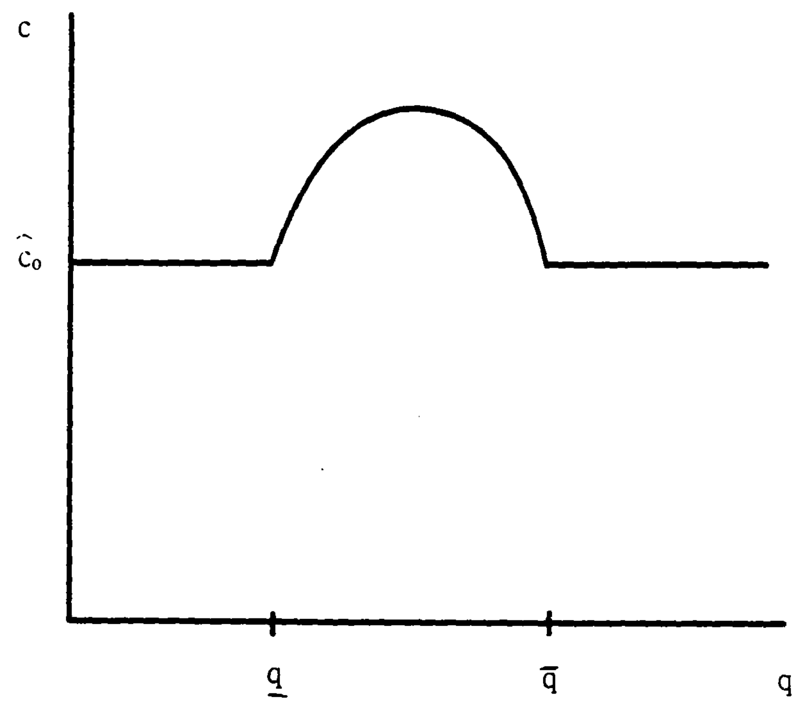


Figure 3

THE RELATIONSHIP BETWEEN TRUE QUALITY AND ADVERTISING CONTENT WITH COMMON CONSUMER TASTES BUT DIVERSE QUALITY EXPECTATIONS



References

Advertising Age, various issues.

R. Caves 1983. "Information Structures of Product Markets," D.P. 1002, Harvard Institute of Economic Research, Harvard University, Cambridge, Massachusetts.

K. Cowling and A.J. Raynor 1970. "Price, Quality and Market Share," *Journal of Political Economy* 78 (Nov.-Dec.) 1292-1308.

Department of Insurance, Government of Canada, *Report of the Superintendent of Insurance*, various years.

A. Dixit and V. Norman 1978. "Advertising and Welfare," Bell *Journal of Economics* 9(1): 1-17.

Federal Trade Commission 1979. *Life Insurance Cost Disclosure: Staff Report to the Federal Trade Commission* by the Bureau of Consumer Protection and the Bureau of Economics, Washington, D.C., U.S. Government Printing Office.

L. Golden 1977. "Attribution Theory Implications for Advertisement Claim Credibility," *Journal of Marketing Research* 14(1): 115-117.

J. Healy and H. Kassarjian 1983. "Advertising Substantiation and Advertiser Response: A Content Analysis of Magazine Advertisement", *Journal of Marketing* 47: 107-117.

S. Kellner and F. Mathewson 1983. "Entry, Size Distribution, Scale and Scope Economies in the Life Insurance Industry," *Journal of Business* 56(1): 25-44.

R. Kihlstrom and M. Riordan 1984. "Advertising as a Signal," *Journal of Political Economy* 92(3): 427-450.

Y. Kotowitz and F. Mathewson 1979. "Advertising, Consumer Information and Product Quality," *Bell Journal of Economics* 10(2): 566-588.

B. Klein and K. Leffler 1981. "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy* 89(4): 415-41.

G. F. Mathewson 1983. "Information, Search and Price Variability of Individual Life Insurance Contracts," *Journal of Industrial Economics* 32(2): 131-148.

G. F. Mathewson and R.A. Winter 1985. "The Economics of Life Insurance Regulation: Valuation Constraints" in J. Finsinger and M. Pauly (eds.) *Insurance Market Regulation*, Macmillan.

P. Nelson 1970. "Information and Consumer Behavior," *Journal of Political Economy* 78(2): 311-329.

P. Nelson 1974. "Advertising as Information," *Journal of Political Economy* 82(4): 729-754.

M. Ohta and Z. Griliches 1976. "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis" in N. Terlecky (ed.) *Household Production and Consumption*, NSER Studies in Income and Wealth, No. 40, Columbia University Press: New York.

A. Resnic and B. Stern 1977. "An Analysis of Information Content in Television Advertising," *Journal of Marketing* 41: 50-3.

R. Schmalensee 1982. "Advertising and Market Structure" in IEA Conference on *New Developments in Market Structure*.

C. Shapiro 1982. "Consumer Information, Product Quality and Seller Reputation," *Bell Journal of Economics* 13(1): 20-35.

C. Shapiro 1983. "Premiums for High Quality Products as Returns to Reputations," *Quarterly Journal of Economics* 98(4): 659-680.

T. Shimp 1975. "Comparison Advertising in National Television Commercials," *Combined Proceedings*, American Marketing Association: 504-8.

H. Simon 1957. *Models of Man*, John Wiley and Sons: New York.

Stone and Cox 1980. *Life Insurance Tables*, Toronto, Stone and Cox Publishing.

The Demand for Advertising by Medium: Implications for the Economies of Scale in Advertising

*Timothy F. Bresnahan**
Stanford University

1. Introduction

This paper seeks to explain the growth in recent times of national brands as an economic phenomenon. Why is it that firms such as Procter and Gamble and Anheuser-Busch have come into existence and persisted as successful forms of economic organization? The entire organizational capital of P&G is marketing capital: the firm exists to create and earn the returns to nationally-branded consumer products, mainly in consumer products like laundry detergent and toothpaste. No serious observer has suggested sufficient economies of scale in manufacturing to support the continuing existence of such a firm. The explanation must lie not so much in the soap as in the soap opera: P&G's advantages over smaller firms arise, in this view, from economies of scale in national distribution of branded products. Similarly, Anheuser-Busch is a ten-brewery seller of branded beer. The market success of such a large firm is not the return to superior manufacturing technique, but rather to a fundamental marketing innovation made 25 years ago: the idea of national, branded, "premium" beer.[1] Thus one possible explanation of the growth of the extremely large consumer-products firm is the exploitation of scale economies in the distribution function of the firm, particularly in advertising. This view is somewhat controversial, as not all scholars agree with the proposition that there are scale economies in advertising.

Such an explanation of the recent growth of large firms with nationally branded products must go on to explain why the distribution-system scale economies only came into existence in recent times or why they were only recently exploited. In this paper, that explanation will arise from a view of the distribution system for consumer products and of the long-run economic forces affecting that system. The core observation in this view is that high-service retailing and advertising are substitutes (in the long run) in the production of distributive services. One way to inform consumers about the nature of products and their uses is at point of sale: an informed, helpful sales person could perform this function. Another way is to have the product simply sit on a shelf, and have all communication about the product's attributes flow directly from manufacturer to consumer

*I would like to thank Tim Taylor for exceptional research assistance, and Gerard Butters, James Ferguson, Pauline Ippolito, Richard Kihlstrom, and Ray Olszewski for helpful comments. All opinions and errors are my own, not those of the FTC or any other agency or person.

[1] Similarly, the size of the second-largest brewer, Miller, is an indication of the introduction 10 years ago of modern advertising and distributing methods, previously developed in the cigarette industry, into brewing.

through advertising media. (The word "inform" can be replaced with "convince," "sell," or "mislead" without altering this argument in the slightest.) In this highly stylized view, there are two different forms of distribution system—traditional, in which the retailer has a great deal of informing or convincing value-added, and convenience, in which mass media perform that function and the retailer serves only to provide low-cost access to the product for the consumer. The two different forms of distribution system are substitutes; within the convenience form, mass media and low-service, low-cost retail outlets are complements.

Since the traditional distribution method is labor-intensive, it has grown substantially more expensive in our century. Since the convenience method is communications-intensive, it has grown substantially cheaper. Specifically, extremely cost-effective means of communication, the electronic mass media, have been invented. These have lowered the effective cost of advertising very substantially. As a result, the cost-minimizing distribution system has shifted away from high-cost, high-service retail outlets toward low-service retail plus manufacturer's advertising (or cooperative advertising between manufacturers and retailers). Over time, the number of messages received by a typical consumer has increased, and the quantity of point-of-sale informing/convincing services has decreased. These are simply the responses of the distribution system to the nature of 20th-century economic development in the already industrialized nations.[2]

To the extent that the mass communications media are more effectively used by large firms, this change in the nature of the optimal distribution system also increases the degree of economies of scale in distribution. Any reason for the more effective use of mass media will suffice, whether it is easily conceived of as "economies of scale in advertising" or not. For example, if high costs of writing high quality advertisements lead smaller firms to write less effective ones, that will lead to the scale economies. Similarily, if network television is cheaper than the sum of spot rates in all the cities (a somewhat controversial assertion),[3] then national firms will have lower distribution costs than regional or local ones. In another dimension, broadcast media tend to reach customers and others somewhat indiscriminately: a firm whose products appeal to more people might be able to use them in a more effective fashion. If there are externalities in the consumption of branded products (as when, for example, Michelob beer is to be valued not for its taste but as a demonstration of the buyer's taste) the value of reaching a large body of potential customers indiscriminately may

[2] The crucial distinction drawn here, that between expenditures on advertising and the quantity of advertising, has been recognized by Ehrlich and Fisher, Bowman, and Rosse. None of these studies attempts to investigate the relationship between changing relative prices and changing degrees of returns to scale over time.

[3] *See* Peterman's extremely careful review of the evidence for the case that this effect is small. My reading of Peterman's results is that it is 10 to 15 percent cheaper for a national than a regional firm to use television.

be doubled. Any of these reasons, if true, will provide a basis for the explanation of the recent growth in large national-brand firms.

This overall conception of the role of advertising in the distribution system and of the possible economies of scale that have only recently arisen is complementary to some recent thinking in the economics of advertising. In a large body of work, Porter has recognized the interaction between the retail sector and manufacturer advertising. He draws the useful distinction between "convenience" goods—those which can be bought in low-service retail environments, and nonconvenience goods, in which the retailer, not the manufacturer, retains a substantial part of the informing/convincing function. In the language of this paper, we would say that not all goods are alike in the ability to substitute mass communication for face-to-face communication, and that at any time the extent of replacement of high-service retail with low-service plus manufacturing advertising is determined by the characteristic of the particular good. Thus, over time, more and more goods will become "convenience" goods as the price of the newer distribution form falls relative to the traditional one. Porter goes on to make the case that it is primarily the information structure of the market—frequency of purchase, (dis)value of buying the wrong product, extent of objectively verifiable quality attributes, etc., which determine the form of distribution for any particular good. Blair, and Mueller and Rogers make the empirical case that much of the increase in concentration in recent times have come in advertising-intensive industries, especially in industries where television advertising is important.

In related work, Steiner emphasizes that the role of the retail sector is crucial in understanding the economic role of advertising. A more heavily advertised product can expect to be carried by more dealers, to have smaller differences between wholesale and retail prices, and to be more effectively pushed by retailers. From the point of view of this paper, this establishes a crucial link in the way the transition to the low-cost retail sector is established. The manufacturer does not need to explicitly force a change in distribution channel type. A move to a long-term strategy of higher advertising will lead to an adjustment in the retail sector. Another observation by Steiner suggests that the interaction between advertising and retail services may lead to further scale economies in the use of mass media. Suppose that retailers do not allocate shelf space proportionally to anticipated demand, but instead give no space to unimportant brands and extra space to highly demanded brands. This will lead to a tipping phenomenon in manufacturer advertising. The manufacturer whose messages produce the most retail store floor traffic will find that they generate sales more than proportionally.

The empirical part of this paper has two main goals: empirical investigation of the overall conception of the distribution system and empirical investigation of the question of returns to scale in advertising.

The first goal is pursued by an investigation of the nature of the demand for advertising services by manufacturers of consumer goods. Is it true that advertising, particularly in mass media, is a substitute for high-service retail and a complement for low-service retail? This is the implication of the theory of the distribution system outlined above. The elasticity of the demand for advertising services with respect to the "price" of low-service retail services should be positive, with respect to the "price" of high-service retail should be negative. In the empirical work, the price of retail services will be taken to be the retail gross margin.

The second goal is pursued along two fronts. First, an attempt is made to resurrect the fact that larger firms in the same industry have lower advertising/sales ratios as evidence for more effective use of mass media by larger firms. Second, a new form of analysis is made of the quality of audiences sold to advertisers by different media. It argues that if "size of audience" is a quality variable for network TV, then one of the stories of more effective use by large firms must be true. The empirical strategy is then to see for which media audience size represents quality (above and beyond CPM) and for which ability to address a specific niche represents quality.

2. The Costs of Distribution Services

The total social costs of consumption of manufactured goods contain a large distribution cost component. This includes expenses directly paid by manufacturers or wholesalers, such as advertising, the sales force which visits retail outlets, shipping, etc. It also includes the retailer's gross margin, the difference between wholesale and retail price.[4] In this paper, all of the costs of distribution services, both retail gross margin and advertising, are viewed as if they are paid by the manufacturer.[5] This is the familiar treatment of distribution costs used in the economic analysis of vertical restraints in antitrust.[6] Thus the language throughout will speak of the manufacturer "choosing" the distribution system to minimize costs. This is merely a shorthand for the tendency for manufacturer decisions (as on the amount and form of advertising) plus market forces (in retail trade) to minimize the social costs of providing distributive services.

In the case of distributive services, "output" must be defined to have both a quantity and a quality element.[7] The price the marginal consumer is willing to pay for the good depends both on total quantity offered for sale and on amount of distributive services, summarized as "quality." I will use the notation output is Q, retail quantity sold, and quality is v, value of distributive services to consumers. The manufacturer is thought of as choosing both the level of production, Q, and the quality, v. The inputs include x, a vector of physical inputs

[4] This is appropriately viewed by the manufacturer plus the consumer as a cost, even if retail market power raises it above social costs.

[5] Other distribution costs, such as direct salesmen, promotional expenses, etc., are ignored in the empirical work.

[6] *See* Posner, Chapter 7. Note that that analysis establishes that it is in the manufacturers' interest to minimize total distribution costs, the principal behavorial assumption of this paper.

[7] This point appears to have been made first by Dorfman and Steiner.

used in manufacturing, a, advertising messages in various media, and R, retail services of various kinds. There is a "production function" for Q and v:

$$(Q,v) = \mathrm{F}\ (\mathrm{x},a,R) \tag{1}$$

Here "production function" is in quotes because the "production" of retail quantity sold and of the value of distributive services by advertising and retail services depends not only on technology but also on consumers tastes. The manufacturer (with help from the market for retail services) will minimize total manufacturing plus distribution cost subject to producing a given level of sales and of quality.

$$\min\ <P_x,\ x> \ +\ <P_a,\ a>\ +\ <P_R,\ R> \text{ subject to } (Q,v) = \mathrm{F}(x,a,R) \tag{2}$$

where $<, >$ represents the inner product and P_x are the prices of factors of manufacturing production, P_a the prices of advertising messages delivered, and P_R the prices of retail services. The solution to (2) is the cost function for the entire firm, including both manufacturing and distribution. This is labeled C_e:

$$\text{Cost of entire firm} = C_e(Q,\ v,\ P_x,\ P_a,\ P_R) \tag{3}$$

The demand curve is written in inverse form as

$$P = P(\mathrm{Q},\ \mathrm{r})$$

The firms' overall problem is to maximize profit by choice of Q and v: this always implies cost-minimization in the "production" of Q and v.

The usual variables appearing in empirical analyses of advertising economics have been implicit in the discussion to this point. A, "advertising," appears as expenditures on advertising $<P_a, a>$. Sales, S, is PQ. It is essential to maintain the distinction between the quantity of advertising messages delivered, a, and expenditures on advertising, A. Since the price of an advertising message delivered has been changing over time, A is not a sensible indicator of the amount of advertising firms are doing.

There appears to be some confusion in the literature about the very notion of economies of scale in advertising (or in distribution). Arndt and Simon, for example, seem to feel that it is "conceptually meaningless" to "isolate the effect of advertising" as a determinant of scale economies. This argument turns on an incorrect view of the definition of scale economies. The incorrect view poses the question: if all inputs are increased in the same proportion, would output increase more or less proportionately? If more, then there are scale economies. What is wrong with this view is that it may not be sensible for a firm to increase all outputs proportionately. This definition is therefore only correct when the production function is homothetic. An alternative definition of scale economies (for the single product firm) is the slope of the average cost curve: downward-sloping AC is increasing returns to scale. This can always be derived, even when the production function is inhomothetic.

Suppose that the total cost function can be written in this (separable) form:[8]

$$C_e(Q, v, P_x, P_a, P_R) = C_m(Q, P_x) + C_d(Q, v, P_a, P_R) \qquad (4)$$

where C_m is manufacturing cost and C_d distribution cost. In this case, it is perfectly sensible to define economies of scale in distribution, independent of the question of economies of scale overall. If average distribution costs, $C_d(Q, v, P_a, P_R) / Q$, are declining in Q, then there are perfectly well-defined economies of scale in distribution.

Thus we expect the concept of economies of scale in distribution to be a sensible one.

The empirical hypothesis of this paper is that $C_d(Q, v, P_a, P_R) / Q$, average distribution costs, is now more steeply declining in many consumer goods markets than it was long ago. The origins of this in the technology of distribution follow from a slightly more detailed view of the high-service retail/convenience plus advertising dichotomy suggested in the introduction. The conception of the distributive system is as of providing two distinct services to the consumer: access to the product and "information" broadly construed. In figure 1, there are exactly two ways to provide either access or information: On the left ray, informing/convincing is done through the retail function, and only a small amount of advertising is done. The same amount of sales may be obtained by a point on the right ray, using "convenience" retail and more advertising. The dotted lines are "isoquants," although only the points on the two rays are available. Since access must be provided for every buyer, either high-service or convenience retail services are used in approximately fixed proportion to output. The left ray is higher, which means that a great deal more retail services are used in it.

[8] This is defined as independence of the marginal rate of substitution between any two elements of x from the level of either R or a, and independence of the MRS between any two distinctive factors (a or R) from the level of x. Separability is obviously true in the usual conception of the manufacturing firm, in which the product is produced by one decisionmaking unit and sold by another. For separability to fail, the technology of the firm needs to have effects like this one: Within the firm, marketing both sets the distribution system and has some key input into the manufacturing process, for example, by writing sales forecasts. This dual role links the marginal rates of substitution because the more accurate the sales forecast, the more capital-intensive (*e.g.*, less overtime) production techniques can be used, and the more high-service retail is used, the better the sales forecast will be. I find this particular story unconvincing: such effects are likely to be very substantial only in high-volatility demand markets with extremely strong marketing functions within the firm. These circumstances are met in very few consumer-goods markets, though they are not all that rare in high-technology intermediate goods markets. Strictly speaking, more than separability is needed for this result. The additional assumption needed in the present analysis is that manufactured goods (the retailer's intermediate input) are used in fixed proportion to Q.

Two assumptions about the nature of the distribution technology are contained in the way the figure is drawn. First note that advertising messages expand much less than proportionately to output along each ray. This is the underlying assertion about economies of scale in advertising. Second, retail services are proportional to output along both rays. Since the left ray uses more retail services, it is less curved. As a result, the degree of returns to scale in distribution changes with the price of retail services. At early-twentieth century prices, the left ray is the cost-minimizing expansion path, so that the AC curve is essentially flat. When advertising's relative price falls, the cost-minimizing expansion path becomes the other ray, so that the late-twentieth century AC is now falling.

The assertion behind figure 1 turns on the inhomotheticity in the distribution-services part of F( ). One could draw a more general figure 1—in which there were more than two technologies of distribution, so that the isoquants were smooth. As long as retail services of fixed quality are used in approximately fixed proportions to output and mass media are lower cost (per efficiency unit of message delivered) for large-scale distribution, a fall in the price of advertising services will lead to an increase in the degree of returns to scale in distribution. It clearly is correct to label this phenomenon "increasing returns in distribution" or "increasing returns in marketing"; it is only a very slight misnomer to call it "increasing returns to advertising." Further, it is caused by a complementarity in distribution between convenience retail services and mass media.

3. Observable Implications for Industry Demand for Advertising

Empirical work on the demand for advertising services by medium[9] can shed some light on the truth or falsity of the last section's assertions about the nature of the distribution system and advertising's role in it. In the next section, demand curves for advertising by medium for ten consumer-product industries in a fifteen-year time series will be estimated. In the following section, such demand curves will be estimated for a group of brewing firms.

Data sources on advertising by firms and industries report A_m, expenditures on advertising by medium, so that division by a price index is required to get a_m, advertising messages sent by medium.[10] The definition of price used in the advertising industry is CPM, (cost per thousand), the price of one thousand "exposures" of the message.[11] The definition of an "exposure" varies somewhat

[9] In the empirical work in this paper the media are those covered by *Leading National Advertisers*, an annual trade publication. These are network television, spot television, radio, magazines, newspaper supplements, and outdoor advertising. Though *LNA* reports advertising expenditures, the underlying raw data are surveys of media counting advertisements: these data are multiplied by estimated prices to obtain the expenditure figures.

[10] The functional form—translog—used in estimation here never in fact calculates a_m, since the demand is represented by an expenditure share. Thus the remarks of this paragraph apply to the definition of a "message" that is implicit in the data.

[11] The McCann-Erickson CPM figures used here are price indexes based on a sample of prices of specific advertising services. They are converted to a CPM basis by adjusting by the size of the audience each message will reach.

by medium. In the print media, newspapers and magazines, it is circulation (not readership). In the broadcast media, radio and network and spot television, it is the number of receivers tuned to the program in question (not the number of viewers or listeners). For the outdoor medium, exposures are determined by indexes of automobile traffic. In general, the CPM for an individual audience "Newsweek back page" or "Gilligan's Island 30 seconds" is the price of access to that audience divided by the number of exposures of the message. The CPM indexes for a medium are weighted averages of the CPMs for specific audiences within the medium.

The assumption that the right indicator of efficiency units of advertising is the number of messages delivered is obviously suspect. But this is the assumption used when A_m is divided by CPM to get a_m. The correct treatment adjusts for the changes in the nature of the audience each advertising medium can deliver over time. In this treatment, CPM is the cost of one exposure, not necessarily the cost of one unit of advertising services in the medium.

A great deal of applied marketing research concerns itself with the fine detail of individual program selection within the advertising medium. This research and the time of the manager with responsibility for an advertising budget are focused on the quality aspects. They ask questions like which specific magazines can best deliver which specific kinds of audiences, or what kind of television program is the best context for an advertisement using a celebrity to recommend the product personally. Obviously, these kinds of quality attributes will not show up in a study of media aggregates.

One quality attribute that be can measured at the level of the medium is what marketers call "reach," the size of the audience delivered by the medium.[12] The advantages of greater reach to advertisers is not spelled out with any precision in marketing texts. But excellent nationwide reach is said to be the primary advantage of network television as an advertising medium.[13]

In any given medium, reach could be either a quality or a disadvantage, or neutral. When advertisers are willing to pay less per exposure for greater reach, we infer that they are in a region of diminishing returns to audience size. This might arise because the larger audience contains relatively fewer people who are natural consumers of the product. In this case, "selectivity" not reach is the quality attribute. In the highly selective media, such as radio (excellent targeting of audience by age) or magazines (by interests), total medium reach may be irrelevant or negative. If advertisers are willing to pay more per exposure when reach is greater, we infer that larger audiences are more valuable. Any of the arguments for economies of scale in distribution raised on p. 136, or

[12] Reach is sometimes understood to mean the fraction of a population in a particular geographic area that will be exposed to the message. This definition is particularly attractive in directing retailers' advertising expenditures. The alternative definition, and the one which makes sense for our purposes, is nationwide total population reach, the number of people in total who will be exposed to the message.

[13] *See* Marketing and Media Decisions.

the Steiner argument about advertising/retail interactions (p. 137) imply that the manufacturer will value reach positively. Thus the inclusion of message reach as a quality variable allows a new inference about advertising scale economies.

In addition to the econometric evidence on reach as a quality measure in the sections below, some further direct evidence is available for the network television medium. Television audience sizes and CPMs vary across time of day, time of year, and across individual programs. Is there any persistent relationship between audience size and advertisers' willingness-to-pay for time on the program? Three kinds of evidence are available (in order of increasing convincingness). First, popular programs that attract large audiences have higher CPMs than less successful ones.[14] This is consistent with the hypothesis that advertisers are willing to pay for reach, but also with the hypothesis that that audience attracted to popular programs is the most faddish, and therefore susceptible, audience for advertisers. Second, audience size and CPM are positively correlated across time of day. Table 1 is extracted from Peterman's Table 1 with some further A.C. Nielsen data added. Note that the prime-time audience is not only more expensive than for the other times, it is more expensive per exposure. Since the reach of prime-time programming much exceeds that of other times of day, this is in part evidence for reach as a positive quality attribute. But this time-of-day evidence is also ambiguous: the quality of the audience, as well as the size, could vary by time of day.[15] The observable indicators of audience quality in the last two columns of Table 1 do not strongly suggest this, but unobservable audience quality variation could be driving the price differences.

Further evidence by season of the year is shown in Table 2. The source is again Peterman.[16] The first row gives the ratio of spring and summer network television reach to the winter audience. (The temperature elasticity of viewing appears to be substantial, and to be higher in prime time than daytime.) The second row gives the ratio of "participation" (buying ad time without sponsoring the entire program) prices to the peak winter ones. These fall off more than proportionately to the size of the audience. In the third row, CPM is calculated (from the first two rows) as a fraction of winter CPM. Note that as the audience size declines, CPM falls. That is to say, as the audience size falls, the cost of advertising time falls more than proportionately. Thus the smaller audience must be viewed as somehow less valuable by advertisers. No obvious quality of audience explanation arises for this time-of-year data.

---

[14] *See* McGann and Russell, chapter five for typical figures.

[15] For example, daytime CPMs are much lower than evening CPMs. But very different kinds of people are at home during the day and in the evening. For some advertisers (as sellers of detergent) the daytime audience is more valuable because it contains a higher fraction of homemakers. The average advertiser, however, seems to value the access to entire households (and thus to entire decisionmaking units) that evening television affords. The evening audience is also on average somewhat better off financially, another on-balance positive quality feature of an audience.

[16] From an unnumbered table in his note 6.

There is a consistent relationship between reach and CPM that reveals that television advertisers are willing to pay more for greater reach, though the effect is not extremely large.[17] Some of the evidence for this variation is susceptible to alternative interpretations, but the weight of the evidence seems to be that there is a small premium paid for greater reach. This suggests that reach may be a quality attribute in network TV, and thus that economies of size in the use of network TV may be present.

4. Empirical Specification: Industry Data

In this section, a system of demand equations for advertising by medium is specified for estimation on data from consumer goods industries in time series. The system will then be estimated for 6 industries.[18] The data are annual time series for 1967-1981.

The core empirical assertions are threefold. First, mass-media advertising and convenience retail are asserted to be complements in the demand for distributive services, and mass-media advertising and high-service retail are asserted to be substitutes. This can be tested by looking at the cross-elasticities of demand: changes in the relative prices of high- and low-service retail should shift the demand for advertising by medium. In particular, an increase in the relative price of high-service retail services should increase the demand for mass-media advertising by manufacturers. Second, larger firms should differ from smaller ones in their use of the media because the distribution technology is inhomothetic. This implication will not be testable in a demand analysis at the level of industries. The empirical specification includes an indicator of the size distribution of firms for each industry. The purpose of this variable is not so much to detect any inhomotheticity but to correct for its presence in an econometric sense. Third, reach is asserted to be a quality attribute for mass media, particularly television. This assertion is tested by including reach as a medium-enhancing input quality index, following the method suggested by Berndt.

---

[17] The size of the advertisers' willingness-to-pay differential cannot easily be inferred from the prices alone. If the supply curve of advertising time is upward sloping (because of rising ad congestion costs or other reasons), supply responses will tend to even out price relative to advertisers' willingness to pay differentials.

[18] They are Food Products, Confections and Soft Drinks, Beer, Wine and Liquor, Toiletries, Drugs and Remedies and Soaps, Cleaners and Polishes. The definition of an "Industry" here is taken from the *LNA*. Several observations should be made about this. First, *LNA* "industries" are not always comparable in definition to those used in other data sets or in the SIC. The 6 used here were selected in part because they were easily comparable either to single two- or three-digit SIC industries or to combinations of four-digit industries. Also, all service industries were not included because of difficulties in putting together sensible indexes of distribution costs other than advertising. The tobacco industry was not included because of the television advertising ban. Second, *LNA* figures cover only advertising by the larger firms. In practice, this is the national brands within the industry or the larger regional brands. As a result, the estimates are interpreted as the demand for advertising by such firms.

The functional form used here is the translog. It has its dependent variables s_{mi}, the share of medium m in total advertising expenditures of industry i.

$$s_{mi} = \frac{A_{mi}}{\sum_{n=1}^{M} A_{ni}}$$

for M-1 of the M media.[19] In this definition, m stands for the medium at hand and n indexes all M media. The exogeneous variables are:

1. Media Prices: p_m is the log of the CPM index for medium m, divided by the CPM index for medium M (outdoor).[20]
2. Retail Prices: p_r is the log of the relative price of retail services in high-service versus low-service outlets (also in logs).[21]
3. Reach: R_m is an index of audience size for each medium. For the print media, it is the (log of the) total circulation of all newspaper supplements or all magazines.[22] For the broadcast media, it is the (log of the) total audience of all stations at peak times of day.[23]
4. Size distribution: The four-firm concentration ratio, F_i within the three-digit SIC industry most closely corresponding to the industry definition is used.[24]

The actual estimating equation takes the form:

---

[19] In the empirical work, M is 6 and it is the equation for the outdoor medium which is dropped. Thus the dependent variables are the shares of newspapers (supplements), magazines, network tv, spot tv, and radio in total advertising expenditures by the industry. The practice of dropping the last equation follows from a difficulty if all are estimated—the error structure is singular because the last share can always be calculated from the rest. When estimation is by 3SLS, as here, the results do not change depending on which equation it is that is dropped, because of an algebraic identity.

[20] As a result, the homogeneity of the cost function in prices is imposed.

[21] The definitions of high-service and low-service (convenience) retail outlets are taken from the Technical Appendix to Porter. The price of retail services in a particular outlet form is calculated as the gross margin in that outlet times the wholesale price index for the industry. The mix of retail outlets used within the high or low service retail sector by a particular industry is (again following the Porter Technical Appendix) calculated from the 1963 Census of Business. Thus the price index for each type of retail services varies by industry for two reasons. First, the (fixed) weights vary because different industries used different kinds of outlets in 1963. (The confections/soda pop industry does not use any shoe stores, nor does the beer, wine and liquor industry use any department stores.) Second, the producer price indexes for the different industries increase at different rates, so that the same gross margin is relatively more expensive for an industry with higher prices later on.

[22] Source: *Circulation*.

[23] Source: A.C. Nielsen Audience Demographics, Radio Fact Book. Note that this definition gives the same "reach" to network and to spot TV.

[24] It is unlikely that there exists a form of inhomothetic cost function that leads to this specification on an exact-aggregation criterion. A natural generalization of the translog cost function leads to inclusion of the sums of squares of logs of output as a size distribution measure. *See* Lau for a discussion of exact aggregation analysis in the face of inhoimothetic production technology.

$$s_{mi} = \alpha_{mi} + \sum_{n=1}^{M} \beta_{m,n}(p_n + \gamma_n R_n) + \beta_{m,r} p_r + \delta_{m,t} t + \delta_{mF} F_i$$
$$\text{for } m = 1,\ldots,M - 1 \qquad i = 1,\ldots,I$$

Interpretation of the parameters α, β, γ, δ shows that this specification captures our main concerns. Note first that α varies both by industry and by medium. Thus, the shares of the different media can vary in an unrestricted way across industries. Second, note that wherever the price of an exposure, p_n, enters, it enters in the form $p_n + \gamma_n R_n$. To interpret this, consider the demand functions for all media written in unlogged form. Let $p_m = log(P_m)$, so that P_m is CPM for medium m, and let $R_m = \log(K_m)$, so that K_m is the audience size. Then the price of medium m would appear as $P_m(K_m)^{\gamma_m}$ in the unlogged demand functions. If γ_n is negative, an increase in reach lowers the effective price of the medium. This means that reach is being treated as a purely own-medium enhancing quality variable.[25] If the β's are all zero, then changes in the relative prices of different media do not affect the expenditure shares. In that case, all of the own-price elasticities of demand for media will be -1, and all of the cross-price elasticities will be 0. The parameters β_{mn} allow the elasticities to vary. β_{mm} determines the own-price elasticity of the demand for media m, while β_{mn}, m distinct from n, determines whether m and n are substitutes or complements in demand. Symmetry ($\beta_{mn} = \beta_{nm}$) has been imposed, as have the adding-up restrictions on the αs and the β_{mn}s.[26] Similarly, β_{mr} determines whether m is a substitute or a complement for a high level of service at retail.[27]

The parameters δ are included to account for the absence of Q or v as exogeneous variables. In large part, size and quality disappear from the right-hand side because the dependent variable is share, not level. Thus only the extent to which large size or high quality tends to use the medium to an extraordinary extent is there any need to include them. The proxies used here are F_i, which as suggested above is intended to capture the size-effects (in this aggregate data) and time, which is intended to capture any long-term trends in distribution quality not resulting from movements in relative prices.

Estimation is on the five media equations simultaneously by 3SLS. The prices of the media are taken to be exogeneous, as are R_m and the retail prices.[28]

[25] Lau's and Berndt's results show that this form will hold whenever the "production function" has a_m and K_m in it only though $a_m / (K_m)^{\gamma_m}$. That is, increases in the quality of a medium have the same effects on the marginal product of other media as an increase in quantity would have. The interesting practical case that this rules out is direct impacts of the quality of one medium on the effectiveness of another. In their work, for example, Lau and Berndt are concerned that the quality variable "fuel efficiency" attached to machines directly affects the marginal physical product of fuel, not that of machines, so that this quality variable is incorrectly treated as purely machine-enhancing. There are obvious analogs to this situation in the advertising demand.

[26] The adding-up restrictions do not appear explicitly because the demand for the retail services equations are not explicitly treated.

[27] The sign of β_{ij} does not alone determine whether i and j are complements or substitutes. *See* Denny et al. for the relevant formulae.

5. Empirical Results: Industry Data

The results of this estimation are presented in table 3. The first thing to note about this table is that several extremely insignificant variables have been deleted from the analysis. These are the concentration ratios and all of the reach variables except that for network television. The insignificance of the concentration ratios is disappointing, because it suggests that the inhomotheticity was not present. The insignificance of the reach measures is less troubling, since the most compelling reach-as-quality story is that for network TV. The coefficient of $R_{TV\text{-}Network}$ is negative and quite significant statistically, but quite small in economic terms. This reinforces the impression of section 3 that audience size is a positive quality attribute to buyers of network TV ad time. But it also reinforces the impression of that section that the size of this effect is not great.

The other coefficients of primary interest relate to the substitution between retail services and advertising demand by medium. The coefficients $\beta_{m,\ retail}$ for each of the five advertising media imply that a fall in the price of low-service retail relative to high-service retail will increase the fraction of ad budgets going to magazines and to network TV, and will decrease the fraction going to radio, newspapers, and spot TV. This strongly supports a view of the distribution system in which the nationwide media—magazines and network television—are complements for low-service retail, and local media—radio, newspapers and spot TV—are complements for high-service retail.[29] Although this was not exactly the view the introduction suggested, it is nonetheless consistent with the overall view of the distribution system exposited in this paper.

It is important at this juncture to say precisely what has been established. The inhometheticity of the demand system does not necessarily imply increasing returns.[30] Instead, it implies that the degree of returns to scale rises when the relative price of national advertising media falls. It is possible that the results imply only that the degree of *decreasing* returns to scale in distribution is now less than in earlier times. Only under the further (untested) assumption that retail services are used in fixed proportion to sales does the increasing returns result follow.

[28] The defense of the assumption that price is exogeneous is the usual one in the systems of demand equations literature: errors in the demand equations of the firms at hand are unlikely to be sufficiently large to affect prices. A slightly superior version of that defense is available in this case because there does seem to be some evidence that the long-run movements in ad and retail prices are driven by forces other than shifting demand by advertisers. The clearest long-run trends are the fall in broadcast media CPMs relative to print media, which is clearly consistent with cost shifts in the electronic age, and the rise of high-service relative to low-service retail costs, probably driven by long-term cost trends. Similarly, the response of newspaper ad prices to the 1970's bulge in the price of newsprint is marked. The fall in television ad rates following the ban on cigarette advertising is demand-driven but can clearly be taken to be exogeneous to the other industries studied here.

[29] The sense of "local" here must be understood precisely. The advertising figures used here are manufacturer advertising. They include cooperative expenditures, in which the ad time is bought for the manufacturer by a retailer (or in which costs are shared) but exclude advertising by retailers.

[30] I owe this point to James Ferguson and Richard Kihlstrom.

6. Empirical Results: Brewing Firm Data

This section estimates the same model as the last one, but on a time-series cross-section of six brewing firms[31] rather than whole industries. The same 15-year time-series of annual data is used. The primary advantage of this change is that it will allow the explicit treatment of firm size—the inhomotheticity of the distribution "production" function can be directly assessed. This is accomplished by including firm quantity[32] as an exogeneous variable in each equation. Thus the estimating equation now takes the form:

$$s_{mf} = \alpha_{mf} + \sum_{n=1}^{M} \beta_{m,n}(p_n + \gamma_n R_n) + \beta_{m,r} p_r$$
$$+ \delta_{m,t} t + \beta_{mQ} Q_f$$
$$\text{for } m = 1, \ldots, M - 1$$
$$f = 1, .., 6$$

where f now indexes firms, and Q_f is the (log) quantity. For brewing firms, maintaining the six separate advertising media is not feasible. Instead, magazines, spot TV and network TV remain, but the other three are aggregated into a single medium for which a price index is constructed.[33] Thus M = 4 for this analysis. Otherwise, estimation is as in the last section.

The insignificant variables in this analysis include all of the reach measures, including network television, as well as the time variable in each of the three estimated share equations. The price index for the combined fourth medium is also everywhere very insignificant. As before, the specification without the extremely insignificant coefficients is the one reported. The insignificance of $R_{TV\text{-}Network}$ may arise because it measures the wrong audience. It was defined as peak-time audience, and brewers do not buy much prime television time. Their expenditures tend to be concentrated in sports programming, a distinct audience.

The results about substitution between retail services and advertising media confirm those of the preceeding section. Here, magazines and network TV are complements for low-service retail, spot TV is a complement for high-service retail.

The demand functions show considerable inhomotheticity in the estimates of β_{mQ}. An increase in firm size will tend to increase the share of network TV at

[31] Pabst, Anheuser-Busch, Coors, Schlitz, Miller, Olympia.

[32] Defined as the log of production of beer in barrels.

[33] This is done because the fourth "medium," outdoor plus radio plus newspapers, rarely exceeds one-tenth of any of these brewers ad expenditures.

the expense of spot TV and of magazines. This result shows up even though the firm-specific intercepts for each of the demand equations have been included. Thus the cross-section variation (A-B tending to use more network TV than Coors) does not enter into the estimated coefficient on Q. Only the tendency of the individual firm to change its media mix when it grows larger affects this estimate. (It also is independent of the simple advance of time, as shown by the insignificance of the δ_ts.)

Thus the brewing-firm data tend to confirm a central part of the story: that the distribution "production" function is inhomothetic, with large scale tending to be Network-TV and Magazines using. It is not surprising that these national media are the ones more favored as brewing firms grow larger. Over the sample period, an important source of growth for brewers was conversion of regional brands into national ones.

7. Advertising Scale Economies and the A/S-Size Relationship

There is a great deal of controversy about the interpretation of a common empirical finding in the advertising literature. Within industries, the advertising/sales ratio is higher for smaller firms.[34]

The controversy is between two competing intuitions. One intuition is usually presented by the scholar making the finding. It treats A/S as the "unit cost of advertising" and says that it is lower for larger firms, hence larger firms are able to buy the "needed" advertising services cheaper. This intuition is then used to interpret the finding in a way that suggests that the "need" to advertise constitutes an important source of lower average costs for larger firms in the industry(ies) in which the phenomenon is observed. This will lead to more concentration and thus more market power. The other intuition is built around the observation that advertising is a choice variable of firms. The uses of "unit," "cost" and "need" in the preceding argument were all clearly incorrect. A/S is the ratio of two dollar quantities, so that the "unit" label is clearly wrong. Advertising is a choice variable, and presumably influences sales through the

[34] See for example Brown. For the purposes of this paper, I will take this to be an established fact, and will take it to mean that all other things equal, a "larger" firm will spend less on advertising per dollar of sales. Here larger might mean demanded by more persons, or with larger manufacturing capacity, or available in more regions. This obviously sweeps under the rug a great many real concerns about whether this finding is a finding. For example, many scholars have suggested that larger/smaller firms are in different market segments. Thus the tendency of larger firms to behave differently may not necessarily have represented an all other things equal experiment. See Albion and Farris, Ch. 5, and Simon and Crain for this criticism. Nothing in this paper will resolve these issues: only the question of interpretation is addressed. I note in passing, however, that the simple observation that large firms do less advertising because they have more goodwill does not constitute a case against declining A/S finding. If larger firms generate more goodwill because they are large, then that is a clear economy of scale in distribution. If, as I presume these critics are actually saying, firms are large because they have good products, and they therefore need to advertise little as well, the critic's case is in fact made.

demand curve, so that the "cost" label is clearly wrong. If larger firms do indeed get more services per advertising dollar, then they might choose to do more advertising as a result, so that the "need" language is clearly wrong. (This counter-argument is made in very abbreviated form in Simon-Arndt.) Thus the conclusions of the first-intuition are not warranted.

This question can be posed in a simple way: will the dollar amount of advertising "A" divided by dollar sales "S" be larger or smaller for larger firms if there are economies of scale in advertising? The two intuitions are directly opposed: if advertising services, in efficiency units, are cheaper on the average for big firms, then they will spend less on them. On the other hand, the firm should respond to a cheaper input by buying more of it, so that we should expect big firms to buy more efficiency units of advertising. The unifying concept here is that of the elasticity of demand for advertising. Expenditures on an input go up or down with price depending on the elasticity of demand for that input, whether it is greater or less than -1. Thus intuition 1 above corresponds to thinking in which advertising is inelastically demanded, and intuition 2 to thinking in which it is elastically demanded. A simple model shows that this unifying analysis is correct, but that -1 is not the crucial cutoff for the demand elasticity in the A/S case.

The analytical model for this issue conceptualizes advertising scale economies in a simple way: larger firms can buy efficiency units of advertising cheaper. This model is designed to represent a situation somewhat like the following: suppose there could be either regional and national firms. Each regional firm faces the same demand curve for its product as a function of price and of efficiency units of advertising. There is no demand interaction between regions. Since regional firms do not value messages delivered outside their region, each buys spot TV exclusively. The national firm, on the other hand, would have as its demand curve the sum of those of all the regional firms—meaning that sending the total of all the regional firms' advertising messages and charging the same price will lead to the same sales as all of the regionals together. But the national firm can buy network TV, which is cheaper per message delivered than spot. When would A/S be larger for the national firms? In this (obviously stylized) situation, it is easy to get the answer in analytical terms.[35]

For simplicity, assume that advertising in a single medium is the only choice

[35] This analysis can be reinterpreted very easily for a slightly less stylized situation, that of a "popular" versus an "unpopular" brand. Suppose that the only available media are mass media, that is, that the manufacturers of unpopular brands cannot buy audiences which contain a large fraction of their natural consumers. Then the analysis goes through as written, with P_a interpreted as cost per thousand exposures ***to the brand's customers***. Thus the "price" will be higher for the unpopular brand. The analysis needs some small alteration to deal with another slightly less stylized situation. Suppose that the expenditure function per efficiency units of advertising, A(a), does not take the linear form $A = P_a a$ but instead exhibits increasing returns. Now suppose that large firms and small firms differ in the size of the market over which they can spread those diminishing returns. This (different) model yields results essentially like those reported in this section, but its analytics are not reported here.

variable in the firms' distribution system. Label the price of advertising messages delivered as P_a and the quantity of messages delivered as a, so that a is efficiency units of advertising for this case. Then advertising expenditures, A, are given by A = $P_a a$. There is also a demand curve (written in inverse form) given by $P = P(Q,a)$, where Q is quantity sold and P price. We interpret $P(\ )$ either as the demand curve of the representative regional firm or of the national firm—an interpretation which makes sense given that there are no regional demand interreactions, and if the demand elasticities are the same in all regions. We assume only that marginal revenue is downward sloping in Q and increasing in a. The firm picks Q and a so as to maximize

$$QP(Q,a) - mcQ - P_a a \tag{5}$$

where mc is the (constant) marginal cost of production. Note that in this simple situation the intermediate step of defining a distribution cost function is skipped. Instead, the approach is to directly maximize profits. The solution to the manufacturer's problem is given by the a and Q which solve

$$P(Q,a) + QP_1(Q,a) = mc \tag{6}$$

and

$$QP_2(Q,a) = P_a \tag{7}$$

where subscripts to $P(\ )$ indicate partial derivatives. These equations directly imply two familiar results. Let $-\eta_p$ and η_a be the elasticity of quantity demanded with respect to price and a respectively. (Thus these are the elasticities of the structural demand curve, not of any equilibrium relation.) Then we have:

$$P(1 - \frac{1}{\eta_p}) = mc \tag{8}$$

directly from the (6). To manipulate (7), note this relationship between the derivatives of the inverse demand curve ($P = P(Q,a)$) and of the direct demand curve ($Q = Q(P,a)$):

$$\frac{Q_2}{Q_1} = -P_2 \tag{9}$$

Substituting this into (7) above yields:

$$-Q\frac{Q_2}{Q_1} = P_a \tag{10}$$

Multiplying and dividing on the left by P, Q and a leaves:

$$-(\frac{Q}{P}/Q_1) \cdot (\frac{a}{Q} \cdot Q_2) \cdot P \cdot \frac{Q}{a} = P_a \tag{11}$$

which is the familiar Dorfman and Steiner result,

$$\frac{\eta_a}{\eta_p} = \frac{A}{S} \tag{12}$$

An immediate implication of (12) is that, when η_p and η_a are constant, A/S will be invariant in P_a. So for the constant demand-elasticity case, variations in A/S by size of firm cannot indicate either that larger firms buy advertising services more expensively or more cheaply. They must indicate some other difference in the environment.

By going on to get the comparative statics of A/S in P_a, we can see that, depending on the demand system, A/S can either be increasing or decreasing in P_a. Thus a monotone relationship between A/S and S could indicate either more or less expensive efficiency units of advertising for larger firms. But the comparative statics analysis will also reveal the observable conditions under which each inference is warranted. First define $MR(Q,a) = P(Q,a) + QP_1(Q,a)$. Then a slightly involved comparative statics exercise shows that:

$$\frac{\partial a}{\partial P_a} = \frac{MR_1}{\triangle} < 0 \tag{13}$$

$$\frac{\partial Q}{\partial P_a} = \frac{MR_2}{\triangle} < 0 \tag{14}$$

where

$$\triangle = (MR_1 \cdot QP_{22}) - (MR_2)^2 \tag{15}$$

With this in hand, we can move toward the comparative statics of A and S. First note that

$$\frac{\partial A/S}{\partial P_a} \gtrless 0 \text{ as } \frac{1}{A}\frac{\partial A}{\partial P_a} - \frac{1}{S}\frac{\partial S}{\partial P_a} \gtrless 0 \tag{16}$$

Using the definitions of A and S, it is easy to see that

$$\frac{1}{A}\ \frac{\partial A}{\partial P_a} = \frac{1}{P_a}(1 + \eta^a_{P_a}) \tag{17}$$

where the notation $\eta^a_{P_a}$ denotes the elasticity of a with respect to P_a, determined by the firm's optimum, this is the demand elasticity for advertising services, and:

$$\frac{\partial S}{\partial P_a} = P\frac{\partial Q}{\partial P_a} + Q\left(P_1\frac{\partial Q}{\partial P_a} + P_2\frac{\partial a}{\partial P_a}\right)$$

$$= P(1 - \frac{1}{\eta_p})\frac{\partial Q}{\partial P_a} - \frac{\partial a}{\partial P_a}\,P\frac{Q_2}{\eta_p}$$

where the last equality uses MR $= P(1 - 1/\eta_p)$. Hence:

$$\frac{1}{S}\frac{\partial S}{\partial P_a} = \frac{1}{P_a}\left((1 - \frac{1}{\eta_p})\eta^Q_{P_a} - \frac{A}{S}\eta^a_{P_a}\right) \tag{18}$$

These are sufficient to prove the result of central interest:

$$\begin{aligned}\frac{\partial A/S}{\partial P_a} < 0 &\leftrightarrow (1 + \eta^a_{P_a}) - (1 - \frac{1}{\eta_p})\eta^Q_{P_a} - \frac{A}{S}\eta^a_{P_a} < 0\\ &\leftrightarrow |\eta^a_{P_a}| > \frac{1 + (1 - \frac{1}{\eta_p})(-\eta^Q_{P_a})}{1 + A/S}\end{aligned} \tag{19}$$

As expected, the more elastic the demand for advertising services is, the greater the tendency for A/S to be ***higher*** for the firm that buys advertising more cheaply. But the cutoff elasticity is not unity. Instead, it depends on market power, on the responsiveness of the size of the firm to the price of advertising, and on the advertising sales ratio. The counterintuitive result is that the low-P_a firm has the high A/S. This result is made more likely by a higher A/S, by more market power, or by a lessened responsiveness of the size of the firm to P_a.

The meaning of this analysis for the interpretation of the empirical finding should now be made clear. Instead of being evidence for economies of scale in advertising, the higher A/S numbers for smaller firms could be evidence for the reverse. This counterintuitive possibility does not require extreme values for the elasticities at all. It is an empirical question whether the higher A/S for small firms is evidence for or against scale economies.

This question can, fortunately, be resolved empirically. The brewing industry studied in the last section allows a sensible test, since it exhibits the declining

A/S ratio. Anheuser-Busch, by far the largest firm, has an A/S (in 1982. Source: Advertising Age.) of .013. The average of the next four firms is over .005. Is this evidence that Anheuser-Busch faces a decreasing returns to advertising, as Arndt and Simon suggest, or that efficiency units of advertising can be bought more cheaply by the giant firm, as Brown and Peles would interpret it?

The question turns on the sign $\partial(A/S)/\partial P_a$, which can go either way analytically. This section finishes by estimating that sign, using the brewing firm data.[36] The model of this section, unlike the cost-minimizing model behind the estimating equations of the previous two sections, is a model of the equilibrium of the whole firm. That is, the derivative whose sign we wish to determine is along the reduced-form of the entire firm's equilibrium, not any structural equation. Thus we regress (A/S) on quantitative measures of the firm's environment, but not on any choice variables of the firm.

For this test, we use the advertising and sales data on only three brewing firms: Anheuser-Busch, Coors, and Pabst[37] for the same 15-year period. The dependent variable is log(A/S). The intercept is allowed to vary across the three firms, but all of the slope coefficients are the same. It is not obvious what the determinants of A/S over time should be. The ones used here are the (log, real) price of advertising,[38] disposable income, brewing average costs,[39] and a dummy for the period 1974 on to capture the effect of the invention of "light" beer.[40]

The results are presented in table 5. The question of primary interest, the elasticity of A/S with respect to P_a, seems resolved in favor of a positive coefficient. The estimate is positive in all three specifications. The exact magnitude of $\partial(A/S)/\partial P_a$ is not clear from the estimates, since it seems to be quite sensitive to specification.

The R^2 figures presented in table 5 are the percent of variances in the (A/S) data explained by the exogeneous variables other than the three firm dummies. That is, they are calculated from the residual of the errors in the specification and the residual variance in a specification which has only the three firm dummies in it. The fit is excellent, given this definition.

In time series, when the real price of advertising falls, brewing firms A/S fall

[36] Note that the evidence of the last section did not bear on this question, since it asked only about the allocation of advertising expenditures within the total ad budget. It did not estimate any equation determining the size of the total.

[37] In the brewing industry, S is rather more difficult to observe than Q. I am grateful to Professor George Foster of the Stanford Business School for help in obtaining data on the dollar sales of beer of these three firms.

[38] This is a Tornquist price index of the media CPMs using brewing-industry weights. *See* Baker and Bresnahan for precise descriptions of all of the data used in this section.

[39] As defined in the census of manufacturers but excluding advertising. This is the average for all brewers.

[40] The purchase of Miller Brewing by a cigarette company was followed not only by a substantial increase in advertising per barrel by that (and later other) firm but by a burst of new-product introductions. These are broadly believed to have changed the nature of competitive interaction in the industry.

as well. Analytically, this means that the demand for advertising services by brewers is not very elastic. But it also means that the tendency of larger brewers to have lower A/S is consistent with their efficiency units of advertising cheaper than their smaller competitors.

9. Conclusions

In this paper, evidence on the role of advertising in the distribution system has been assembled. The overall conception of the distribution system suggested in the introduction—that high-service retail is a substitute for, and low-service retail a complement with advertising—has been modified slightly. If it is taken to mean advertising in the national media, network television and magazines, then the conception is correct. For other forms of manufacturer advertising, the substitution/complementarity results are reversed.

Some further evidence that the social production function for distributive services is inhomothetic has also been presented. There is direct evidence of this from the brewing firms' demand for advertising. This implies that a fall in the price of network television, or of its complement, low-service retail, will increase the degree of returns to scale. As a result, we can conclude either that there are increasing returns to distribution at modern prices or (at least) that decreasing returns have grown less important over time. Either finding is consistent with the theory that changes in the prices of distributive services contributed to the growth of concentration in consumer-product industries in the postwar era.

Two other kinds of evidence were presented. Further evidence that network television reach is a quality attribute was available from the relative prices (per exposure) of advertising time for different size audiences. A conceptual difficulty with studies using the A/S ratio was resolved in favor of the interpretation that large firms' lower A/S indicates increasing returns to advertising.

In short, a great many methodologies for assessing economies of scale in advertising suggest that there are in fact some scale economies. Of these, only the A/S studies alluded to rely in any way on differences between large and small firms in cross-section. None of the results suggest extremely large economies, but all of the analysis was done on fairly substantial firms. Thus the possibility of more important scale economies at smaller scale cannot be ruled out.

The empirical results provide an explanation of the rise of the firm specializing in national brands through an increase over time in the degree of manufacturer scale economies in the use of the distribution system. That increase is driven by changes in the relative prices of different parts of the system, especially high— versus low—service retail, and by the invention of new national media, such as television. The evidence here is from a relatively short time period after the rise of the national-brand firms, and was not limited to firms already national in scope. Thus it is independent of the long-term trend, tending to confirm one theory of that trend.

Since a recent review of the literature on economies of scale in advertising concluded that the assertion that economies exist is a "myth," (Arndt and Simon) it is appropriate to relate the results found here to alternative approaches.

1. *Production Functions for Sales*. Studies in which advertising combines with other inputs to produce sales, and in which this production function is directly estimated, are unlikely to reveal much about scale economies. The results presented here indicate that the distribution technology is quite inhomothetic. In these circumstances, drawing a returns-to-scale inference from an estimated production function is extremely difficult.

2. *Advertising Response Functions*. Studies of the way in which consumers respond to increasing levels of advertising messages (a special case of the production function approach) cannot reveal anything about economies of scale in distribution. A rational firm will always operate in a region of diminishing return to advertising expenditures. Further, advertising response functions implicitly ask the question, would sending more messages to the same audience have increasing or diminishing returns? This is simply the wrong question. For example, it does not capture any distribution cost savings that might arise if regional firms were replaced by a national firm. In this case, both the size of the firm and the size of the audience increase. The to-be-explained fact is the rise of the national branded consumer product, and of the large firms selling such products. Thus the change from regional to national firms is more than hypothetical.

I conclude by observing that this paper has no direct welfare-economic implications. Just as changes in transportation costs in the last century led to exploitation of scale economies in manufacturing, changes in communication and distribution costs led to exploitation by national-brand firms of scale economies in the distribution function. In both cases, there is probably an increase in market power. But there is also a social cost saving due to the increase in firm size. Thus, no welfare inference is warranted.

Figure 1

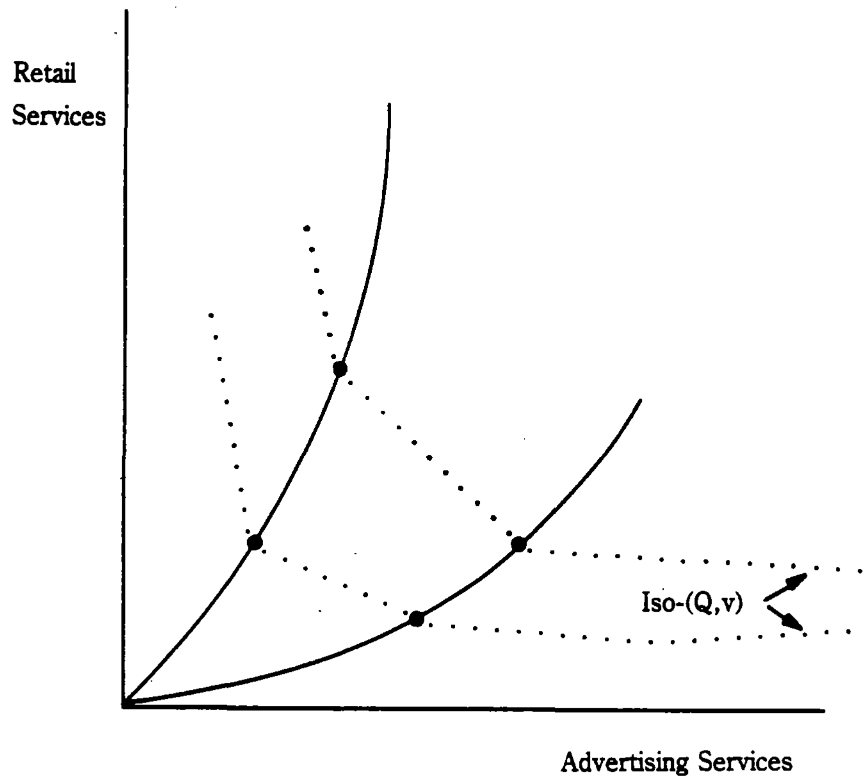


Table 1

AUDIENCE SIZE AND CPM BY TIME OF DAY: NETWORK TV

WINTER 1976-1977

Time of Day	Audience Size[1]	CPM/[2] Prime Time CPM	Household Income[3] > $15,000	Small[4] Household
Prime Time[5]	64.5	1.00	112	112
Late Evening[6]	26.9	.96	107	109
Early Evening[7]	49.0	.90	97	108

[1] In millions of HUT: year-round average. Source: Nielsen Audience Demographic Report.
[2] NBC and CBS. (Network ad rate at this time divided by network ad rate in prime time)/(Network HUT at this time / Network HUT in prime time.) Adapted from Peterman, Table 1, p. 553.
[3] (Percentage of this type of household watching at this time) / (percentage of all households.) Source: Nielsen Audience Demographic Report.
[4] As in 3. "Small" is one or two persons.
[5] 8:00 to 11:00 p.m.
[6] 11:00 a.m. to 12:00 a.m.
[7] 6:00 to 8:00 p.m.

Table 2

AUDIENCE SIZE AND CPM BY TIME OF YEAR: NETWORK TV

1965-1966

	Daytime Spring	Daytime Summer	Evening Spring	Evening Summer
HUT/(Winter HUT)	92.4	88.5	85.7	72.3
Price/(Winter Price)	89.7	86.7	83.3	68.3
CPM ratio	.97	.98	.97	.94

Table 3

DEMAND FOR ADVERTISING MEDIA BY INDUSTRY

Parameter	Estimate	Standard Error
α_{mi}	(suppressed)	(suppressed)
$B_{\text{Newsp.}, \text{Radio}}$	-0.18372E-01	0.19586E-01
$B_{\text{Maga.}, \text{Radio}}$	0.14193E-01	0.24872E-01
$B_{\text{TV-Netw.}, \text{Radio}}$	-0.63726E-01	0.26681E-01
$\gamma_{\text{TV-Netw.}}$	-0.22769E-01	0.52237E-02
$B_{\text{TV-Spot}, \text{Radio}}$	0.20660E-01	0.23594E-01
$B_{\text{Radio}, \text{Radio}}$	0.28526E-01	0.42035E-02
$\delta_{\text{Radio}, \text{time}}$	0.25435E-02	0.10973E-02
$B_{\text{Radio}, \text{retail}}$	-0.14764E-01	0.50591E-01
$B_{\text{Newsp.}, \text{Maga.}}$	0.25941E-01	0.23265E-01
$B_{\text{Maga.}, \text{Maga.}}$	-0.14488E-01	0.12542
$B_{\text{TV-Netw.}, \text{Maga.}}$	-0.40546E-01	0.87374E-01
$B_{\text{TV-Spot}, \text{Maga.}}$	0.38283E-01	0.12815
$\delta_{\text{Maga.}, \text{time}}$	0.65724E-02	0.10377E-02
$B_{\text{Maga.}, \text{retail}}$	0.18128	0.12200
$B_{\text{Newsp.}, \text{TV-Spot}}$	0.13750E-01	0.20110E-01
$B_{\text{TV-Netw.}, \text{TV-Spot}}$	0.17328	0.12701
$B_{\text{TV-Spot}, \text{TV-Spot}}$	-0.24928	0.19634
$\delta_{\text{TV-Spot}, \text{time}}$	-0.13260E-01	0.73404E-02
$B_{\text{TV-Spot}, \text{retail}}$	-1.18253	0.37009
$B_{\text{Newsp.}, \text{Newsp.}}$	-0.32778E-02	0.27066E-01
$B_{\text{TV-Netw.}, \text{Newsp.}}$	0.11582E-01	0.83022E-02
$\delta_{\text{Newsp.}, \text{time}}$	-0.5202E-03	0.95783E-03
$B_{\text{Newsp.}, \text{retail}}$	-0.42560E-01	0.43255E-01
$B_{\text{TV-Netw.}, \text{TV-Netw.}}$	-0.48554E-01	.14701E-01
$\delta_{\text{TV-Netw.}, \text{time}}$	0.66011E-02	0.39253E-02
$B_{\text{TV-Netw.}, \text{retail}}$	0.21223	0.62745E-01

Summary Statistics:

Dependent Variable: Radio Share
R^2: .63
Durbin-Watson Statistic: 1.1276
Dependent Variable: Magazine Share
R^2: .75
Durbin-Watson Statistic: 2.4622
Dependent Variable: Spot Television Share
R^2: .69
Durbin-Watson Statistic: 1.6522
Dependent Variable: Newspaper Share
R^2: .53
Durbin-Watson Statistic: 1.2291
Dependent Variable: Network Television Share
R^2: .84
Durbin-Watson Statistic: 1.6126

Table 4

BREWING FIRMS DEMAND FOR ADVERTISING MEDIA

WINTER 1976-1977

Parameter	Estimate	Standard error
$\alpha_{\text{Mags}, \text{A-B}}$	0.20028	0.49580
$\alpha_{\text{Mags}, \text{Pabst}}$	-0.20151	0.50313
$\alpha_{\text{Mags}, \text{Coors}}$	-0.25789	0.50552
$B_{\text{Mags}, \text{Mags}}$	-0.31692	0.29903
$B_{\text{TV-Netw}, \text{Mags}}$	-0.65124	0.38490
$B_{\text{TV-Spot}, \text{Mags}}$	0.35175	0.39498
$B_{\text{Mags}, Q}$	-0.32992	0.50464E-01
$B_{\text{Mags}, \text{retail}}$	0.57239	0.39481
$\alpha_{\text{TV-Spot}, \text{A-B}}$	-0.84237	1.3166
$\alpha_{\text{TV-Spot}, \text{Pabst}}$	-0.42146	1.3090
$\alpha_{\text{TV-Spot}, \text{Coors}}$	-0.26393	1.3067
$B_{\text{TV-Netw}, \text{TV-Spot}}$	-0.60879	0.74739
$B_{\text{TV-Spot}, \text{TV-Spot}}$	0.49092	0.81756
$B_{\text{TV-Spot}, Q}$	-0.95173	0.15061
$B_{\text{TV-Spot}, \text{retail}}$	-1.1127	.51148
$\alpha_{\text{TV-Netw}, \text{A-B}}$	2.0307	1.5464
$\alpha_{\text{TV-Netw}, \text{Pabst}}$	1.9833	1.5433
$\alpha_{\text{TV-Netw}, \text{Coors}}$	1.7850	1.5420
$B_{\text{TV-Netw}, \text{TV-Netw}}$	-2.4139	0.60778
$B_{\text{TV-Netw}, Q}$	0.18626	0.74880E-01
$B_{\text{TV-Netw}, \text{retail}}$	0.94625	.29407

Summary Statistics:

Dependent Variable: Magazine Share
R^2: .80
Durbin-Watson Statistic: 1.2721
Dependent Variable: Spot Television Share
R^2: .71
Durbin-Watson Statistic: 2.0239
Dependent Variable: Network Television Share
R^2: .64
Durbin-Watson Statistic: 1.6536

Table 5

BREWING FIRM A/S REGRESSION RESULTS[1]

A-B	7.9073	6.3817	7.6860
	(9.0977)	(9.8792)	(0.77538)
PABST	10.130	8.6135	9.9181
	(9.0976)	(9.8795)	(0.75586)
COORS	7.7072	6.1862	7.4906
	(9.1000)	(9.8819)	(0.79574)
Y	1.0800		
	(.53996)		
LRAC	-0.42108	0.44358	
	(3.1033)	(3.3488)	
LITE	-0.46221	0.35668	0.36540
	(.47023)	(.25203)	(0.23614)
LRPA	3.5136	.27754	.29679
	(3.0693)	(.0962)	(.0479)
R-SQUARED	0.912	0.889	0.882
DURBIN-WATSON	2.0017	1.8134	1.8135

[1] Standard errors in parentheses. Sample size: 45.

Bibliography

Albion, Mark S., and Farris, Paul W., *The Advertising Controversy: Evidence on the Economics Effects of Advertising*, Boston: Auburn House, 1981.

Arndt, John, and Simon, Julian L., *Advertising and Economies of Scale: Critical Comments on the Evidence*, Journal of Industrial Economics, December 1983, Vol. XXXII, 229-242.

Baker, Jonathan, and Bresnahan, Timothy F., *Estimating the Demand Curve Facing a Single Firm*, Mimeograph, Stanford University, 1984.

Berndt, Ernst R., *Quality Adjustment in Empirical Demand Analysis*, Mimeograph, MIT Sloan School of Management, 1982.

Blair, John, *Economic Concentration: Structure, Behavior and Public Policy*, New York: Harcourt Brace Jovanovich, 1966.

Bowman, Gary W., *Demand and Supply of Network Television Advertising*, Bell Journal of Economics, Spring 1976, Vol. 7, 258-267.

Brown, Randall S., *Estimating Advantages to Large-Scale Advertising*, Review of Economics and Statistics, August 1978, Vol. 60, 428-437.

Dorfman, Robert, and Steiner, Peter O., *Optimal Advertising and Optimal Quality*, American Economic Review, December 1954, Vol. 44, 826-836.

Ehrlich, Isaac, and Fisher, Lawrence, *The Derived Demand for Advertising*, American Economic Review, June 1982, Vol. 72, 366-88.

Lau, Lawrence J., *The Measurement of Raw Material Inputs*, Chapter 6 in Smith and Krutilla, 1982.

Lambin, Jean-Jaques, *Advertising, Competition and Market Conduct in Oligopoly Over Time*, Amsterdam: North-Holland, 1976.

Leading National Advertisers: *Ad $ Summary*, Annual issues.

McGann, Anthony F., and Russell, J. Thomas, *Advertising Media: A Managerial Approach*. Homewood, Illinois: Richard D. Irwin, 1981.

Mueller, Willard F., and Rogers, Richard T., *The Role of Advertising in Changing Concentration in Manufacturing Industries*, Review of Economics and Statistics, February 1980, Vol. 62, 511-520.

Peles, Yoram, *Economies of Scale in Advertising Beer and Cigarettes*, Journal of Business, 1971, Vol. 44, 32-7.

Peterman, John L., *Differences Between the Levels of Spot and Network Television Advertising Rates*, Journal of Business, 1979, Vol. 52, 549-61.

Porter, Michael E., *Interbrand Choice, Strategy, and Bilateral Market Power*, Cambridge: Harvard University Press, 1976.

Posner, Richard A., *Antitrust Law: An Economic Perspective*, Chicago: University of Chicago Press, 1976.

Rosse, James N., *Estimating Cost Function Parameters without using Cost Data*, Econometrica, March 1970, Vol. 38, 256.

Simon, Julian L., and Crain, George H., *The Advertising Ratio and Economies of Scale*, Journal of Advertising Research, September 1966, Vol. 6, 37-43.

Smith, V. Kerry, and Krutilla, John V., editors, *Explorations in Natural Resource Economics*, Baltimore: John Hopkins Press for Resources for the Future, 1982.

Steiner, Robert L., *Understanding the Consumer Goods Economy*, Paper presented at the Federal Trade Commission, 1978.

Comments

James M. Ferguson
Federal Trade Commission

I

Bresnahan poses an interesting question — what explains the relative growth of national brands in the twentieth century? The author makes three arguments. First, he argues that using advertising has become relatively less expensive over time compared to using retailers to provide information to consumers. Second, he argues that the "production" of distribution services, including the information services provided by advertising and by retailers, is inhomothetic. Third, he argues that the electronic mass media are more effectively used by large firms. As a result, the fall in the relative price of advertising services has led to the substitution of advertising for retailer information services and to an increase in the degree of returns to scale in distribution. The author claims he demonstrates that economies of scale in advertising and distribution are greater than before, especially for large firms, the development of which explains the relative growth of national brands. His evidence does not support his claim.

According to Bresnahan, there are two basic forms of retail distribution. In the traditional form the retailer spends a great deal of time informing or convincing the consumer to buy the brand. In the convenience form the retailer only provides low-cost consumer access to the brands. The traditional retail distribution method is labor-intensive, so it has become relatively more expensive with the rise in wage rates. The convenience method of retail distribution is national advertising-intensive, and has become relatively less expensive with the growth of the electronic advertising media. Thus, the cost-minimizing distribution system has shifted away from high-cost, high-service retail outlets toward low-cost, low-service retail outlets "plus manufacturers' advertising (or cooperative advertising between manufacturers and retailers)."

The author argues that while the two different forms of retail distribution are substitutes, high-service retailing and national advertising are long-run *substitutes* in distribution while national advertising and low-service retail outlets are *complements*. Apparently, this argument is based on the view that national brand advertising, especially in the more efficient electronic media, will increase retail sales at low-service retail outlets. However, I see no reason why it would not also increase sales at high-service retail outlets. I see national advertising as a technical complement in distribution to both types of retail outlets as well as a substitute in response to a relative price change. Thus, I disagree with the basic hypothesis of the paper.

Furthermore, nowhere does the author mention the growth of retail advertising in response to the rise in the cost of personal retailer information services. Retail advertising, especially in daily newspapers, exceeds national advertising in television. Since local advertising rates are often substantially lower than national advertising rates, I would expect a major, if not primary, effect of the rise in wage rates and the cost of personal retailer information services to be to substitute retail advertising for personal retailer information services.

The author chooses to define economies of scale as a downward sloping average cost curve, the definition of which does not require a homothetic production function. However, in the presence of specialized resources, average cost is determined by price, so neither the shape nor the level of average cost is very informative. Nevertheless his hypothesis is that average distribution cost now declines more steeply in many consumer goods markets than it did long ago.

He uses the sophisticated translog functional form in estimating a system of industry equations for advertising services by media in each of six consumer-goods industries and additionally for six firms in the brewing industry. Unfortunately, I doubt the relevance of this empirical analysis to the preceding theoretical discussion. I also find the evidence on the relationships between the ratio of advertising to sales and firm size to be even less useful.

The data cover six consumer product categories — food products; confections and soft drinks; beer, wine, and liquor; toiletries; drugs and remedies; and soaps, cleansers, and polishes. I doubt that any of these products involve much retailer informing of the consumer during the period covered by the data, 1967-1981. (The data do not include newspaper advertising, except newspaper supplements.)

Since I have doubts about the approach underlying the empirical analysis, I will only briefly comment on the empirical results. The dependent variable is the share of advertising expenditures in each media in each of six so-called consumer product industries. One of the independent variables is the ratio of the price of retail services in high-service versus low-service outlets. The price of retail services for each type of retail outlet is its gross margin times the wholesale price index for the consumer products industry. It is difficult to understand what this variable measures. It does not measure the retail margins for the specific consumer products in question, nor does it reflect the possibly different prices charged different types of retail outlets by producers.

The author reports that the results indicate that a fall in the relative price of low-service retail distribution will increase the fraction of ad expenditures in these six consumer product industries going to magazines and network television and decrease the share going to radio, newspapers, and spot television. The author claims this evidence strongly supports a view of the distribution system in which the nationwide media — magazines and network television — are complements for low-service retail, and local media — radio, newspapers, and spot TV-are complements for high-service retail. Why only these two media show a rise is puzzling. As he recognizes the theoretical discussion suggested a substitution of all media advertising for retail consumer information services in response to a rise in the price of retail information services, I have an alternative explanation. The advertising data used by the author exclude cooperative advertising purchased by the retailer, whether or not paid for by the manufacturer. Since such cooperative advertising is concentrated in local media, the omissionof this sizable component of advertising can distort the results.

The author recognizes that establishing the inhomotheticity of the demand for distribution services does not necessarily imply increasing returns, because the untested assumptions that retail services are used in fixed proportions to retail sales and that advertising messages increase much less than proportionately to sales may not be correct. The author's troubles are compounded, because the reported lack of significance of the coefficients of the concentration ratio variable suggests inhomotheticity was not present. (He does find it present for the brewing firm data.) He claims that analysis of data for six brewing firms for the same 15-year period confirms that magazines and network television are favored by brewing firms as they grow larger.

In his conclusions, Bresnahan claims that either there are increasing returns to distribution at modern prices or that decreasing returns are less than before. "Either finding is consistent with the theory that changes in the prices of distributive services contributed to the growth of concentration in consumer-products industries in the postwar era." This is incorrect, because the so-called economies of scale are not industry specific. As I pointed out in my book, *Advertising and Competition: Theory, Measurement, Fact* (Ballinger, 1974, pp. 16-17), if there are economies of scale which are not industry specific, the greater minimum efficient size of firm will reduce the number of efficient firms in the economy but need not reduce the number in any industry and thus need not have any effect on industry concentration.

II

Kotowitz and Mathewson argue that advertising may mislead consumers in the short run, while consumers are learning about experience quality. In this situation, price premiums exist because advertising is a substitute for experience quality, not a signal of experience quality. Is it possible, they ask, to reconcile rational expectations with observations of misleading advertising?

They introduce these issues by presenting alleged tests of three competing explanations of advertising using data for two experience good industries, automobiles and life insurance. They interpret the theories in the following way. According to Nelson, greater advertising expenditures indicate better buys — offer more quality per dollar of expenditure. According to Klein-Leffler, greater advertising expenditures signal higher experience quality assured through a larger price premium over average salvageable cost. The Kotowitz-Mathewson hypothesis is that greater advertising expenditures can be a (deceptive) substitute for product quality conditional on the rate of learning by consumers.

The first problem is that the so-called third theory of advertising is a hypothesis about disequilibrium conditions during the process of consumer learning while the other two are statements about an equilibrium. So I don't understand how the same test can be valid for all three theories — either the test assumes full adjustment or does not, but it can hardly do both.

Second, I do not believe the authors have properly characterized the Klein-Leffler model. The relevant variable which measures the incentive to produce high quality is not current advertising expenditures but the stock of advertising capital (compared to anticipated future sales). In fact the relevant variable is the stock of nonsalvageable capital — which includes more than advertising. Furthermore, the price premium that tends to assure quality is the price premium over average recoverable cost, not average total cost.

Third, with respect to the first automobile regressions using the brand dummy coefficients from the Ohta-Griliches hedonic price equations as the dependent variable, the expected sign of the advertising variable depends upon the interpretation of these brand dummy coefficients. Do they represent systematic quality differences between the brand and the base brand not captured by the other variables or do they represent higher costs per unit of quality for the brand? The authors opt for the latter interpretation. Finding that these dummy coefficients are positively correlated with electronic media advertising and with the ratio of electronic media advertising to sales, the authors reject the Nelson best buy hypothesis of advertising, even though the coefficients of total advertising and of total advertising to sales are not significant.

Nevertheless, the authors propose an interesting test to distinguish between their view of advertising and the Klein-Leffler view. The difference in depreciation rates of auto brands from the group average during the first year of ownership is taken as a relative measure of consumer disappointment. If advertising is a signal of experience quality assurance, then according to the authors this depreciation should be negatively correlated with advertising intensity of the model in the previous year. If, however, persuasive advertising is enticing consumers to believe the brand is better than experience shows it to be, then depreciation should be positively correlated with advertising intensity in the previous year. Again, the test does not use the stock of advertising capital. A more serious objection is that the measure of depreciation used is the difference between list price and the value after one year of ownership. This confounds two distinct components — discounts from list price which are given at the time of purchase — a reduction in the quality assuring premium — and depreciation. If the two components could be separated and only the actual post-purchase depreciation used, the authors would have an excellent test. The coefficients of the advertising to sales variables, both total and only electronic advertising, are positive and significant, while the coefficients of the advertising variables are not.

The authors also present some regressions using Canadian data for participating whole-life insurance. The discussion is too abbreviated to ascertain the variables. Even if life insurance is an experience good, I question the advertising variable used, which is the sum of advertising expenditures and salesmen's commissions with commissions apparently a large part of the total. "If advertising is a sunk 'signal' of quality, then so are these sales expenditures." This is incorrect. While salesmen obviously play an informative role, they do not play a quality assuring role, because salesmen's commissions are not a sunk expense.

The lengthy remainder of the paper attempts to develop a model of consumer learning consistent with rational expectations but also consistent with the presence of deceptive advertising in the short run. In my opinion this effort is unsuccessful. No testable implications are derived. Furthermore, the model provides for no policing function by consumers — it never pays consumers to punish advertisers for false advertising.

Comments

*Richard E. Kihlstrom**
University of Pennsylvania

In these remarks I propose to comment on an issue raised by the Kotowitz-Mathewson paper. In concluding their discussion of the reliability of the information provided by advertising, they remark that

> .... in the absence of truth in advertising regulations, no correlation between advertising claims and true quality can be established and hence advertising claims contain no useful information.

In many cases, this observation is undoubtedly correct. It is, however, possible to imagine circumstances in which there are no truth in advertising regulations but in which sellers have incentives to provide buyers with advertising messages that are convincingly informative. In these remarks, I propose to discuss one such set of circumstances. The discussion of the advertiser's incentives will follow the analyses of disclosure by informed sellers contained in the papers of Grossman (1981), Milgrom (1981) and Milgrom and Roberts (1984). It will also be closely related to the recent contributions of Stiglitz (1984), Farrell-Sobel (1983), Farrell (1985) and Matthews-Postlewaite (1985).

Grossman extends the argument of Grossman-Hart (1980) to demonstrate that, when there are significant penalities for lying, sellers who know the quality of their products have an incentive to fully disclose their information to potential buyers. As Grossman notes, for his argument to be correct, communication must be essentially costless and it must be possible to costlessly and accurately detect lying *ex post*. The same argument can be used to demonstrate that if false advertising is subject to penalty and can be costlessly detected *ex post* and if communication is costless, then all information possessed by advertisers will be accurately disclosed. Truth in advertising regulations are, of course, one way of penalizing false advertising.

The question I propose to address is whether, in the absence of truth in advertising regulations, there exists any other economic mechanism that can be expected to substitute for these regulations. The situation I will describe is one in which such a mechanism indeed operates. In this situation, the seller himself will unilaterally take actions that have the effect of introducing a penalty for false advertising. In the particular circumstances to which I am referring, the seller of a product has the option of receiving an informative signal of his product's quality. Once the seller receives the message provided by the signal, he must decide whether or not to disclose it to the buyer. The crucial feature of the situation is that the seller does not, when he makes the decision to become

*In the process of preparing this comment, the author has benefitted from conversations with Franklin Allen, Andrew Postlewaite and Michael Riordan.

informed, know the message that will be conveyed by the signal. Another crucial feature of the hypothesized situation is that the profits the seller receives from an uninformed buyer are less than the expected profits he can anticipate receiving from a buyer whose decision will be made with the aid of the information conveyed by the signal.

Thus, in the situation envisioned here, the seller is assumed to be better off if he sells to an informed buyer than he is when he sells to an uninformed buyer. This assumption effectively means that if *ex ante* the seller could convincingly commit himself to truthfully disclose any message he might receive he would choose to do so. If false advertising can be costlessly detected, the seller can make such an *ex ante* commitment by posting a sufficiently large bond with a third party, who will henceforth be referred to as the trustee. The commitment will be convincing if the seller agrees to forfeit the bond whenever false advertising is detected.

It should be emphasized that, in the particular formalization described below, false claims by the seller about the nature of the information acquired will be included under the heading of false advertising. Thus, the asumption that false advertising can be detected must be interpreted in a strong sense. It asserts that the buyer can costlessly validate claims the seller makes about the reliability of the information he has acquired.

It is not, however, necessary to assume that the buyer is capable of verifying that no information has been acquired if the seller claims that to have been the case. There is a potential problem when the buyer cannot verify such claims. Specifically, if the seller receives an unfavorable message, he can simply report that he has not acquired information; that is, he can claim that no message has been received. This problem can be avoided if the trustee is instructed to award the bond to the buyer whenever the seller fails to make a report concerning the information he has received. In this way the posting of the bond commits the seller to the acquisition of information and requires him to make some form of report on the results to the buyer.

The mechanism just described will function effectively as long as it is possible for the buyer to convince the trustee that false advertising has taken place. The cost of this communication need only be finite. If this cost is significant, the bond will simply have to be large enough to permit the buyer to recoup these costs if he ever has to incur them as the result of a false claim. In the equilibrium described below, no false claims will ever, in fact, be made.

For the purpose of formalizing the argument, I will introduce a simple but general model of the seller's advertising decision. I will assume that the advertising message has content in the sense that it is informative about the quality of the product. The model of the consumer is virtually the same as that used in Section 3 of Grossman's paper. Thus, there are two quality levels, b_1 and b_2, and

$$b_1 > b_2.$$

Initially, both the buyer and the seller believe that the probability of quality level b_1 is π, where

$$1 > \pi > 0.$$

Thus, if the buyer receives no additional information and pays p dollars for one unit of the product, his expected utility is

$$\pi u(b_1 - p) + (1 - \pi)u(b_2 - p).$$

If the buyer's reservation utility level is u_o, the seller, who throughout this analysis is assumed to be a monopolist, will charge him $p(\pi)$ where $p(\pi)$ is the p value at which

$$\pi u(b_1 - p) + (1 - \pi)u(b_2 - p) = u_o.$$

Suppose now, however, that the seller can provide the buyer with an informative signal of quality. In fact, suppose that there is a continuum of possible informative signals that can be supplied to the buyer. Also suppose that these signals can be indexed by their reliability. We let Θ be the index of reliability and use x_Θ to denote the random message provided by an advertising signal of reliability Θ. The advertising messages are assumed to lie in some set X. The probability that a signal of reliability Θ brings message $x \epsilon X$ when the quality level is b, that is, the probability that $x_\Theta = x$ when quality is b, is denoted by $f(x;b,\Theta)$. For Θ to be a justifiable measure of the information's reliability we must assume that, when $\Theta > \Theta'$ observations of x_Θ are more informative (in Blackwell's sense) about quality than observations of x_Θ. Finally, we assume that more reliable information is more expensive to provide than less reliable information. Thus, we let $C(\Theta)$ be the cost to the seller of providing the buyer with information of reliability Θ, and we assume that $C'(\Theta) > 0$.

For concreteness, we can assume that $X = \{x_1, x_2\}$ and that

$$f(x_k;b_j,\Theta) = \begin{cases} \Theta, \text{ if } k = j \\ 1 - \Theta, \text{ if } k \neq j, \end{cases}$$

where $\frac{1}{2} \leq \Theta \leq 1$. In this example, the receipt of the message that $x_\Theta = x_j$ can be regarded as a prediction that the quality level is b_j, and Θ can be interpreted as the probability that the prediction is correct. It is also possible to interpret Θ as the size of a sample taken by the seller.

If the buyer receives the message that $x_\Theta = x$, his *a posteriori* probability of b_1 will be $\Phi(x;\Theta)$ and his *a posteriori* expected utility when he purchases one unit of the good at a price p will be

$$E[u(b - p);x,\Theta]$$
$$= \Phi(x;\Theta)u(b_1 - p) + [1 - \Phi(x;\Theta)]u(b_2 - p).$$

In this situation, the seller will be able to charge $p(x,\Theta)$ where $p(x,\Theta)$ is the p for which

$$E[u(b - p);x,\Theta] = u_o.$$

As noted earlier, we assume that the seller makes his information acquisition decision before he knows the content of the message that will be provided by any information he might choose to acquire. For the moment, we will assume that the buyer knows the reliability of the information supplied by the seller and that the seller discloses the information when it is received. Under these assumptions, if the seller is risk neutral, and if he chooses to acquire information it will be of reliability level Θ where

$$\Theta = \underset{\Theta}{\text{Argmax}}\ Ep(x_\Theta,\Theta) - C(\Theta).$$

In computing the expected price

$$Ep(x_\Theta,\Theta),$$

the seller uses the marginal distribution

$$g(x;\Theta) = f(x;b_1,\Theta)\pi + f(x;b_2,\Theta)(1 - \pi).$$

The seller will choose to acquire information and supply it to the buyer if

$$\underset{\Theta}{\text{Max}}\ Ep\ (x_\Theta,\Theta) - C(\Theta) > p(\pi). \tag{1}$$

When inequality (1) is satisfied, the assumption, discussed earlier, that the seller prefers to sell to an informed buyer is satisfied. We can now show that, under this assumption, the seller will acquire information of reliability Θ and then, as asserted earlier, post a sufficiently large bond with the trustee with instructions to award the bond to the buyer if the seller is ever found to have made a false advertising claim or if no report on the information is made.

For the purpose of defining a bond that is, indeed, sufficiently large, let p_u be the price the buyer would pay if he knew that quality was b_1 and let p_L be the price he would pay for a good of quality b_2; that is, let p_u and p_L satisfy

$$u(b_1 - p_u) = u_o$$

and

$$u(b_2 - p_L) = u_o$$

respectively. The bond G will, as we shall see, be sufficiently large if it exceeds or equals $p_u - p_L$. Recall that the bond must also be large enough to cover any costs the buyer might conceivably incur in convincing the trustee that a false claim has been made.

Suppose then that the seller does post such a bond with the instructions indicated. Suppose also that he invests in information of any degree of reliability Θ. We can now apply the argument of Grossman (1981) to establish that the seller will indeed disclose both Θ and the message when it is received. Note that, when G is large in the sense defined above, no profits obtained because of a false advertising claim can ever make up for the loss incurred when the lie is detected and G is forfeited.

Clearly if the seller is going to follow the strategy of posting a bond and investing in information he will prefer to invest in information of reliability Θ. Inequality (1) implies that the expected profits achieved in this way exceed the price obtained when no information is acquired and there is no advertising. When we interpret $p(\pi)$ as the profits associated with no information, we are assuming that if the seller doesn't acquire information then he doesn't post a bond. We are also assuming that if buyers do not observe the seller post a bond, they ignore any advertising claims he makes. This attitude on the part of buyers is justified by the fact that when no bond is posted, the seller has no incentive to truthfully disclose what he knows.

There are of course strategies other than the ones just discussed that sellers could follow. They could, for example, acquire information and post either no bond or a small bond. It is easy to demonstrate that, because of the assumed attitudes of buyers, the acquisition of costly information by a seller who doesn't post a bond will leave the seller worse off than he would be if he didn't acquire any information. An argument that is slightly more involved can be used to demonstrate that the seller will never prefer to post a small bond.

Concluding Remarks

The analysis presented here is intended as an introductory discussion of the incentives faced by advertisers. In particular, the assumption that false advertising can be detected is particularly strong and not especially realistic. It does, I believe, nevertheless provide a useful basis for an introductory formal investigation of the extent to which the Grossman, Milgrom, Milgrom-Roberts analyses can be applied to the study of informative advertising. The reader is referred to the papers of Farrell-Sobel (1983), Farrell (1985) and Matthews-Postlewaite (1985) for more complete discussions of related issues. The informational assumptions of those papers are somewhat different than those considered here, as is the focus of their discussion. In particular, they discuss the impact of disclosure requirements on the seller's decisions.

It should finally be reemphasized that the assumption that the seller's advertising decision is made *ex ante* is crucial for our conclusions. We specifically consider the case of an advertiser who does not yet have the information he is contemplating providing to sellers. Sellers who are already informed face different incentives than those discussed here.

References

Farrell, J., "Voluntary Disclosure: Robustness of the Unraveling Result, and Comments on Its Importance," in *Proceedings of the U. C. Conference on Regulation and Antitrust*, R. Grieson, editor, Lexington Books (forthcoming 1985).

Farrell, J. and J. Sobel, "Voluntary Disclosure," unpublished notes 1983.

Grossman, S., "The Information Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 1981, 24, 461-483.

Grossman, S. and O. Hart, "Disclosure Laws and Takeover Bids," *Journal of Finance*, 1980, 35, 323-334.

Matthews, S. and A. Postlewaite, "Quality Testing and Disclosure," *Rand Journal of Economics* (forthcoming 1985).

Milgrom, P., "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 1981, 12, 380-391.

Milgrom, P. and J. Roberts, "Reliance on the Information of Interested Parties: Does the Invisible Hand Operate in the Marketplace for Ideas?", Yale University and Stanford University, unpublished paper 1984.

Stiglitz, J., "Information, Screening and Welfare," *Bayesian Models In Economic Theory*, M. Boyer and R. Kihlstrom, editors, North-Holland, Amsterdam, 1984.

The Effects of the Advertising Substantiation Program on Advertising Agencies*

Keith B. Leffler
Raymond Sauer, Jr.
University of Washington

Introduction

Federal concern about product advertising dates back to the turn of the century. The Federal Trade Commission Act of 1914 included Section 12 which was designed to control the exhorbitant product claims made by the patent medicine sellers of that era. At that time these medicines were the major advertisers in the economy.[1] In 1938 the FTC Act was amended to eliminate ambiguity concerning FTC jurisdiction over deceptive product claims not affecting competition. Since then the Commission is mandated to protect consumers from the wiles of unscrupulous sellers.

The Commission's actions in regulating deceptive advertising over the last 40 some years are extensive and well documented.[2] For the purposes at hand, the significant era at the Commission commenced in the 1960s. In 1963, the Commission held that an advertiser making a claim related to health or safety "implicitly represents that he has a reasonable and substantial foundation in fact for making the claim."[3] For the first time deception could occur even though a seller's claim is true.

Towards the end of the 1960s, the FTC's efforts in regulating advertising became of interest to the consumer movement. During this era, advertising was generally viewed as an evil of capitalism and a means of persuading the ignorant consumer to follow the dictates of Madison Avenue.[4] Rather than protecting consumers, FTC actions were viewed as "obsessed with trivia" by leaders in the consumer movement.[5] In 1969, Ralph Nader's organization and the American Bar Association both filed reports recommending total revamping of the FTC's regulation of advertising.[6]

---

*James Ferguson and Mark Plummer of the FTC have been most helpful in this research.

[1] *See* F. Presbrey, *The History and Development of Advertising* (1929).

[2] *See*, for example, R. Pitofsky, "Beyond Nader: Consumer Protection and the Regulation of Advertising," *Harvard Law Review* 90 (February, 1977) and S. Peltzman, "The Effects of FTC Advertising Regulation," *Journal of Law and Economics* (December, 1981), and the numerous papers cited by these authors.

[3] Heinz W. Kirchner, 53 FTC 1282 (1963), 1275.

[4] Vance Packard, *The Hidden Persuaders*, and J.K. Galbraith, *The Affluent Society*, are typical of prevalent opinions of advertising's role in the economy.

[5] "FTC Drops 70% of Cases," *Advertising Age*, 12/8/69, p. 1.

[6] The Nader Report on the Federal Trade Commission (1969); American Bar Association, Report of the ABA Commission to Study the Federal Trade Commission (1969).

Within this "heated" climate, Casper Weinberger was appointed head of the FTC in late 1969. President Nixon gave Weinberger "carte blanche" to reorganize.[7] Eighteen of thirty-one top level staff members were discharged and the Bureau of Consumer Protection was created replacing the Bureau of Deceptive Practices.[8] Seventy percent of existing cases were dropped in preparation for an overhaul of the regulatory fare.[9]

This activist, revitalized Federal Trade Commission instituted the Advertising Substantiation Program (ASP). In April, 1970 the FTC first signalled an expanded concept of deception that included the lack of substantiation for all product claims.[10] By late 1970, Ralph Nader campaigned in the press for FTC requirements that ad substantiation be available to the public.[11] The McGovern-Moss Bill was introduced in Congress in April, 1971. This bill called for substantiation of ad claims.[12] The clear signal that advertising substantiation was now a primary part of the FTC's arsenal in fighting deceptive sellers' claims came at this time,[13] when a hearing examiner found the Pfizer Company guilty of deception due to the absence of substantiation for their advertised product claims.[14]

Economists have generally been skeptical of the impact of the FTC's regulation of advertising.[15] First of all, under FTC programs deceived consumers receive no redress, that is, there are no damages to the deceived. Traditionally, the end result of a successful FTC complaint has been a consent decree which essentially is a promise by the defendant not to offend again. This obviously limits the incentive for consumers to utilize the regulatory authority of the FTC. Posner has concluded that the FTC's efforts at regulating advertising are without effect.[16] A leading industrial organization text concludes that "the FTC impact on advertising remains slight."[17]

On the other hand, major companies and major advertisers have been subjected to FTC proceedings with non-trivial litigation costs. Sterling Drug, the manufacturer of Bayer products, has undoubtedly spent hundreds of thousands

[7] K. Clarkson and T. Muris, "The FTC Since 1970" (Cambridge University Press, 1981).

[8] *Ibid.*, p. 4.

[9] *Advertising Age*, 12/8/69, p. 1.

[10] *Wall Street Journal*, 4/14/70, p. 4: as reported in R. Higgins and F. McChesney, "Truth and Consequences: The Federal Trade Commission's Ad Substantiation Program," mimeo, undated, p. 35.

[11] *Advertising Age*, 12/14/70, p. 3.

[12] *Advertising Age*, 4/15/71, p. 1.

[13] Four cases had been filed as the substantiation program was being developed. *National Dynamics Corp.* (82 F.T.C. 488), filed 11/21/69; *Pfizer, Inc.* (81 F.T.C. 23), filed 4/6/70; *Firestone Tire* (81 F.T.C. 398), filed 6/29/70; and *Standard Oil of California* (84 F.T.C. 1404), filed 12/29/70. The *Pfizer* case was the first in which a decision was reached.

[14] Discussed in the 1972 decision of the Commission (81 F.T.C. 23), at 25. The examiner's opinion was reported in *Advertising Age*, 6/14/71, p. 1. The hearing examiner's decision was subsequently overturned by the Commission though the substantiation theory was upheld.

[15] *See* S. Peltzman, "The Effects of FTC Advertising Regulation," *Journal of Law and Economics* (December, 1981), p. 447.

[16] R. Posner, *Regulation of Advertising by the FTC*, Americn Enterprise Institute, 1983.

[17] W. Shepherd and C. Wilcox, *Public Policies Towards Business*, Irwin, 1979, p. 539.

of dollars fighting a substantiation complaint filed in February, 1973 but still not decided. The FTC spends significant amounts on ASP, about ¾ million dollars in 1979.[18] Private self-regulatory bodies that have arisen to forestall government controls, have direct enforcement expenditures of over 2-1/2 million dollars annually.[19] The costs of substantiation documentation and compliance with FTC information requests are not known but are undoubtedly significant.[20]

The advertising substantiation program has the potential to impact numerous economic markets. Producers relying on advertising to disseminate information about their products are the most directly affected. Due to the diligence of the FTC, consumers may find advertising to be a more credible, reliable source of product information. As a consequence, producers of "better buys" may increase advertising, market sales and profits. If ASP significantly impacts the value or cost of advertising, the effects will extend to the producers and disseminators of the ads, namely ad agencies and the media (TV, radio, magazines, newspapers and billboards).

To date, little is known about the effects of the FTC's advertising regulation. However, the limited empirical results suggest a potentially large effect. Peltzman studied FTC advertising deception cases from 1960 to 1975.[21] As we discuss further below, the issues involved in deception and substantiation cases are very similar so that Peltzman's findings should be indicative of the substantiation program's impact. Peltzman found that a deception complaint lowered the stock value of the involved producer by over 3 percent. Quoting Peltzman, "the size of the effect is...astounding,"[22] since the companies involved are usually quite diversified with the complaint involving a small percent of products sold. The results become even more mysterious in light of Peltzman's general failure to show significant effects on product sales as a consequence of FTC action.[23]

Higgins and McChesney have studied the capital market effects on both advertisers and advertising intermediaries from the substantiation program. They formed portfolios of both large advertisers and large ad agencies. The advertiser portfolio, 54 firms with the largest advertising expenditure in 1970 in drugs,

[18] C. Guerard and J. Niemasik, "Evolution and Evaluation of the Ad Substantiation Program Since 1971," mimeo, FTC 12/1/78, p. 35.

[19] Priscilla LaBarbera, "Analyzing and Advancing the State of the Art of Advertising Self Regulation," *Journal of Advertising*, V. 9, N. 4, 1980, p. 32, reports that the Council of Better Business Bureaus budgeted $1,475,000 for advertising self regulation in 1979 and the National Association of Broadcasters budget was $950,000. Self regulation activity is discussed in Section 1 below.

[20] On March 3, 1983, the FTC requested comments on their substantiation program. The FTC included request for data on costs. The respondents to this question (Whirlpool, Procter and Gamble, General Mills, Inc., and Sears, Roebuck) all responded that the program did not not impose significant additional costs of substantiation.

[21] S. Peltzman, "The Effects of FTC Advertising Regulation," *Journal of Law and Economics*, (December, 1981), pp. 401-448.

[22] *Op. cit.*, p. 418.

[23] Peltzman found no effect from complaints against Crest and Colgate toothpastes; little effect from complaints against Blue Bonnet margarine and Alcoa Wrap, Standard Oil and Sun Oil, some effects of questionable importance for two unidentified food items.

toiletries, home products and building products submarkets, had 27 percent excess returns over the period April, 1970 to July, 1972.[24] If this was a result of ASP, the order of magnitude is also "astounding." Direct sales effects should be readily observed.

Higgins and McChesney also examined the stock market value of four, large publicly traded advertising agencies.[25] They report that from April, 1970 to June, 1971 these companies earned 91 percent excess market returns which they attribute to the implementation of the advertising substantiation program.[26] This finding is also striking in magnitude. It suggests either a very substantial increase in the value of advertising due to ASP or possibly, as Higgins and McChesney argue, a regulatory bias towards the larger agencies. Either way, the source of the increased value should be readily observed in disaggregated advertising data.

In this paper we pursue the effects of ASP on the advertising agencies. The advertising agencies have been vocal opponents to former FTC Chairman James C. Miller's suggestion to reevaluate and perhaps weaken the ASP Program.[27] This supports the Higgins-McChesney finding of beneficial effects for these firms. Here we wish to examine more detailed data to see if we can uncover the source and distribution of any such benefits to advertising agencies.

In the first section we review the temporal pattern of FTC activity in regulating advertising since the mid-1960s. Our ultimate objective is to develop variables proxying the level of FTC substantiation scrutiny over advertising content. In order to have some control for the extent of "suspect" advertising, we also discuss the activity of the private self regulatory National Advertising Review Board of the Council of the Better Business Bureau. Comparing private and FTC case activity, we find ASP to be a program of the mid-70s. Advertising substantiation appears to have been substantially deemphasized independently of the Reagan retrenchment.

In the second section, we briefly discuss the many ways that FTC regulation of advertising might impact the advertising agencies. This section also details the structure of the advertising industry and relates the concentration of U.S. agencies to international concentration trends and FTC activity.

The aggregated structural data shows a possible differential favoring the mid-

---

[24] Higgins and McChesney, "An Economic Analysis of the FTC's Ad Substantiation Program," this volume.

[25] Higgins and McChesney, "Truth or Consequences: The Federal Trade Commission's Ad Substantiation Program," mimeo, undated.

[26] These were the only four publicly traded agencies. They were J. Walter Thompson; McCann-Erickson; Foote, Cone and Belding Communications; and Wells, Rich and Greene. They ranked 1, 7, 11 and 19 in 1971 U.S. billings according to *Advertising Age*, 2/21/72.

[27] *Advertising Age*, "Miller Questions Ad Proof Policy," 11/2/81; Stanley Cohen, "A Healthy Skepticism About Public Policy," 11/21/81, p. 48; Gordon, "Miller Asks for Ad Rule Review," 10/25/82, p. 1; Larry Edwards and Todd Findell, "FTC's Miller: Ad Proof Still Concerns Me," 10/16/81, p. 3. Cohen's biweekly column "Washington Beat" makes frequent unreferenced reports of industry support for ASP.

sized agencies, those with ranks below the top 8. Any such effect is, however, small. No impact on the relative growth of the leading eight firms can be discerned. In addition, any impact of ASP on advertising levels and agency use seems to be negative. While major limitations in the data make any conclusions most tentative, the failure to find very significant beneficial impacts is inconsistent with the Higgins-McChesney attribution of 91 percent excess returns for large agencies to the advertising substantiation program.

The third section of the paper studies individual agency accounts. An empirical model of account retention is developed to test for the impact of ASP.

The concluding section discusses further research on the impact in the advertiser or product markets and the media markets that is needed. Only by assembling the effects in all of the advertising "sub-markets" can a complete, consistent picture of the effects of ASP be assembled.

I. FTC and Private Advertising Substantiation Activity

With its seed from the consumerist movement, the advertising substantiation program's major role was conceived as a consumer education device.[28] Industries were selected for substantiation requests. The responses were to be publicly available and to be used for possible follow-up complaints. Table A lists the industries subject to these substantiation "rounds" along with the number of complaints resulting from staff investigation of the substantiation responses. The peak of this program was obviously in 1971-1972 when 11 industries and 148 companies were scrutinized. The far right column of Table A simply provides an indicator of the importance of round activity. The number of rounds and generated complaints are simply added.

The industry "round" approach was not particularly successful. "The public education goal was largely frustrated because the substantiation was too technical...(and)...deterrence would not occur simply by placing the documents on the public record."[29] After internal evaluation of the industry rounds in the early 1970s, the emphasis of ASP shifted from education to law enforcement. Industry rounds were still utilized as a source for cases, but Table A shows the decreasing importance of these rounds.

Subsequently, ASP has focused on individual products. The source of particular complaints is unknown: presumably the staff reacts to consumer complaints. In addition, the FTC staff generates investigations from their "weekly ad monitoring meetings."[30] Table B summarizes the activity of the FTC in advertising substantiation cases. The cases are divided into those involving national

---

[28] *See* C. Guerard and J. Niemasik, "Evolution and Evaluation of the Ad Substantiation Program Since 1971," FTC memo to the Commission, mimeo, 12/1/78, for an "inside" discussion of ASP.

[29] *Op.cit.*, p. 23.

[30] *Op.cit.*, p. 25, report that "(From January to December, 1978), the Division of Advertising Practices has conducted nearly fifty preliminary investigations of claims which surfaced in the weekly ad monitoring meetings."

Table A

FTC ADVERTISING
SUBSTANTIATION INDUSTRY ROUNDS

Year	Industry	Number of Companies	Number of Complaints	Industry Round Activity Variable*
1971	Automobiles	7	0	
1971	Air Conditioners	17	2	
1971	Electric Shavers	4	0	7
1971	Televisions	12	0	
1971	Dentifrices	8	0	
1972	Cold/Cough Remedies	15	0	
1972	Tires	18	1	
1972	Soaps	35	0**	14
1972	Hearing Aids	11	6	
1972	Pet Foods	12	1	
1972	Automobiles	9	0	
1973	Antiperspirants	10	0	
1973	Shampoos	11	0	6
1973	Acne Preparations	14	3	
1974	Auto Tires	13	3	
1974	Color TVs	6	2	
1974	Dental Products	4	1	
1975	Dishwashers	5	1	3
1975	Lawnmowers	5	0	
1976	Preference Polls	4	2	3
1977	Energy	7	0	1

SOURCE: Appendix A, Guerard and Niemasik (1978)
*Sum the number of industries investigated and the number of complaints thereby issued.
**Detergent Rule was established governing ecology claims.

Table B

FTC AND NAD CASE ACTIVITY IN ADVERTISING SUBSTANTIATION ISSUES, 1968 — 1982

	Federal Trade Commission					NAD/NARD
	Substantiation Cases National Adv.*				Deception Cases	Cases
Year	Agency Named	Agency Not Named	Local Adver.**	TOTAL***	TOTAL	
1968	--	--	--	--	1	NA
1969	--	--	1	1	1	NA
1970	1	2	--	3	6	NA
1971	--	2	--	2	11	NA
1972	3	2	1	6	18	****
1973	3	2	5	10	3	138
1974	3	9	8	20	7	132
1975	5	5	4	14	6	160
1976	5	4	2	13	3	168
1977	2	2	3	7	--	151
1978	2	2	2	6	--	164
1979	2	2	1	5	1	163
1980	1	4	2	7	2	129
1981	2	1	2	5	--	143
1982	--	--	2	2	--	*****
Total	29	37	33	101	59	1348

NA = not applicable, program not operative.

*When multiple cases are filed for substantially the same allegation only 1 case is recorded. This includes situations where industry rounds result in numerous complaints (e.g., hearing aids), situations where both a retailer and a manufacturer are named (Payless and Lear Siegler), and situations where separate complaints are filed for the advertising agency.

**Includes both door-to-door and local newspaper advertising.

***Includes only cases involving substantiation type claims. Examples of excluded practices include "bait-n-switch," fictitious sale prices, and labelling issues. Included cases are exemplified by Ocean Spray Cranberries' nutritional claims, Profile Bread's diet claims, and Firestone Tire's safety claims. The inclusion in the substantiation rather than deception column is based on the FTC's list of substantiation cases used by Higgins and McChesney (1984).

****Reporting methodology not comparable to subsequent years.

*****Not available yet.

advertising, generally television or magazine ads, and those local in scope, typically newspaper ads. The national cases are further broken down into those where the advertising agency for the product is named either within the complaint or in a separate complaint, and those cases where the agency is not named.

The list of substantiation cases was obtained from the Federal Trade Commission.[31] In reviewing these and other FTC advertising activity, we were struck by the similarity of many "deception" cases and the substantiation cases. For example, claims about Ocean Spray Cranberries' nutritional value, Vivarin's potency effects, and Sunoco gasoline's performance characteristics lead to deception complaints. However, claims about Gainesburgers' nutritional value, Rev-Up Vitamin's health effects, and Standard Oil gasoline's pollution characteristics lead to substantiation complaints. We, therefore, have surveyed all FTC advertising cases and consent decrees since 1968 and found 65 "deception" cases we judged to involve issues basically similar to substantiation questions. The annual number of such cases are also given in Table B. It is noteworthy that the pattern of these deception cases over time is basically the same as that of the substantiation cases. The "hey day" of deception activity was from 1971-1975. This suggests that the commonly held view that ASP substitutes for deception activity (with a shifted burden of proof) is not correct. Apparently a "consumerist's" concern with advertising leads to activity of either type.

Measured by filed complaints, the substantiation program peaked during the years 1974-1976 when half of the cases filed since 1969 were issued. By 1982, case activity has nearly ceased with no cases of national scope being filed. The decline in FTC involvement could be the result either of policy changes within the agency or of the program's success in eliminating unsubstantiated advertising claims.

One way to distinguish between these two possible explanations for the change in FTC advertising case activity is to find some independent measure of the prevalence of unsubstantiated advertising. In 1971, the National Advertising Review Board was set up by the Council of the Better Business Bureaus, the American Advertising Federation, the Association of National Advertisers, and the American Association of Advertising Agencies as a private self-regulatory body to scrutinize advertising content for product claims.[32] The National Advertising Division (NAD) of the NARB handles most of the complaints about advertising. If this staff cannot resolve the complaint with the advertisers, the NARB itself steps in. This procedure occurs in less than 4 percent of the cases.

NAD cases originate mainly from staff monitoring (44 percent) and competitor complaints (30 percent).[33] Table B gives the number of cases handled by the

[31] The list of complaint and consent decrees were obtained from Richard Higgins of the FTC as assembled by Susan Campbell of his staff.

[32] *See* Gary Armstrong and Julie Ozanne, "An Evaluation of NAD/NARB Purpose and Performance," *Journal of Advertising*, V. 12, n. 3 (1983), pp. 15-26.

[33] Complaint source percentages are from Armstrong and Ozanne, *op. cit.*, p. 19, for 1979-1981.

NAD from 1973 through 1981.[34] The pattern of the cases is obviously very different from FTC complaint activity. These case numbers are not directly comparable to FTC complaints. About 52 percent of NAD cases involve ad claims substantiated by the advertiser. Another 17 percent involve ad claims dropped prior to a hearing.[35] Presumably these cases would usually not reach the complaint stage in an FTC investigation. However, the number of NAD cases where ads were held unsubstantiated remained relatively constant over time. From 1973-1975, the NAD ruled 132 advertisements unacceptable; from 1976-1978, 143 advertisements were ruled unacceptable; and from 1979-1981, 138 ads were found unacceptable.[36]

From a review of FTC advertising complaint activity as compared to NAD case activity, we conclude that FTC regulation of advertising has changed markedly over the period 1965-1982. The activity is well proxied by the pattern of round activity and complaint activity given in Tables A and B. Activity peaked in the mid-1970s. The Reagan Administration has continued the deemphasis of advertising regulation that was discernible by 1977.

II. The Advertising Substantiation Program and the Structure of the Advertising Agency Industry

Federal regulation of advertising can impact advertising agencies in numerous ways. The demand for advertising and for advertising intermediaries' services can increase in general. Alternatively, the increased federal scrutiny of ad content can increase the use of intermediaries who specialize in regulatory compliance, even if the overall demand for advertising falls. In addition, the regulatory activity can cause shifts in demand among the agencies. The regulators may focus efforts on the large, visible agencies relative to the smaller agencies or a program such as ASP may cause the advertiser to favor larger, experienced agencies. Another possibility is that the regulatory program can increase agencies' efficient scale due to a fixed cost of the regulatory compliance division. Finally, the entry cost into the agency business may rise due to the increased value of reputation of having dealt with authorities.

The actual effects on the advertising industry from the introduction of and variance in enforcement of substantiation requirements are not, in our view, determinable from a priori theorizing. Rather, descriptive, empirical evidence is required on the effects in the advertiser, advertising and advertising intermediary sectors of the advertisinq market. Here we provide only a limited set of the evidence necessary to understand the impact of ASP on ad agencies. In the conclusion, we suggest the additional evidence from the advertiser and

[34] The case count is from annual *NARD Activity Reports* listing all cases and their disposition.

[35] Armstrong and Ozanne, *op. cit.*, Table 8, p. 24.

[36] These numbers are calculated from Table B in conjunction with case disposition percentages in Armstrong and Ozanne, *op. cit.*, Table 8, p. 24. An unacceptable ad includes either a finding of deception or lack of substantiation.

advertising sectors that would substantiate our preliminary conclusions reached here.

The general structure and size of the advertising agency industry is given in Table C. The first set of numbers, columns II and III, gives the concentration ratios for the largest eight agencies, CR8, and for the next largest twelve, CR9-20 for years 1965-1981.[37] While no strong pattern is present in either set of concentration ratios, there is a tendency towards decreasing eight-firm concentration from 1965 to 1973 and increasing eight-firm concentration from 1974-1981. The nine to twenty firm concentration shows a minor increased concentration after 1972.

The next set of numbers in Table C, columns IV and V, gives a general indication of the size of the advertising industry. The fourth column gives the total advertising expenditures on television, radio, magazine and local newspaper ads for the years 1965-1981, while the fifth column presents the total U.S. billings of the largest 50 agencies for the same years.[38] Advertising expenditures appear to increase in 1972 and then accelerate in 1976, while real billings generally fall until 1976. Neither the structural or the size data indicate obvious impacts of ASP. We now shall attempt to relate ASP to these data in more detail.

The structure of the advertising agency industry is presumably affected by numerous factors including technology, industry growth, luck, entrepreneurial skills and regulatory activities other than FTC advertising policies (such as antitrust). To control for technological factors, we have calculated three alternative variables measuring the structure of non-U.S. advertising agency billings. These are a 25-firm Herfindahl index in Great Britain GHHI, a 20-firm index for Japan JHHI, and concentration measures of U.S. agencies' foreign billings, FCR.[39] The foreign agency Herfindahls can measure only relative shifts among the very large firms since data on middle and small foreign agencies is unavailable.

[37] These concentration ratios refer to the percentage of domestic billings of the top 50 firms accounted for by the top 8 and top 9 to 20 firms as calculated from "Estimated Billings," as reported in *Advertising Age*, annually, various weekly issues. The standard concentration ratios would use agency revenue or gross income rather than billings and take the percentage of all agencies rather than the top 50. The income estimates are unavailable prior to 1972. Domestic billings are broken out from international billings only for agencies billing more than 10 million dollars. In addition, the number of smaller agencies responding to the *Advertising Age* survey varies by 10-15 percent per year. We checked the percent of domestic billings accounted for by the largest 50 firms, assuming all billings less than 10 million are domestic, and ignoring non-reporting agencies. This percent shows the same pattern as the CR9-20 numbers. If the implied CR50 estimate is valid, our conclusions about CR9-20 can be extended to firms in the 9-50 rank size.

[38] The real advertising expenditures are the sum of television, local newspaper, radio and magazine advertising expenditures as estimated by McCann-Erickson. *See Advertising Age* 1/5/78 for an example. The nominal estimates are deflated by the GNP deflator, 1972 = 100. These expenditures exclude newspaper supplement and billboard advertising among missing major categories. The billings of the top 50 agencies are also deflated by the GNP deflator.

[39] *Advertising Age* reports foreign agency billings annually, the list of reported agency changes over time. We were able to form a reliable series of only 20 Japanese firms and 25 British firms. When a large agency did not report billings for any year over the 1965-1981 period they were excluded. There were two such firms for Japan and three for Britain.

Table C

THE STRUCTURE AND SIZE OF
THE ADVERTISING AGENCY INDUSTRY*

	Industry Structure		Industry Size (000)	
Year	8 Firm Concentration	9-20 Firm Concentration	Real Advertising Expenditures	Real Agency Billings
	Percent		Dollars	
I	II	III	IV	V
1965	42.1	30.0	106,671	587.7
1966	42.5	28.5	114,169	591.8
1967	42.4	29.6	113,833	618.7
1968	41.1	29.7	118,759	619.6
1969	41.4	29.4	124,027	629.7
1970	42.7	30.5	119,016	591.3
1971	40.5	31.5	119,751	571.0
1972	40.1	34.8	130,480	586.5
1973	40.0	33.0	133,538	578.5
1974	41.5	33.0	128,883	550.5
1975	40.6	31.5	127,127	548.0
1976	41.7	32.0	146,150	617.5
1977	42.9	29.5	156,604	655.0
1978	40.2	33.3	169,484	748.9
1979	42.5	32.4	175,317	828.3
1980	43.0	32.5	175,447	878.8
1981	42.9	32.3	180,762	926.7

* *See* notes 37 and 38 to the text for an explanation of these data.

We expect the eight firm concentration ratio of U.S. billings to be positively related to the foreign Herfindahl measures.[40] The 9 to 20 firm concentration ratio should be negatively affected since an increase in the foreign Herfindahls indicate shifts in billings out of the smaller agencies in the top 20 or 25. The 20 firm U.S. concentration should not be related to the foreign 20 or 25 firm Herfindahl. It is more difficult to predict how U.S. firms' foreign billings concentration should impact domestic concentration. On the one hand, increased foreign billings may substitute for domestic due to resource constraints.[41] On the other hand, foreign billings concentration changes may measure technological factors also affecting U.S. concentration. A final problem in the U.S. agencies foreign billings variable is that ASP may affect both domestic and foreign practices of the U.S.-based agencies. We have not attempted to account for entrepreneur-specific events influencing agency concentration.

We also are interested in the possible effect of ASP on U.S. concentration. Two alternative measures of ASP activity were used. The first, ASPD, is simply a dummy variable with a value of 0 prior to 1971 and value of 1 thereafter. The second, ASPA, is a measure of enforcement activity defined as the sum of case activity in Table B and round activity in Table A.[42]

Table D reports the results of representative least squares regressions of various agency concentration measures.[43] Keeping in mind the limited number of observations, it appears that the substantiation program did not affect the eight-firm concentration level. The measures of ASP always negatively affect the eight-firm concentration but the effect is never significant. The second set of regressions in Table D.I. have the 9 to 20th rank agency concentration as the dependent variable. All coefficient signs are sensible and ASP has a significant positive effect on these "next-to-the-largest" set of firms.[44] We conjecture that perhaps ASP leads to general increased agency economies of scale along with increased scrutiny of the largest most visible agencies.

Section II of Table D reports the results of agency billings regressions. Here we included a time trend, the level of real expenditures from Table C and the alternative ASP variables. We find that ASP negatively impacted the largest eight agencies and that the impact is statistically significant. The second set of regressions is for the billings of the ninth to twentieth ranked agencies. We do not find a significant impact of ASP. In addition, while the estimated coefficient on ASP is negative, it is, relative to billings, far smaller than the coefficient for

[40] The foreign concentration and the foreign agency Herfindahls were found to be unrelated to any measure of ASP.

[41] *Advertising Age* reports numerous stories where agencies do not bid accounts due to worries of spreading their talent too thin.

[42] We hypothesized the industry rounds generating complaints might be viewed as particularly tight enforcement. Hence the double counting of round generated cases.

[43] Time is included to account for any trend effects in domestic concentration not taken account of by the foreign structure variables.

[44] The firms ranked 9 to 20 have billings about 75 percent of the top 8 billings.

Table D

I. REGRESSIONS OF AGENCY CONCENTRATION

Dependent Variable	Time	Billings	FCR	JHHI	ASPA	ASPD	Cons.	R^2
	Independent Variable Coefficients and (t-statistics)							
CR-8	-.180	.006	-.138	.002	-.004		56.3	.350
	(-1.1)	(1.4)	(-1.4)	(.6)	(-.07)			
	.01	.004	-.119	.003		-1.33	40.4	.437
	(.06)	(1.2)	(-1.3)	(1.2)		(-1.3)		
			-.119**	.003	-.007		43.6	.251
			(-2.2)	(1.3)	(-1.4)			
CR-20	.075	.006	-.078	.004	.146**		30.0	.665
	(.3)	(.9)	(-.6)	(-1.13)	(2.2)			
	-.028	.001	-.146	-.007**		3.17**	48.0	.693
	(-.1)	(.3)	(-1.4)	(-2.1)		(2.5)		
			-.145**	-.006**		3.06**	46.5	.737
			(-2.5)	(-2.85)		(4.6)		

II. REGRESSIONS OF AGENCY BILLINGS

Dependent Variable	Time	Advertising Expenditures	ASPA	ASPD	Cons.	R^2
	Independent Variable Coefficients and (t-statistics)					
Billings of	-3.40	2.47**	-2.01		196.7	.841
the Top 8	(-.7)	(2.46)	(-2.0)			
Agencies						
	-3.54	3.29**		-52.5**	110.3	.863
	(-.9)	(5.3)		(-2.6)		
Billings of	-3.80	2.30**	-.240		168.3	.858
the 9-20th	(-1.1)	(3.3)	(-.3)			
Top	-2.25	2.27**		-17.5	68.6	.870
Agencies	(-.7)	(4.7)		(-1.1)		

**Significant at the .05 level.

the leading eight firms.[45] These results suggest that ASP generally adversely impacted advertising agencies but impacted the biggest agencies more.

It is difficult to reconcile these results with the Higgins-McChesney substantial capital value increase findings mentioned above. One possibility is that ASP affected different agencies in very different ways and that their four agency sample happened to be a favored bunch. We note, however, that the share of the billings of the top 20 accounted for by the four publicly traded agencies declined from .235 in 1971 to .173 by 1974. This mainly resulted from the poor performance of McCann-Erickson. We simply do not find any evidence in the aggregate data that suggests very substantial benefits from ASP to large ad agencies as argued by Higgins and McChesney. In the next section, we examine the pattern of individual account activity to see if impacts of ASP on various agency categories are therein manifest.

III. The Effects of ASP on Agency Accounts

The analysis of Section II indicated that different size agencies were likely impacted differently by the advertising substantiation program. Differential impacts can arise because of FTC selectivity in enforcement and/or economies of scale impacts favoring particular sized agencies. In investigating specific accounts of the agencies, we are interested in determining whether the ASP program and ASP complaints differentially impact agencies' probabilities of gaining or losing an account according to agency size.

Table E.I. presents a list of products involved in substantiation complaints where the advertising agency was named in the complaint. Section II of Table E summarizes the size classes of named agencies. The data indicate that the largest and the smallest agencies are disproportionately represented. We hypothesize that the smallest agencies, with less reputational capital at stake, are more willing to take a chance with unsubstantiated claims. The disproportionate naming of larger agencies however is more likely due to their visibility and size itself.

It is not obvious that being named in a substantiation complaint adversely impacts an advertising agency's industry position. For the 15 agencies in Table E with industry rank data available before and after a complaint, 8 improved their industry position one year after the complaint as compared to one year before.[46] However 10 of the 15 agencies did relatively poorer in the two-years interval surrounding the complaint than in the two-year period three years to one year before the complaint. In addition, 8 of 11 agencies for which data is available, did relatively better in the two-year period one to three years after the complaint than during the complaint period. Since there is less than a 1 percent prob-

[45] The estimated negative effects on the top 8 billings from ASP are from 10 times (ASPA) to 3 times larger (ASPD) than those on the top 9-20.

[46] About half the agencies improved their rank and half regressed regardless of which before and after years are compared in the interval three years before and three years after the complaint.

Table E

I. Advertising Agencies Named in FTC Substantiation Cases

Year	Product	Agency	Size*
1970	Gas—F310	Batten, Barton, Durstine, Osborn	4
1972	Sugar (Hawaiian)	Leo Burnett**	3
1972	Lysol	Sullivan, Stauffer, Colwell, Bayles	17
1972	Vega/Opel	McCann Erickson***	7
1973	Vanquish	Benton and Bowles**	16
1973	Volvo	Scali, McCabe	78
1973	Bayer	Dancer-Fitzgerald	12
1973	Anacin	C.T. Clyne Maxon	50
1974	Whirlpool AC	Doyly, Dane, Bernbach**	7
1974	Ford	J. Walter Thompson**	2
1974	Eggs	Richard Weiner	NR
1975	Rev.Up (Vitamins)	Levine, Huntley, Schmidt	184
1975	Yamaha	Batsford, Kelchin	NR
1975	X-11 (Diet Pill)	Kelly, Ketting, Furth	174
1975	Poli-Grip	Grey Advertising	5
1975	Bridgestone Tires	Parker Advertising**	49
1976	STP Additives	Stern, Walters, Simmons	123
1976	Perma-Strate	Merrill, Kremer	NR
1976	Mr. Cool	Franklin Lett	NR
1976	Ultra Sheen	Bozell and Jacobs	23
1976	Adolph's Salt	Nagle, Spillman and Bergman**	NR
1976	Vital-E Diet	Advertising Unlimited	467
1977	Hawaiian Sugar	Foote, Cone and Belding	7
1977	Kenmore D/W	J. Walter Thompson	2
1978	Womack MO	S.W. Advertising	NR
1978	Acne-Statin	National Media Group	NR
1979	G.R. Valve	Ad Marketing	101
1979	Water Pik	J. Walter Thompson	2
1980	Encare O/C	Shaller Rubin	116
1980	Semicid O/C	Sorga	NR
1981	Fleischmann's	Ted Bates	6
1981	Blackman's Shaver	DKG Advertising	47

*Rank by U.S. annual billings one year before the complaint.
**Separate complaint in agency's name issued.
***Complaint dismissed.
NR = Not reported in *Advertising Age* or in the *Standard Directory of Advertising Agencies*.
SOURCE: *Advertising Age*, "Agency Billings," various issues.

II. Size Classes of Named Agencies

	Size Classes			
	1-8	9-20	21-100	-100**
Number of Complaints	9	3	4	13
% of Complaints	28.1	9.4	12.5	40.6
% of Billings*	26.7	22.5	29.0	21.8

*Estimated for 1972 Agency Billings, *Advertising Age*, 2/26/73.
**Agencies that are not reported in *Advertising Age* are assumed to be small agencies with ranks above 100.

ability that this pattern is random, it appears that being named in a complaint does have adverse consequences.[47]

The impact of ASP on agencies is not likely to be limited to those accounts named in complaints or to those agencies named. If advertisers anticipate FTC actions are more likely directed towards particular agency types or if certain agency types are less expert in avoiding FTC sanctions, the advertisers will presumably be less likely to use these agencies.

In order to investigate this possible effect of ASP, we studied a sample of 200 advertising accounts over the 1967 to 1976 period. These included the 36 products named in ASP complaints prior to 1976 and 164 other products selected to ensure a reasonable number of accounts for agencies of different size classes. The final sample was reduced to 136 accounts because information on agency size and/or account size was unavailable.[48] Through consultation with the *Standard Directory of Advertisers*, we noted whether, when and to whom these accounts switched agencies, 1967-1976.

Table F summarizes the agency account changes that occurred between agencies of various sizes before and after ASP. Out of 1,360 possible agency changes, there were 68 switches, 28 before ASP, 40 after. There is a 25 percent probability these numbers could be generated by the same binominal process. By inspecting the cells of Table F, we can see no obvious inter size changes before versus after ASP. Within each cell, the upper left number is the before ASP switches in accounts and in the lower right is the post ASP number. It appears that there are no major differences other than the top 8 firms losing more accounts to firms of size rank *below* 50 after ASP.

A more detailed regression analysis confirmed the absence of any significant effect of ASP on account retention. A logit regression of the following general form was estimated.

$$\begin{aligned} PLA_{tKJ} &= \alpha_0 + \alpha_1 ASP_t + \alpha_2 S_K \cdot ASP_t + \alpha_3 S_K + \alpha_4 C_{tJ} \\ &\quad + \alpha_5 R_{tJ} + \alpha_6 E_{tJ} \end{aligned}$$

where PLA_{tKJ} is the probability of agency K losing account J in year t, a 0,1 dummy variable with a value of 1 in a year where the account was lost; ASP_t is a measure of the advertising substantiation program, ASPA or ASPD; S_K is the size rank of agency K; C_{tJ} is a dummy variable for whether account J was subject to an ASP complaint; R_{tJ} is a dummy variable for whether account J was subject to an industry round; E_t is the size of advertising account J.

[47] If we assume that it is equally likely that an agency does relatively better or poorer during the complaint interval as before or after, there is only a .0062 probability that 18 of the 26 observations would show the agency doing poorer during the complaint interval.

[48] The final sample included for 1972 agency sizes, 47 accounts of the top 8, 41 accounts of the ninth to twentieth agencies, 24 accounts of agencies ranked from 21 to 50, and 24 accounts of agencies smaller than rank 50. Twenty-three accounts that were named in complaints were included in the 136 account sample.

Table F

SUMMARY OF ACCOUNT CHANGES, VARIOUS AGENCY SIZE CLASSES, BEFORE AND AFTER ASP

Size Class	Top 8 Before	Top 8 After	9-20 Largest Before	9-20 Largest After	21-50 Largest Before	21-50 Largest After	Smaller than the 50 Largest Before	Smaller than the 50 Largest After	Total Before	Total After
Top 8	4	1	4	3	0	0	1	5	9	9
9-20	3	3	3	5	1	2	2	2	9	12
21-50	0	0	0	1	1	2	1	4	2	7
Smaller Than 50	0	1	2	2	0	1	6	8	8	10
TOTAL	7	5	9	11	2	5	10	19	28	40

Table G reports the results of maximum likelihood (logit) estimates of this regression equation. The reported equations use dummy variables for the listed size classes. While we find ASP to have reduced account stability, the effect is not significant. None of the pre or post ASP size variables is found to impact significant upon account stability. Generally, the regression analysis is very unsuccessful in explaining the account changes and neither the ASP program ASPA, a specific account complaint C_J, industry round R_J, agency or account size is found to affect an agency's account retention. Various alternative variable definitions, inclusions and exclusions and functional forms confirm this conclusion.

Table G

LOGIT REGRESSIONS OF AGENCY ACCOUNT LOSSES

INDEPENDENT VARIABLE COEFFICIENTS AND COEFFICIENT/STANDARD ERROR

	Agency Size Ranks									
	Top 8		9-20		21-50					
ASPA	Pre-ASP	Post ASP	Pre-ASP	Post ASP	Pre-ASP	Post ASP	Complaint	Round	Account Size	Cons.
.04	-.07	.006	-.01	.02	-.6	.03	.02		.005	-3.5
(1.4)	(-.2)	(.3)	(-.04)	(.08)	(-1.4)	(1.2)	(.04)		(1.2)	(-4.6)
.04	-.08	.006	(-.01)	.02	-.6	.03		.18	.005	-3.3
(1.4)	(-.3)	(.3)	(-.04)	(.9)	(1.4)	(1.3)		(.8)	(1.2)	(-5.8)
.014									-.004	-3.0
(1.0)									(1.1)	(-12.9)

Dependent Variable: 0 if account is retained.
1 if account is lost.
Number of observations: 1,360.

Concluding Remarks

We have found that the advertising substantiation program was followed by a slight increase in concentration of second level, size rank 9-20, advertising agencies' billings. At the same time, the biggest 8 firms showed a very modest fall in concentration. We also find FTC activity to be directed somewhat disproportionately towards the biggest and the smallest firms. However, try though we may, we cannot find ASP activity to significantly impact account retention or account switching.

It would seem that any effects of ASP on advertising agencies are very modest. This is a puzzling conclusion if the Higgins-McChesney finding of 91 percent excess returns to publicly traded advertising agencies from April, 1970 — June, 1971 is properly attributed to the FTC's advertisinq substantiation program. We cannot reconcile these findings and rather shall conclude by describing ongoing research, the results of which may provide a more complete picture of ASP's effects.

The most obvious explanation for the Higgins-McChesney finding is that it is not a distributional result but rather an industry-wide increase in agencies' wealth. We are currently examining this in two ways. First, by using over-the-counter stock market prices, we have expanded the Higgins-McChesney portfolio to include middle and smaller advertising agencies. We shall test for any sensitivity of the excess returns to firm size and also confirm (or refute) the general finding of excess returns.

The other, more difficult research requires a study of the impact of ASP on the demand for advertising. This we are doing by (1) examining advertising-sales functions for over-the-counter drug markets, (2) expanding the Higgins-McChesney portfolio of advertiser stocks and studying complaint and industry round impacts, (3) examining excess returns patterns in media stock prices, and (4) estimating demand equations for magazine advertisements.

We believe that only by assembling the results of all these studies will a conclusive answer to the question of the benefits and costs, both private and public, from the advertising substantiation program be available.

An Economic Analysis of the FTC's Ad Substantiation Program

Richard S. Higgins and *Fred S. McChesney**
Federal Trade Commission — Emory University

I. Introduction

As the economic theory of regulation has developed, our understanding of the distribution of gains and losses from government intervention in the market place has advanced considerably.[1] This article draws from the economic theory of regulation to examine preliminarily some of the allocative and distributive consequences of the Federal Trade Commission's (FTC) "advertising substantiation" doctrine. Ad substantiation has sharply altered the focus of advertising regulation at the FTC. The FTC had traditionally been required to prove after the fact that advertising claims were false. With the coming of substantiation in the early 1970s, advertisers henceforth were required to possess a "reasonable basis" for advertising claims prior to making them.[2] The FTC now enjoins advertising claims it finds exaggerated in the light of the evidence available at the time the claims were made.

Section II of this paper explains the origins and workings of the ad substantiation doctrine in greater detail. In Section III we present a simple model of advertising regulation by the FTC, and we show how the ad substantiation program in particular may have enhanced the FTC's ability to transfer wealth from small to large firms. Section III also derives some testable implications, one of which distinguishes empirically between public interest (wealth enhancing) and merely redistributional applications of ad substantiation. Section IV tests these implications using data from the stock market and on FTC case incidence. Section V summarizes our findings and suggests some policy options.

---

* The views expressed here are the authors'. They doubtless do not reflect the views of some Commissioners, and do not necessarily reflect the views of the Commission itself. We received helpful comments on earlier drafts from Ronald Bond, Gerard Butters, David Haddock, Cotton Lindsay, Robert Mackay, Michael Maloney, the late Steven Marston, Robert McCormick, William Shughart, Robert Tollison and Bruce Yandle. Research and computational assistance from Susan Campbell, Kim Garman and Kathleen McChesney is also very gratefully acknowledged.

[1] Stigler, *The Theory of Economic Regulation*, 2 Bell J. Econ. 3 (1971); Peltzman, *Toward a More General Theory of Regulation*, 19 J.L. & Econ. 211 (1976).

[2] As explained below, ad substantiation thus focuses on the firm's inputs, in the form of surveys, scientific tests or other evidence deemed to constitute a "reasonable basis" *ex ante* for advertising claims. On the differences between regulation of inputs and outputs, *see generally* Wittman, *Prior Regulation versus Post Liability: The Choice between Input and Output Monitoring*, 6 J. Legal Stud. 193 (177). The ad substantiation doctrine complements the more traditional FTC monitoring for falsity *ex post*, which focuses on outputs (*i.e.*, on the actual performance of the product relative to claims made for it). In principle, that is, the FTC may challenge a claim as either unsubstantiated or false; it often alleges both falsity and lack of substantiation.

II. The Advertising Substantiation Doctrine

A. *Origins of Ad Substantiation*[3]

Since its creation in 1914 the Federal Trade Commission has brought cases alleging false and deceptive advertising. There have been two essential elements in deception cases: interpretation of the actual claim or claims made in the advertisements, and determination of the truth or falsity of the claims as interpreted by the Commission. Then and now the federal courts, which review FTC cases on appeal, have generously deferred to the Commission's expertise in interpreting advertising claims and determining their truthfulness.

In April 1970, however, the Commission announced that it would issue a complaint against Pfizer, Inc. for claims about its Unburn sunburn lotion that were allegedly ***unsubstantiated***. As discussed further below, the Commission's complaint was a distinct break with its traditional regulation of advertising.

Other substantiation actions ensued. In December 1970, Ralph Nader's Center for the Study of Responsive Law petitioned the Commission for a trade regulation rule requiring national advertisers routinely to submit to the Commission the documents substantiating advertising claims. The FTC instead adopted a resolution in June 1971 requiring advertisers to submit to the Commission on demand relevant substantiation materials in their possession when they made their claims. The submissions, eventually to be released publicly, would inform consumers of unsubstantiated claims and encourage competitors to challenge them.

From the outset, the substantiation program distinguished factual (objective and empirically verifiable) from non-factual ("puffing") claims, requiring substantiation only for the former. Several industries were selected and required to submit substantiation for all factual claims. These so-called advertising "rounds" continued until 1981, but the results were disappointing. The expected consumer interest and competitor challenges did not materialize. An avalanche of paper descended on the Commission. And, internal sources show, the FTC was unable to evaluate much of the substantiation because it was too technical or complex (which doubtless explains consumers' lack of interest as well).

Nevertheless, the rounds continued, but the Commission also began to issue more complaints under Section 5 alleging lack of substantiation. The Commission modified the program in late 1973, formally making it a part of the FTC's law enforcement mission. Substantiation rounds, while still available for public viewing, were to be used principally in generating cases. As the use of rounds has diminished, Commission staff has relied on advertisers' "voluntary" responses to its requests for substantiation and, where necessary, compulsory process to obtain materials supporting ad claims. In the decade since the Com-

[3] This section is developed from several internal and external sources. *See, e.g.*, Cohen, *The FTC's Advertising Substantiation Program*, J. Marketing, Winter 1980, pp. 26-35.

mission issued its first complaints, the substantiation program has resulted in over a dozen litigated matters and over 100 consents. Indeed, substantiation is now the principal focus of the FTC's advertising regulation.

The elements of the substantiation doctrine remain essentially as they were established by the Commission in July 1972 in *Pfizer*, its first opinion on substantiation.[4] Under Section 5 of the FTC Act it is unfair and deceptive to make an advertising claim for a product without possessing a "reasonable basis" for the claim. The Commission explicitly held that substantiation must exist *prior* to dissemination of the claim: "the fact that [a] test was not conducted *prior* to making the affirmative product claims ... precludes it from being considered as a defense ...."[5] In other words, subsequent tests might establish the truth of the claim, but they would not defeat liability; this has been the informed interpretation of the Commission's opinion.[6]

As a matter of FTC staff practice, the prior-substantiation doctrine was later broadened to include not just the existence of prior substantiation, but advertiser *reliance* on it:

> The requirement of substantiation is a requirement that substantiation *be possessed and relied upon at the time of the first dissemination* of the ad. Ex post facto substantiation is not a legal defense. Documents in the corporate basement or library that were not in the corporate mind when the ad was prepared; opinions of general knowledge that was possessed by individuals within the corporation but not communicated to those actually involved in the preparation of the ad are all insufficient.[7]

Further, the advertising agency hired by an advertiser has almost always been held liable once the client-advertiser's liability has been established, on the theory that the agency is a partner in designing, preparing and running the typical advertising campaign and so shares responsibility for required substantiation. (Curiously, the Commission has never proceeded against a testing agency hired by an advertiser to produce substantiation when that substantiation later proves inadequate, even if the inadequacy results in the advertiser and its ad agency being liable.[8])

Violation of Section 5 through inadequate substantiation subjects the advertiser to a cease-and-desist order. The cease-and-desist order carries no monetary penalty, but advertisers incur the costs of litigation and adverse publicity, and sometimes the cost of interrupting a productive advertising campaign. Violation

[4] *Pfizer, Inc.*, 81 F.T.C. 23 (1972).

[5] 81 F.T.C. at 67 (emphasis in original).

[6] 81 F.T.C. at 63, 67. "[U]nfairness may exist even if the claim is true or if the product performs as advertised." Cohen, *supra* n. 4, at 27.

[7] Address of Richard Herzog (Asst. Director for Advertising Practices) before the Institute of Advanced Advertising Studies, American Assn. of Advertising Agencies, New York Council (June 24-27, 1976), at 19, quoted in Cohen, *supra* n. 4 (emphasis in original).

[8] *National Dynamics Corp.*, 82 F.T.C. 488 (1973), *aff'd* 492 F.2d 1333 (2d Cir.), *cert. den.*, 419 U.S. 993 (1976); *Litton Industries, Inc.*, 97 F.T.C. 1 (1981), *aff'd*, 676 F.2d 364 (9th Cir. 1982).

of an outstanding order, however, is typically punished by civil penalties.[9]

B. *Economic Rationale for Substantiation*

In holding that factual advertising claims must be supported by prior substantiation, the Commission relied on an economic model of the market for information and an assessment of its efficiency:

> Generally, the individual consumer is at a distinct disadvantage compared to the producer or distributor of goods in reaching conclusions concerning the reliability of product claims .... In other cases, the complexity of a consumer product, and accordingly the large amount of detailed product information necessary to an informed decision, makes the costs of obtaining product information prohibitive. This problem is further magnified by the large number of competing products on the market . . . .
>
> The manufacturer has the ability, the know-how, the equipment, the time and the resources to undertake such information by testing or otherwise—the consumer usually does not . . . .
>
> The consumer is entitled, as a matter of marketplace fairness, to rely upon the manufacturer to have a "reasonable basis" for making performance claims. A consumer should not be compelled to enter into an economic gamble to determine whether a product will or will not perform as represented. The economic gamble involved in a consumer's reliance upon affirmative product claims is created by the vendors' activities, and cannot be easily avoided by consumers.[10]

This assessment of the information market remains the rationale for the substantiation doctrine today. Consumers have sub-optimal information; there is no effective way for them to ameliorate this problem; it is more efficient to force producers to substantiate their claims.

In effect, the substantiation requirement has put the FTC in the business of enforcing offers made through the media.[11] The need for third-party enforcement of most (if not all) contracts is not self-evident, however. For third-party

---

[9] Interestingly, more than half the recidivism at the FTC has involved violation of advertising orders. Penalties assessed typically run between $10,000 and $100,000, but may be higher. *See*, Shughart and Tollison, ***Antitrust Recidivism in FTC Data: 1917-1982*** (mimeo, 1982); Altrogge and Shughart, ***The Regressive Nature of Civil Penalties***, 4 International Review of Law and Economics, pp. 55-66 (June 1984).

[10] 81 F.T.C. at 61-62. Although the ***Pfizer*** holding was based on Section 5 unfairness, lack of prior substantiation has also been held deceptive under Section 5. *National Dynamics Corp.*, *supra*, n. 10. The standard for evaluating substantiation is the same under either an unfairness or deception theory. *Crown Central Petroleum Corp.*, 86 F.T.C. 1493, 1548 (1974), *modified*, 530 F.2d 1093 (D.C. Cir.), *reissued*, 88 F.T.C. 210 (1976).

[11] The FTC thus treats an advertising message differently from the common law. At common law, an advertisement usually is not deemed an offer; it is merely an invitation to make an offer. *See* 17 Am. Jur. 2d sec. 334 for numerous cases. The terms of the contract are found in the purchaser's ensuing offer and its acceptance by the seller. Variance of the product or its performance from advertising claims would not be actionable. The FTC breaks with the common law tradition by routinely treating advertising claims as part of the performance promised by the offeror/seller of a unilateral contract, which becomes binding when accepted by the consumer's purchase. *See also* Section 2-313 of the Uniform Commercial Code, which also includes in the contract any statement of fact that is "part of the basis for the bargain."

enforcement of contracts to be efficient, parties must be unable to enforce their own agreements at lower cost. Self-enforcement includes (1) consumers' ability to avoid being deceived, and (2) producers' incentives to avoid deceiving. The lower the costs to consumers of avoiding deception and the lower the incentive to producers to try to deceive in the first place, the more efficient self-enforcement is.

For example, Nelson has shown that the kinds of advertising claims made for different goods, and hence the potential for deception, are a function of the characteristics of the good being promoted.[12] The characteristics of some goods ("search goods") are largely verifiable before purchase, and so are not likely to be advertised deceptively. The characteristics of other goods ("experience goods") are better evaluated after a purchase. Thus, consumer search, while it may not completely alleviate deception in advertising, effectively constrains it.

Likewise, consumers can purchase seller honesty in two ways: paying a higher price in any given transaction, or making repeat purchases.[13] Breach of contract through deception results in a reduction in consumer willingness to pay price premiums or in the flow of repeat business over time to the deceptively-advertising seller. Even if a good is characterized more by "experience" qualities, the likelihood of repeat sales reduces producers' incentives to deceive.[14]

The value of a firm's brand-name capital or reputation provides a third market check on producer incentives to deceive. A reputation established over time indicates that the firm has not disappointed its customers. Thus, venerable firms frequently note their time in business; newcomers rarely point out their recent arrival in the market. The firm's capital also includes the investment value of advertising, which yields a stream of returns over time. The value of the firm's investment in advertising falls if consumers find that the claims are untrue.

Even if the costs and rewards to buyers and sellers are insufficient to maintain honesty in the market, private third-party enforcement of the sales contract may be cheaper than government intervention. Deceptive advertising harms a dishonest firm's competitors, which then have an incentive to run corrective ads of their own or bring legal action against the false advertiser.[15] Consumer demands for information about seller honesty also create profit opportunities for outside suppliers of that information, like *Consumer Reports*, local shopping newsletters, local Better Business Bureaus, and more specialized consumer journals and magazines (*e.g., Stereo Review*).

---

[12] Nelson, *Information and Consumer Behavior*, 78 J. Pol. Econ. 311 (1970); Advertising and Information, 82 J. Pol. Econ. 729 (1974).

[13] Klein & Leffler, *The Role of Market Forces in Assuring Contractual Performance*, 89 J. Pol. Econ. 615 (1981); Telser, *A Theory of Self-Enforcing Agreements*, 53 J. Bus. 27 (1980).

[14] The situation is analogous to a consumer installment contract, where the consumer's performance (payment) follows the seller's. A deceived or defrauded consumer can refuse performance in the event of producer dishonesty, as the seller well knows. *See* R. Posner, *Economic Analysis of Law*, 81 n. 2 (2d ed. 1977).

[15] Jordan and Rubin, *An Economic Analysis of the Law of False Advertising*, 8 J. Legal Stud. 527 (1979).

In short, it is frequently assumed that because of information costs government must regulate advertising.[16] This was the leap of logic made by the FTC in *Pfizer* and subsequently used to justify the Commission's ad substantiation program, despite several self-correcting tendencies evident in the market.

III. An Economic Analysis of Advertising Substantiation

This section hypothesizes that FTC regulation of truth in advertising enables it to redistribute wealth from marginal to inframarginal firms. We show how wealth distribution opportunities are altered when the FTC regulates truth *ex ante* according to the prior-substantiation doctrine, instead of regulating truth *ex post* according to the traditional deception standard. Finally, we derive some implications of the economic theory of regulation. One prediction in particular distinguishes between the public-interest hypothesis and the redistribution hypothesis and suggests an empirical test.

A. *A Simple Model of Ad Substantiation*

The FTC's prior reasonable-basis standard is applied to advertising that implicitly or explicitly contains verifiable factual statements about product performance or attributes.[17] Factual advertising claims convey incomplete information about product performance. Advertising is limited to summary information because the cost of complete information exceeds its value. The model here focuses on the amount of statistical and experimental evidence sellers subject to FTC regulation will collect, and on the truth of the advertisers' claims given the available support.

The appropriate level of truth of ad claims is a statistical decision problem. Advertising claims are not absolutely true or false; some erroneous inference is always possible. Claims are more or less true, depending on the cost of error in statistical inference. There is a trade-off between more truth in advertising and the cost of achieving it. In other words, truth in advertising is an economic problem.

The FTC's ad substantiation program regulates the level of truth in advertising. Advertising claims are deemed true *ex ante* only if the advertiser possess-

[16] *E.g.*, Pitofsky, *Beyond Nader: Consumer Protection and the Regulation of Advertising*, 90 Harv. L. Rev. 661 (1977); Note: *The FTC Ad Substantiation Program*, 61 Geo. L.J. 1427 (1973). *But see* J. Galbraith, *The New Industrial State*, 332-3 (1968) (no need for regulation of advertising of non-lethal products, as consumers understand that advertising "requires well-considered mendacity").

[17] The line between puffery and factual claims is not always a bright one. By focusing on factual advertising we are not disparaging the information content of puffery which, according to Nelson, by its mere existence conveys valuable information to consumers about the utility of offers. *See* Nelson, *Information and Consumer Behavior*, 78 J. Pol. Econ. 311 (1970); *Advertising and Information*, 82 J. Pol. Econ. 729 (1974).

ed and relied on a reasonable basis in making the claims. Upon investigation, the FTC determines whether or not the sellers' claims were adequately substantiated. If the claims are found to be unsubstantiated, the Commission issues a complaint enjoining the advertising. Anticipation by advertisers and ad agencies of an FTC injunction and order, which if subsequently violated places the advertiser and the ad agency at the risk of incurring monetary fines, disciplines firms to supply the level of truth (*ex ante*) required by the FTC.[18] The FTC's ability to mandate truth is limited though, since it requires substantiation only for factual advertising claims. Advertisers can forego factual claims and escape FTC scrutiny. Nonfactual advertising ("puffery"), however, is not a perfect substitute for factual advertising.

We show in a stylized context how the FTC can redistribute wealth among advertisers in a market by regulating the level of truth in factual advertising. We suppose that all firms supply an identical product, that no firms have accumulated brand-name capital and that all advertising is factual advertising.

We also assume that the firms exhibit variable returns to scale and in the absence of any substantiation, all firms have the same minimum average cost but different optimal sizes. The industry supply curve is flat in the absence of substantiation. We further assume that for a given product and type of claim the requirements of substantiation (*e.g.*, a scientific test to substantiate a performance claim) impose identical fixed costs across the market. The per-unit costs of substantiation rise more for the smaller producers. Substantiation thus imparts an upward slope to the industry supply curve. In other words, marginal firms' costs are raised relative to those of their inframarginal rivals—all other things equal.[19]

The commodity has one quality dimension that is fixed and cannot be ascertained by consumers prior to purchase.[20] Producers also do not know quality with certainty. They must "test" to determine the quality of the commodity, and testing is subject to error. Thus, for a given amount of testing and a particular test outcome there is an optimal claim that the firms are justified in making. We suppose that FTC ad substantiation regulation mandates the level of testing and the appropriate claim for the specific test outcome. For any level of prior support imposed uniformly on all firms in the market there will be a distribution of "optimal" advertising claims. We assume that consumers rely on the average of these claims to estimate product performance. Thus, con-

[18] It is assumed that the cost imposed by FTC remedies is large enough that the FTC can regulate the level of truth in factual advertising. This assumption is apparently reasonable: Peltzman found that FTC challenges to advertising caused serious capital losses for the firms named. Peltzman, *The Effects of FTC Advertising Regulation*, 24 J.L. & Econ. 403 (1981). The cost of punishment to advertisers, discounted by the probability of being caught, includes the foregone benefits of the ad campaign enjoined, the cost of legal defense and investigation, and the loss in reputation occasioned by the adverse publicity.

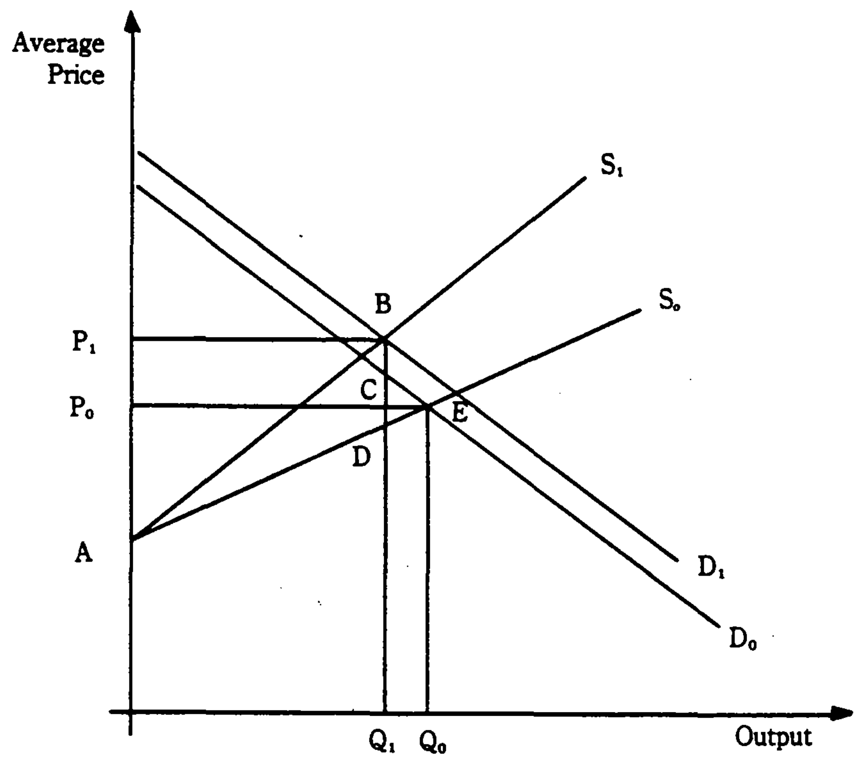
[19] *See* S. Salop, D. Scheffman and W. Schwartz, *A Bidding Analysis of Special Interest Regulation: Raising Rivals' Costs in a Rent-Seeking Society*, FTC Conference Volume on *The Political Economy of Regulation: Private Interests in the Regulatory Process*, (March 1984).

[20] For example, x% of the units sold "work," or the commodity works for x% of consumers.

sumers are assumed to rely on the FTC to assure that firms are making advertising claims that have a reasonable basis.

The welfare and wealth-redistribution effects of our stylized version of FTC advertising regulation are illustrated in Figure 1 for two different levels of required substantiation. Curves S_0 and S_1 depict supply curves in a competitive market for low and high levels, respectively, of required substantiation. By assumption, demand (D_1 and D_0) is linear in the average quality claim in the market. $D_1 > D_0$ because the accuracy of the average product claim based on a higher level of substantiation (testing) exceeds the accuracy of the claims made with less support.[21] We have not assumed anything about risk aversion to get this result; demand is raised because the cost of error from relying on statistical claims is smaller when the claims are more substantiated.

Figure 1



[21] To obtain this result we are implicitly assuming that true product quality (which is unknown) actually exceeds the minimum, that consumers know the level of substantiation the FTC typically requires for such claims (in fact, some advertising describes the tests on which performance claims are based), and that consumers assume the worst when there is no substantiation.

By imposing higher standards of truth in Figure 1, the FTC raises rent for the inframarginal firms from area AP_0CD to area AP_1B at a cost of CDE at the margin. By varying the required level of truth the FTC is able to divert wealth from marginal firms (and, ultimately, consumers) to inframarginal firms.

We allow that ad substantiation regulation provides benefits as well as costs. Because advertising claims based on greater support are more informative, additional reasonable-basis requirements shift demand and supply. But the marginal contribution of substantiation to demand price falls with greater substantiation while the cost at the margin rises with greater substantiation. Thus for initial low levels of substantiation additional substantiation requirements may benefit both marginal and inframarginal firms as demand shifts more than supply. However, we hypothesize that the FTC will avail itself of the opportunity to redistribute wealth to the inframarginal firms and push substantiation requirements beyond the point where consumer welfare is maximized.

Our hypothesis about the implementation of the ad substantiation program by the FTC is based on a simple view of regulation. The demand for truth in advertising is filtered by the political process so that regulation goes to the highest bidder. In other words, if by imposing more truth the inframarginal firms gain a dollar at the loss of less than a dollar to the marginal firms, then the additional truth is imposed. We realize that this is a very naive picture of the political process; we do not have a supply-side theory. In our defense, we note that the typical impressionistic description of the opposing interests manifest in the political process would tilt the regulatory solution even further toward the inframarginal firms. For example, it is not uncommon in these accounts to ignore consumer interests, which would press towards less substantiation, supposing consumers have too little at stake to justify organization.

B. *The Pfizer Doctrine and Some Predictions*

Based on the foregoing hypothesis of how advertising regulation works, we now note that *Pfizer* represents a regulatory innovation that altered the redistributive potential of FTC advertising regulation. Before *Pfizer*, an advertiser could avoid meeting FTC standards of truth *ex ante*, anticipating that, were he investigated, he could successfully supply, with some positive probability, the required substantiation of the truth of his claims (or the FTC interpretation of his claims, about which he may have been uncertain).[22] After *Pfizer*, truth is not a defense, making the probability of vindication *ex post* zero. As a consequence, the Pfizer doctrine raises the expected cost of punishment for a given penalty for violation, probability of detection, standard of truth and level of prior substantiation.

[22] Of course, it would be rare indeed for an advertiser's claims to be false *ex ante* and true *ex post*. But we are not dealing with all or nothing. Claims are more or less true; the appropriate margin is at stake. It is not unreasonable to suppose that an advertiser would make claims that are false *ex ante* according to FTC standards and are true *ex post* according to these same standards.

Thus the *Pfizer* doctrine represents a regulatory innovation that increases the FTC's ability to impose substantiation costs differentially on small and large firms, thereby raising the value of inframarginal firms.

Another prediction of our model of ad substantiation is that the ability of the FTC to transfer wealth depends on the degree of diversity of firm size within a product market. When firms' market shares are uniform the FTC is powerless to redistribute wealth. Thus we predict that the incidence of ad substantiation cases by product market will be positively related to diversity of firm size within the market.[23]

The prediction that the FTC will target markets where there is diversity in size distinguishes the economic model of advertising regulation from what might be termed the public-interest model. According to the latter, the FTC repairs an information-market failure. Suppliers can make their claims credible at a lower cost when the FTC stands ready "as cop on the beat" to penalize firms that exaggerate. The public-interest model predicts that case incidence is related to the supply of false advertising. Since there is no theoretical link between the amount of such advertising and diversity of firm size, *ceteris paribus*, the public-interest model predicts no relationship between case incidence and diversity, in contrast to the economic model. Moreover, the conventional wisdom that it is the small firms without brand-name capital that are most likely to supply false and unsubstantiated advertising, suggests that the public-interest model predicts that average firm size (or market share) in a product market will be negatively correlated with case incidence. Diversity in a market with many firms without reputations or in a market with a low average level of brand-name capital is irrelevant according to this hypothesis.

IV. Empirical Evidence

A. *Effects on Advertisers' Stock Prices*

The economic model of advertising regulation predicts that the ad substantiation doctrine first applied in *Pfizer* raises the value of large advertisers. Financial market analysis is used to measure the effects of the FTC's shift in advertising regulation from falsity *ex post* to substantiation *ex ante*. Financial market analysis is based on the efficient market theory of finance.[24] Security prices at

[23] Of course as a long-run consequence we expect ad substantiation to raise the optimal scale of a firm in its various product lines. The gains to inframarginal firms from the *Pfizer* doctrine thus appear to be transitional. With a positive rate of interest these gains still could be large enough to justify seeking the regulation. And, finally, even after all size adjustment has occurred, the ad substantiation program will continue to provide benefits to incumbent firms by raising the cost of entry.

[24] Schwert, *Measuring the Effects of Regulation: Evidence from the Capital Markets*, 24 J.L. & Econ. 121 (1981). For applications of financial market analysis to analyze regulatory changes, *see* Schwert, *Public Regulation of National Securities Exchanges: A Test of the Capture Hypothesis*, 8 Bell J.Econ. 128 (1977).

a given moment in time reflect all available information and discounted expected future returns, including risk adjustments. As new information affecting a security is received, the resulting wealth effect is reflected in the security's price. The wealth effect of a regulatory change expected to affect a firm's revenues and costs into the future is captured in the present price of the firm's securities as information of the regulatory change develops.

In financial market analysis, then, a principal concern is the timing of events that alter present values. A time "window" is investigated if a series of events contributed to changing expectations. In the case of ad substantiation, the FTC first signaled the policy change away from deception when it announced in April 1970 its intention to file a complaint against Pfizer (which it in fact did three months later). At that time, a Pfizer spokesman noted that the FTC was "seeking to establish new legal theories that have no precedent in cases that have been decided by the courts or by the FTC."[25] This was apparently the first public mention of substantiation; a review of the financial and advertising press in 1969 and 1970 discloses no earlier reference to it.

Thus any adjustment in firms' values would date from April 1970.[26] The adjustment should have been completed when the Commission published its *Pfizer* opinion. The relevant time window for measuring the value changes of firms affected by substantiation is therefore from April 1970 to July 1972.

1. Advertisers' Returns

The investigation of wealth changes due to the new substantiation requirement begins with an examination of monthly rates of return for large advertisers in two categories, "drugs and toiletries" and "home products and building products." Within these product categories, for each of 79 sub-classes (markets) defined in *Leading National Advertisers* (LNA), the largest firm in total 1970 advertising expenditures was selected.[27] This produced a portfolio of 54 firms

[25] "Pfizer's Advertising of Sunburn Reliever is Challenged by FTC," *Wall St. Journal*, April 14, 1970, p. 4, col. 2.

[26] Also well publicized was the Nader petition filed with the Commission in December 1970 that sought a rulemaking to require submission for public disclosure of all substantiation material. By that time, however, the market adjustment process should have been well underway. An FTC spokesman "indicated the proposed [Nader] requirement was in line with the agency's own recent efforts to require advertisers to substantiate claims."

> The proposed rule basically would broaden to an industrywide basis a principle the FTC already is seeking to establish on a case-by-case basis. In test cases, such as one against ad claims for Pfizer Inc.'s Unburn sunburn lotion, the FTC is seeking to require advertisers to have proof in advance that a claim is true rather than continuing the previous practice of placing on the agency the burden of proving that a claim is false.

"Nader's Bid to Require Proof of Ad Claims to be Studied by FTC," *Wall St. Journal*, Dec. 14, 1970, p. 8, col. 3.

[27] Leading National Advertisers, Inc., LNA Multi-Media Report Class/Brand $, January-December 1970; Media Records, Inc., Expenditures of National Advertisers in Newspapers Year 1970; Radio Expenditures Report, First Quarter-Fourth Quarter 1970.

(some firms were leaders in more than one market). The equally-weighted portfolio's monthly rate of return was regressed on the monthly returns for an equally weighted market portfolio of stocks traded on the New York Stock Exchange (NYSE), as reported by the Center for Research in Securities Prices (CRSP), and on a dummy variable for the time of adjustment to the new substantiation requirement.[28] The dummy was coded 1 for the period April 1970 to July 1972, zero otherwise.

The results are shown in Table 1. The coefficient for the market portfolio return is significant and positive, as one would predict.[29] The equation explains most of the variation in the dependent variable, and the Durbin-Watson statistic indicates no problem of first-order serial correlation. Notably, the coefficient for the time of adjustment to the ad substantiation program shows substantiation had a positive effect on the value of the market leaders, significant at the five percent level. These firms experienced an abnormal monthly increase in their rates of return of almost one percent for over 28 months.

B. *Incidence of FTC Substantiation Cases*

The public-interest and the economic hypotheses have opposing implications for the way in which the FTC implements the ad substantiation program. The economic model hypothesizes that substantiation exists to redistribute wealth. The necessary condition for the FTC to redistribute wealth is a differential advantage of large firms to spread the costs of substantiation. Thus the economic model predicts that the FTC will bring more ad substantiation cases in those product markets where firm size and reputation are more diverse, since that diversity gives the FTC greater power to redistribute wealth.[30] The mean level of firm size in the market and brand-name capital is irrelevant under the economic hypothesis. The market-failure model predicts, however, that lower mean levels of brand-name capital are precisely what attracts the attention of a publicly-interested FTC. The variance in size or reputation has no explanatory power under the public-interest model.

In short, the economic model predicts that variation matters (positively) in the incidence of FTC cases and the mean does not; the public-interest model predicts that the mean matters (negatively) but variation does not.

Perfect measures of firms' market size and reputation capital do not exist for appropriately defined product markets. As in the stock-market analysis of advertisers' returns, we adopted LNA's product sub-classes to define markets, and we used as proxies for firm market size and reputation a firm's 1970 total

---

[28] The portfolio for both the dependent and the independent variables exclude dividends.

[29] We also tested for the stability of the beta coefficient on the market portfolio return variable, and found that the introduction of the ad substantiation program caused no significant shift in the beta coefficient.

[30] It is not overall firm size that we refer to but the firm's size in a particular product market.

advertising expenditure in each market.[31] The mean and coefficient of variation for firm advertising expenditures in each market was computed. To define the dependent variable, we assigned FTC ad substantiation cases and consent orders from 1972 to 1980 to the relevant market.[32]

The results of regressing case incidence on the mean and variation of advertising expenditures are shown in Table 2. The only estimated parameter of significance is the coefficient of variation, and its positive sign is correctly predicted by the economic hypothesis. The test results refute the principal implication of the public-interest model. The mean of the reputation/size measure has no significant effect on the FTC's case activity.[33] When the coefficient of variation is the sole regressor, as suggested by the economic model, its significance increases,and the F-statistic rises to 5.15.

V. Conclusion

An economic model of FTC advertising regulation has been presented that views this regulation as a means by which large advertisers differentially impose costs on their rivals. The ad substantiation doctrine in particular is seen as a regulatory innovation that enhanced the FTC's ability to transfer wealth to large firms. The factual advertising not chilled by the higher cost of truth may contain greater truth than before *Pfizer*, but the cost of the extra truth exceeds its value. In any case, the integrity of advertising is a by-product that cloaks the aims of the large advertisers to restrict market penetration and en-

---

[31] There are several theoretical and empirical problems with using a firm's advertising expenditure as a proxy for its size or reputation in a market. *See* Nagle, *Do Advertising-Profitability Studies Really Show that Advertising Creates a Barrier to Entry?* 24 J.L. & Econ. 333 (1981). For example, in the advertising-profitability literature, a critical issue is the appropriate magnitude of the depreciation rate for advertising capital. For our purposes, whether advertising expenditure is a current or capital expense is irrelevant as long as depreciation rates do not vary substantially across products. In fact, there is evidence that retention rates differ across products. *See* Peles, *Rates of Amortization of Advertising Expenditures*, 79 J.Pol. Econ. 1032 (1971) and J. Lambin, *Advertising, Competition and Market Conduct in Oligopoly Over Time* (1976). The unavailability of a sufficiently large set of product-specific estimates of depreciation rates precluded adjustment of advertising expenditures in our product sample. However, we were able to apply a crude test based on evidence that depreciation rates are higher for durable goods. When we estimated the regression model in this section on the subsample, "drugs and toiletries," we got results identical to those for the full sample which included many markets for durables.

[32] We are grateful to Collot Gerard, Julie Niemasik and Bruce Levine of the Federal Trade Commission for the list of ad substantiation cases. There was a total of 45 cases.

[33] We also estimated the regression models of Table 2 including total market advertising expenditure as a regressor. Holding constant total advertising expenditure permits a test of an alternate version of the public-interest model in which mean firm shares, instead of mean size, is one of the key predictors of case incidence. The results do not support the public-interest model any more than those in Table 2; however, the significance of the variance coefficient was reduced to .08 in a one-tailed test. (Total advertising affects case incidence positively, but the effect is statistically insignificant.)

try by their rivals.

In support of our hypothesis we have presented some empirical evidence from the stock market that large advertisers' values did rise with the advent of the FTC's prior substantiation standard. Additionally, by relating ad substantiation case incidence to some characteristics of the product markets targeted by the FTC, we find support for the economic theory of regulation relative to the public-interest hypothsis.

Based on our findings we have several recommendations. We recommend more study; our results are certainly not dispositive. But they suggest that we should suspend judgment about the welfare effects of the program until more research is completed. In this regard, we applaud the efforts of our fellow panelist, and we eagerly await the research being conducted by the Bureau of Consumer Protection.

More daringly, perhaps, our research impels us to suggest a natural experiment. We would revive a policy proposal already made by others some time ago. Let the FTC give advertisers a choice between regulation by the FTC or no regulation. Those advertisers who think that consumers rely on the FTC's regulatory presence in assessing the credibility of their ad claims would submit their substantiation for review by the FTC. When the Commission found the claims adequately substantiated, the advertisers could mark their advertisements with an FTC logo.[34] The choice to undergo FTC scrutiny need not delay introduction of an ad campaign, since the claims could be made with the notation "FTC approval pending." Those advertisers who think FTC regulation is overly burdensome and intrusive could avoid the Commission's oversight altogether by refraining from seeking certification for their ads. With this rule in place we could readily observe the demand for an FTC role in advertising and gauge the relative informativeness of advertising with and without the FTC's sanction. If markets have truly failed, the demand for FTC regulation will be brisk, and the FTC's own claims for its program will be substantiated.

[34] After some experience with the program, the FTC may find that user fees can be charged to cover the cost of enforcement.

Table 1

REGRESSION COEFFICIENTS FOR THE EFFECTS OF ADVERTISING SUBSTANTIATION ON RETURNS TO PORTFOLIO OF LEADING ADVERTISERS

	Coefficient (t-statistic)
Intercept	-.002
	(-1.10)
Market Return	.691**
	(12.78)
Event Dummy	.009*
	(1.71)
F = 182.58**	
R^2 = .78	
Durbin-Watson = 2.26	

*Significant at .05 level in one-tailed test.
**Signficiant at .01 level.

Source: Estimates from monthly returns tape (1964-1978), Center for Research in Securities Prices.

Table 2

INCIDENCE OF FEDERAL TRADE COMMISSION AD SUBSTANTIATION CASES AS A FUNCTION OF ADVERTISING VOLUME, 1972-1980

	Coefficient (t-statistic)
Intercept	-.360
	(-.80)
Mean Advertising Expenditures	$3.1\text{x}10^{-5}$
	(.27)
Coefficient of Variation for Advertising Expenditures	.428*
	(2.27)
F-statistic	2.58
R^2	.07

*Significant at .01 level in one-tailed test.

Comments

*J. Howard Beales III**
Associate Director for Policy & Evaluation
Bureau of Consumer Protection
Federal Trade Commission

The legal theory of the advertising substantiation program requires an advertiser to have a "reasonable basis" before making objective product claims in advertising. The papers presented today have viewed this doctrine as a different species of advertising regulation. In fact, however, substantiation is simply a more efficient regulatory tool to control deception.

For an advertiser contemplating the possibility of a deceptive claim, the likelihood of getting caught, the consequences of conviction, and the standard for judging the truthfulness of the claim combine to determine an expected cost of violation. This expected cost creates an incentive for advertisers to determine the truth of a claim before making it. Incentives to determine whether the claim is true may be adequate or inadequate, but as long as making false claims has some consequences, an incentive to check in advance exists. Advertising substantiation can change the amount of prior checking of the truth of a claim, but it does so by raising the expected cost of punishment if a firm deceives. Nonetheless, there exists some *ex post* regulatory strategy with exactly the same effects as the ad substantiation program. Because an equivalent *ex post* strategy would not alter the resources advertisers devote to verification of claims in advance, the question should be the effects of different strategies, and not whether those strategies are implemented *ex ante* or *ex post*.

The key question in assessing the effects of the advertising substantiation program is the standard of truth advertisers are required to meet. This is perhaps most apparent in the Higgins-McChesney model of substantiation. In their model, the efficient solution is for all firms to make exactly the same claim and conduct a single test. The advertising substantiation doctrine allows this outcome, so any effects of the doctrine must come from changing the level of testing. Changing the level of testing, however, has very little to do with whether the standard is *ex post* or *ex ante*. Rather, it deals with the standard of truth — how certain must we be that a claim is truthful, before the claim is judged permissible?

With this perspective on the program in mind, consider the effects observed by Higgins and McChesney. They first examine the impact of the substantiation doctrine on large national advertisers, as reflected in financial markets. Although they find a significant effect, it is unclear whether that effect is a cost or a benefit. If the amount of testing is suboptimal without a substantiation re-

*The views in this paper are my own. They do not necessarily reflect the views of the Bureau of Consumer Protection, the Commission or any individual Commissioner.

quirement, the program generates net benefits. Because testing costs are fixed, however, increased testing will benefit large or inframarginal firms relative to smaller ones. Thus, whether the advertising substantiation program is efficient or not, it will nonetheless generate wealth transfers.

A differential impact of the substantiation requirement on advertisers of different size does not depend exclusively on the fixity of testing costs, however. If the program works as intended, it should enhance the credibility of advertising. The effect of such an enhancement is analogous to a reduction in the price of advertising, and, of course, large advertisers tend to benefit more. This also suggests a test of the competing hypotheses. The public interest theory, that advertising substantiation is efficient, implies that the financial market effect should be proportional to the number of ads run by a company. The Higgins-McChesney theory, in contrast, implies that the effects should be proportional to the number of units of the product sold. Admittedly such a test is difficult to undertake, but conceptually it can distinguish the competing hypotheses.

The Higgins and McChesney results, in conjunction with the Leffler and Sauer results, also suggest a puzzle. Like Peltzman's study of the impact of FTC advertising cases,[1] Higgins and McChesney find large financial market effects. It seems difficult to find much impact in the product or advertising markets, however. Measured by the expectations of financial markets, the effects of substantiation are enormous. Measured by the apparent actual effects in advertising markets, at least as found by Leffler and Sauer, the effects are almost non-existent.

Higgins and McChesney also seek to evaluate the substantiation program by examining competing theories of what factors should influence the likelihood of an FTC case. They first examine mean advertising expenditures in an industry, arguing that a substantiation program that primarily serves the public interest should bring cases against firms with little advertising capital, while rent seeking implies no such relation. Unfortunately, the theoretical importance of average advertising expenditures under a public interest theory is unclear. Firms with little brand name capital are more likely to engage in deception, as Higgins and McChesney note, but for a given deceptive claim, benefits are greater from suing a larger firm. It is not apparent which effect will dominate.

Superficially more compelling is the Higgins and McChesney test based on the coefficient of variation. The authors argue that a substantiation program that primarily redistributes wealth among large and small firms should bring cases in industries with diverse firm sizes. They further contend that no such implication follows from a public interest theory. There is, however, a theoretical link between diversity of firm size and the likelihood of deception. First, honest firms are likely to grow relative to dishonest ones. Even if all firms are initially the same size, as honest firms grow, the presence of deception will generate disparities in firm size. The effect is analogous to an industry with different firms

[1] Sam Peltzman, "The Effects of FTC Advertising Regulation," *Journal of Law and Economics*, December 1981, Vol. 24, p. 403.

charging different prices. Because firms with lower prices will tend to grow, there is a theoretical relationship between the firm size distribution and the price distribution. Similarly, the existence of market forces that check deception implies that deception will be more common when a fringe of dishonest firms competes with a core of honest companies.

Second, some market forces that limit deception may operate less effectively if firm sizes diverge. For example, large well known firms are probably less reluctant to counter claims by a rival of similar size than to name small, unknown competitors in counteradvertising. To the extent that such factors are significant market incentives for truthful advertising, uniform firm sizes should promote honest advertising.

The Leffler and Sauer paper represents a search for correlates of the advertising substantiation program that may represent effects of the program on advertising agencies. Little impact is apparent in the results. Again, one possible reason is that substantiation cases and deception cases are alternative means to the same end. Leffler and Sauer's suggestion that the theories are not substitutes seems to confuse changes in the level of activity with changes in the theory employed. The years surrounding the emergence of the substantiation doctrine in 1972 represent the peak years of advertising regulation in the period under study. Between 1970 and 1972, 76 percent of the Commission's 46 advertising cases were based on a deception theory. Between 1973 and 1975, 73 percent of 60 total cases were substantiation cases. This remarkable shift in the proportion of cases under each theory would seem to indicate that the approaches are substitutes. An increase in the overall level of activity coincided with the advent of substantiation, but the level of activity tells us little about the effects, if any, of the theory employed.

The primary effect that Leffler and Sauer find is that ad substantiation seems to be correlated with an increase in the market share of advertising agencies ranked between 9 and 20. Strong theoretical reasons for an impact of ad substantiation on agencies that are not too small and not too large are difficult to identify. Such a theory may be unnecessary, however. The Leffler and Sauer results are not overwhelming, and not fully consistent. For example, for large agencies (ranked 1 to 8), the effect of ad substantiation measures is negative but insignificant if market share is the dependent variable, but negative and significant (at least for one measure) when the dependent variable is agency billings. Despite the significant positive effect of substantiation on the market share of middle-sized agencies, substantiation has a negative (but insignificant) influence on their billings. Medium-sized agencies seem slightly less likely to be named in FTC complaints relative to their billings than larger or smaller agencies, but any targeting effect does not seem sufficient to have influenced advertiser choices among agencies generally. Regressions explaining account stability do not indicate any significant effects of agency size, either before or after the substantiation doctrine. The pattern of results can to some extent be rationalized, and might provide persuasive evidence for an otherwise compelling theoretical prediction. They hardly seem to demonstrate the kind of strong empirical regularity that cries out for a theoretical explanation.

The search for effects of the advertising substantiation program must take into account how that program fits into the larger scheme of advertising regulation. Unless we assume, contrary to other evidence, that other forms of advertising regulation have no effects, we cannot learn much about substantiation without taking into account ad regulation based on other legal theories. Unless we control for the overall level of activity, we cannot tell whether or not substantiation is simply a different legal theory to reach the same end as cases based directly on deception. Any observed effects that seem to stem from the substantiation doctrine could just as easily stem from the level of activity, not the legal theory employed to prosecute deception. To examine the overall impact of advertising regulation, we should examine the total number of cases, without regard to the legal theory employed. To examine the impact of the theory, we should seek *differences* in impact between substantiation cases and deception cases. There is little basis in either of these papers for suspecting such an effect.

Comments

Pablo T. Spiller
The Hoover Institution
Stanford University

The two papers on the FTC's Advertising Substantiation Program (ASP) complement each other. While the Higgins and McChesney paper tries to measure the effect of the ASP on advertisers' stock prices and to explain its incidence, Leffler and Sauer try to find the effect of ASP on advertising agencies' performance. The results of the two papers are seemingly at odds: while Higgins and McChesney find large effects on large advertisers' stock prices, Leffler and Sauer find little if any effect on advertising agencies' performance. In the following comments I will discuss each paper and try to provide an explanation for the seemingly different results.

Let me first start with the Higgins and McChesney paper. The authors discuss in sections IIB and IIIA two models of advertising, and analyze the potential implications of ASP. The model they discuss in section IIB is an application of the Klein and Leffler (1981) model and its implications can be summarized as follows: Firms build reputation implicitly to commit themselves to maintain the quality of their products. Thus, firms with brand name will be able to transmit hard information, which consumers may actually base their decisions on. That is, their quality statements are believable. On the other hand, firms which are entering the market and do not have a reputation yet, will not be able to transmit credible statements. The only strategy for them is to advertise a la Nelson, that is with "puffery," since no information message would be believed by customers. Introduce ASP and firms that so far were not able to transmit "hard" information are now able to do so, since consumers, knowing that firms can be prosecuted for unsubstantiated statements will tend to believe what firms say. These firms, then, are better off with ASP, since it essentially reduces their introduction costs. Large firms on the other hand, those with large investments in brand name, will find that they are worse off. Some of their investments in brand name will no longer provide a return, since those investments are not needed to transmit their informational advertising (still reputation may be needed to be able to promise the continuation of the optimal product quality). That is, the Klein and Leffler model predicts that, if consumers believe that the FTC will enforce ASP, firms that are building their reputation will benefit more than those that are trying to maintain it. I will refer to this model as the "reducing entrants' costs" model of the FTC's ASP. As we will see momentarily, this model is consistent with both Higgins and McChesney's and Leffler and Sauer's results.

In section IIIA, the authors develop a model of the Salop and Scheffman (1983) "raising rivals' costs" type. While being sympathetic to the idea, I do not think that the model the authors develop properly describes an industry where firms advertise. First, in their model advertising is about "quality," in spite of all products being identical. Second, consumers assess the "quality" of all identical products by looking at the average of all claims. Thus, if there are costs involved in testing, no firm will advertise any claim, and the introduction of ASP should not have any effect on anyone. The authors, however, conclude that ASP should hurt the marginal (small) firms while benefit — via a price increase — the inframarginal (large) ones. I believe this result would follow from a "raising rivals' costs" model in which firms advertise because it is profitable to do so. My own "raising rival's costs" model is based on the Klein-Leffler model previously described. Since entrants are those that need to advertise factual claims (how otherwise would you know that a Mitsubishi uses one gallon of gasoline for 44 miles), and since those firms may also benefit from ASP (see above), these firms, however, may be exposed to much more bureaucratic scrutiny than incumbent firms. Thus ASP bureaucratic costs may more than compensate for the implied reduction in required advertising in such a way that firms trying to introduce new products may actually be hurt from ASP (observe that while Leffler and Sauer say that the bureaucratic cost does not seem to be large, the firms they base their statement on are well established ones).

We have two theories of ASP, one where the introduction of ASP "reduces entrants' costs" and another where it "raises rivals' costs." I will now discuss the authors' empirics and relate them to these two theories.

Higgins and McChesney's first result (Table 1) is that a portfolio of 54 leading advertisers had from April 1970 to July 1972 an excess return of around 25 percent. The authors attribute those gains to the introduction of ASP and suggest it supports the "raising rivals' costs" view against the "reducing entrants' costs" one. There are two problems with this result. First, a time span of 28 months is extremely large. During that period many other things may have happened that may have affected the returns to large advertisers (for example, the ban on cigarettes' TV advertising may have reduced the price of TV ads and therefore improved the performance of large — non-cigarette — advertisers). I would recommend a much more precise test where each announcement on ASP is analyzed individually and the overall effect of ASP is derived from the cumulation of the effects of the individual announcements (for example, see my (1983) paper using such a methodology). Second, even if we grant the existence of such gains, it is not clear who really benefited from ASP. According to the "reducing entrants' costs" theory, firms introducing new products are the ones that (before ASP) were building the required stock of reputation through (perhaps large) advertising campaigns and they may have also been the larger advertisers. It is then consistent with that theory that these firms should gain from ASP. The raising rival's costs theory, however, implies that those same entrants should be hurt.

The second result relates to the incidence of ASP (Table 2). Table 2 should be made clearer. In particular, the authors do not say whether they also include industries which did not have any ASP cases. If that is the case, then most of the dependent variable values will be zero and thus OLS is inappropriate. If they did not include them, then the standard selection problem is present and again OLS is inappropriate. I encourage the authors to deal with this specification problem. If we assume that the results in Table 2 are robust to the problem just described, then, I think this test cannot differentiate between the two alternative models. Since we should observe large cross sectional variation of advertising expenditures in industries where new entrants or products are present, and since these are the industries where "factual" claims are going to be usually made, it is not surprising that the incidence of ASP cases is larger in those industries. On the one hand these are the industries where truly unsubstantiated claims may appear. On the other hand, these are the industries where incumbents may use the bureaucratic apparatus to deter entry.

The paper by Higgins and McChesney raises interesting issues. However, much still has to done to be able to conclude what the effects of ASP are and why it was introduced.

Let me now discuss the paper by Leffler and Sauer and relate their results to the two theories of ASP.

While Leffler and Sauer do not develop a particular theory explaining the effects of ASP, it seems that they also have a "raising rivals' cost" theory of regulation, which involves one group of advertising agencies promoting regulations that will increase their rivals' cost. They do not explain clearly which group of agencies should benefit nor through which mechanism. They state as their objective to empirically determine whether larger agencies were affected differently than the smaller ones. The authors try to answer this question by analyzing whether ASP affected (a) the degree of concentration, (b) the volume of sales of the different agency groups, and (c) the probability of losing an account.

Table D presents results of regressions which try to explain the degree of concentration as well as the volume of sales. The main results from this table are: (1) the market share of the medium-sized agencies increased, at the expense, it seems, of the smallest and perhaps also of the largest ones; (2) ASP reduced the demand for all advertising agencies, but perhaps more of that of the largest ones. The latter result is very interesting since it fits the prediction of the "reducing entrants' costs" theory. Also, in a study that Tom Ross and myself are currently working on, we found that around the Pfizer case (actually, 20 trading days before to 20 trading days after the first public announcement in the Wall Street Journal of the Pfizer case), a portfolio of media firms (including TV, radio, newspapers and magazines) had an abnormal return of -7 percent (with a t-statistic of around 3.5). Moreover, when we estimated monthly time series models for advertising expenditures by different types of media, we found that there is a change of regime in 1971, not only for TV, but for all other media as well. This change in regime implied a reduction in advertising following 1971 for all types of media. Thus, both our results and those of Leffler and Sauer suggest that ASP did have a negative effect on overall advertising (interestingly, the change in regime is permanent and does not disappear as of 1977 when the FTC's measure of ASP activity seems to have declined). This reduction in advertising is consistent with the "reducing entrants' costs" theory. Since advertising messages are now more credible, firms may need a smaller amount of reputation capital; thus the demand for advertising may fall. Why should this reduction in advertising affect mostly the largest (since their market share did not change so much) and the smallest firms? The authors try to answer this question by looking at the incidence of ASP cases by agency size and whether ASP cases affected sales and turnover accounts. Their answer is inconclusive. While it seems to be true that being named on a case implies some losses of revenue (perhaps because the quantity of ads performed by the account named in the complaint falls), it does not imply that the agency's probability of losing an account increases. Thus, we are left without an understanding of why large and small firms may have lost more sales than medium-sized firms. My impression is that to answer this question, much more has to be learned about the determinants of the size distribution of ad agencies.

To summarize, the effects of ASP on different advertisers and advertising firms are unclear. While the evidence presented here and elsewhere suggests that ASP reduced the demand for advertising, it is nevertheless unclear who, if anybody, actually benefited from its introduction. In neither study did the proposed "raising rivals' cost" theories fare well. The evidence tends to support, albeit not too strongly, the "reducing entrants' costs" one.

Additional References:

Salop, S. and Scheffman D., "Raising Rivals' Cost," *American Economic Review*, May 1983.

Spiller, P. T., "Assessing the Profitability Effects of Airline Deregulation," *Economic Letters*, 1983.

Quality Uncertainty and Bundling*

Charles A. Holt and *Roger Sherman*
University of Virginia

Bundling is a practice of selling certain goods together rather than individually. Probably the most celebrated bundle arose with the "block booking" of motion pictures, a practice from the early days of film exhibition whose legality was disputed off and on from 1938 until the Paramount and Loew's decisions in 1948 and 1962 made it illegal for theatrical and television exhibition.[1] George Stigler has provided a classic price-discrimination rationale for block booking.[2] Interest in the practice of bundling was revived when James Adams and Janet Yellen[3] offered a new framework to show why a two-product monopolist facing independent demands could prefer to bundle its offerings. Richard Schmalensee[4] extended this study of bundling incentives to the monopolist who produces two products but has a monopoly on only one of them. And by introducing diversity of consumers' bundle preferences, Robert Dansby and Cecilia Conrad[5] showed an incentive for a firm to bundle even though it had no monopoly power. All of these analyses of bundling incentives assumed complete information and certainty.

Roy Kenney and Benjamin Klein[6] demonstrated that a motive to bundle can arise when information about the quality of individual items is imperfect. Buyer search can then be economized if bundles of units are sold that are reliable in their average quality level, rather than individual units whose quality must be separately determined. Kenney and Klein illustrated this argument with an analysis of the De Beers Central Selling Organization, a monopoly arrangement for selling diamonds in boxes, or "sights," of numerous presorted groups rather than individual diamonds. Besides avoiding buyer's search of all diamonds in the market, sale of the diamonds in groups prevented rejection of lower quality diamonds within any particular category. Kenney and Klein also showed how block booking of motion pictures was similar to this diamond market practice. Film distributors typically possessed *de facto* monopoly power, so the market

* The authors wish to acknowledge the assistance of Esther Cash in conducting the experiments and preparing the manuscript, and are grateful to James Mackay and Ross Miller for helpful comments.

[1] *U.S. v. Paramount Pictures, Inc.*, 344 U.S. 131 (1948), and *U.S. v. Loew's, Inc.*, 371 U.S. 38 (1962).

[2] George J. Stigler, "*United States v. Loew's, Inc.*: A Note on Block Booking," *1963 Sup. Ct. Rev.* 152 (1963).

[3] *See* W. James Adams and Janet L. Yellen, "Commodity Bundling and the Burden of Monopoly," 90, *Q.J. Econ.* 475 (1976).

[4] Richard Schmalensee, "Commodity Bundling by Single-Product Monopolies," 25, *J. Law and Econ.* 67 (1982).

[5] Robert E. Dansby and Cecilia Conrad, "Commodity Bundling," 74, *American Economic Review* 377 (1984).

[6] *See* Roy W. Kenney and Benjamin Klein, "The Economics of Block Booking," 26, *J. Law and Econ.* 497 (1983).

structure was similar. And both inventory cost savings and film exhibition scheduling economies could be obtained from block booking, which prevented reneging by exhibitors on agreements to show less successful films while retaining rights to show hit movies.

Bundling appears to offer a means by which buyers and sellers can moderate the effects of exogenous uncertainty, and so its use may be motivated in certain cases even though monopoly power is absent. Bundling can discourage selective rejection of items below average in quality, and this is desirable if it is costly or impossible for buyers to inspect prior to purchase or if it is costly or impossible for sellers to resell rejected items. A more detailed discussion of the effects of bundling together units of uncertain quality is contained in section I. One of our objectives in this paper is to specify a simple model in which bundling can be an efficient means of moderating the effects of uncertainty. We do this in section II. Once it is shown that bundling can improve market efficiency in the context of this model, our second aim is to determine whether bundling will emerge in a competitive market situation. We explore this question in an experimental double oral auction market which is described in section III. Units of uncertain quality are bought and sold in the market by buyers and sellers who can choose to transact in individual units or in two-unit bundles. Section IV contains the evidence from the competitive experimental markets which indicates that bundles are preferred when they offer an efficiency advantage. The final section contains a summary.

I. Bundling When Quality Is Uncertain

We consider a market in which there is exogenous production uncertainty. Some units of the product may turn out to be of relatively low quality, and advance identification of low-quality units is either impossible or prohibitively costly. By combining units together and selling them as one bundle, a seller can prevent rejection of lower quality units which might be contained within the bundle. The seller can still allow a right of rejection by buyers, but only of the entire bundle rather than of the least valuable units within the bundle. This practice of bundling can be socially beneficial, as Kenney and Klein have shown for the diamond market,[7] when it reduces the wasteful search buyers would otherwise make to identify the least valuable units. With widespread agreement on the values of individual units that fall within a quality category, the sale of bundles that contain a fairly reliable average quality level avoids separating units according to the quality of every unit. For fostering transactions, sale by bundles can have virtually the same effect as identifying the quality of each unit and therefore can be efficient by preventing every buyer from having to search to determine the quality of each unit. But in addition the rejection practice itself,

[7] *See* Kenney and Klein, *ibid*.

which may be necessary to attract buyers, can be wasteful of the rejected units if their value decreases with time, or if for any other reason recontracting is costly. By making quality more reliable, the practice of bundling can reduce the extent to which the rejection option is exercised, and that can improve efficiency.

As an example of these effects of bundling, consider the negotiations between distributors and exhibitors of motion pictures. New information about the quality of a film may become available to a particular exhibitor between the time a contract is signed and the time the film is delivered to the exhibitor. Motion pictures that turn out to be of low quality, in that they do not appeal to large audiences, may nevertheless be of some value to exhibitors. If contracted for in advance with the right of inspection and rejection, these films might be rejected, though, and so would be wasted, while hit movies would be accepted. An agreement in advance, by which exhibitors would accept a block of movies including only some hits, could offer the advantage that the lower quality films would have exhibition use *without costly recontracting*. The costs of rejection and recontracting could be significant because rejection of individual films might greatly complicate exhibition scheduling. Assuming that a right of rejection is to be offered to the buyer, then it might be better for both parties if that right applies to a set of films rather than to individual films. This right of rejection will protect the buyer while at the same time avoiding for the seller the complicated rearranging of exhibition schedules that might follow when individual films can be rejected. By having the right of rejection apply to the bundle, selective film rejection, with attendant inefficiencies, is less likely to occur.

To summarize, uncertainty about quality can add to the advantages of bundling. It is possible for quality to be uncertain at the time of production, as for example in the making of a motion picture, where no amount of search can determine the ultimate appeal of the film. Bundling units together might then result in a package containing items of reasonable average quality which are all usable to a buyer. On the other hand, sale of individual items with inspection and rejection privileges might cause items of below-average quality to go unsold, or the rejection of low-quality units might result in significant recontracting costs. In a case like this where high transaction costs arise after uncertainty is resolved, the bundle could offer efficiency advantages. The bundle can moderate the effect of uncertainty by fostering transfer of ownership before uncertainty is resolved, in a manner that will allow sale of low-quality as well as high-quality items.

II. A Model of Bundling Efficiency

Throughout our discussion we shall consider a market for a commodity with only two possible quality levels: high or low, each occurring with probability one half. These probabilities are independent across units and are independent of seller decisions, so there is exogenous production uncertainty. Units are sold before quality is known, either singly or in bundles, and buyers have a money-back guarantee on each purchase contract. Buyers and sellers are risk neutral. In any one time period, each seller produces N units at a common, constant unit cost denoted by C, where N is an even integer and $C > 0$. Each buyer may purchase up to N units per period. All low-quality units have a monetary value, V_L, common to all buyers. The monetary value of a high-quality unit depends on the number of high quality units consumed by the buyer in that period. We let V_1 be the monetary value of the *first* high-quality unit for each buyer in a period, and let V_H be the value of all subsequent high-quality units; thus $V_1 > V_H > V_L$.

In this market sellers will sell their units either singly or in two-unit bundles. When sellers offer inspection and rejection rights to buyers, in the case of a bundle purchase the bundle itself must be accepted or rejected; buyers are not permitted to reject part of a bundle. We focus on mandatory guarantees, but the case of optional guarantees will also be considered.

This market structure is rich enough to provide an efficiency advantage for bundled sales, but it is simple enough to serve as the basis for the laboratory experiments that will be discussed in later sections. We shall now elaborate in more detail two models in which quality uncertainty can make bundling an efficient means of transacting. In the first case, discussed in subsection A, market participants have two units to buy and sell in each market period, so the unspecified N in the description above is set at two. In subsection B, buyers and sellers deal with four units each period.

A. *A Model with Two-Units per Trader*

We shall consider first an especially simple case in which each buyer and each seller can deal with at most two units in any one market period, so that $N = 2$. There will be an equal number of buyers and sellers. Besides its simplicity, this setting has the advantage of forcing on every market participant a choice between dealing in single units and dealing in bundles, for with only two units at hand one cannot deal in both. We analyze first the case of single-unit sales, then consider bundles, and finally compare the two possibilities.

1. Single-Unit Sales

Each unit has a 50-50 chance of turning out to be of high or low quality. The probabilities of obtaining a high-quality unit are independent across sellers and across units. For simplicity, we analyze transactions of single units for a paired buyer and seller. Then we shall subsequently consider a bundle transaction for a buyer-seller pair. Units sold by a seller may be rejected, and the rejections experienced will depend on the sale price of the units.

Let p_1 denote the price of a single unit. Obviously, no units sold at p_1 would be rejected if $p_1 < V_L$, but all units would be rejected if $p_1 > V_1$, and indeed none should even be purchased if $p_1 > V_1$. Recall that $V_L < V_H < V_1$, so there are two intermediate cases: $V_L < p_1 < V_H$ and $V_H \leq p_1 < V_1$. The case in which the price is exactly equal to V_H is ambiguous without some convention about whether buyers will purchase a unit that yields a zero profit to them. If buyers sometimes do not accept such zero-profit units, then sellers would prefer to reduce price slightly. For simplicity, we consider only strict inequalities. With $V_L < p_1 < V_H$, all high-quality units ("highs") would be accepted and all low-quality units ("lows") would be rejected. With $V_H < p_1 < V_1$, a buyer would only accept a unit if it is the first high-quality unit for that buyer in that period. We next analyze the seller's expected profit for these cases.

For the range of prices $V_L < p_1 < V_H$, a buyer would reject all lows and accept all highs. The probabilities of obtaining two highs, two lows, and a high/low combination are 1/4, 1/4, and 1/2 respectively. Thus the seller's expected profit as a function of p_1, denoted $E^s(\pi(p_1))$, is computed:

$$E^s(\pi(p_1)) = (\tfrac{1}{2})p_1 + (\tfrac{1}{4})2p_1 - 2C = p_1 - 2C \tag{1}$$

where C is the uniform unit cost. We assume that the cost C is incurred at the time of production, before quality is known and buyer inspection and possible rejection occur.

We shall choose parameter values that force price into the range ($V_L < p_1 < V_H$) just considered. To do so we first determine values of V_1 and V_H so that sellers' expected profits cannot be increased by raising price above V_H. A price increase to a level above V_1 would result in all units being rejected, so instead consider a price increase to a level just below V_1. At such a price, a buyer would accept a unit only if it was the buyer's first high-quality unit. The seller in this buyer/seller pair would therefore sell one unit in every case except that of two lows, and so would sell with a probability of .75. Thus the seller's expected profit would be approximately $(\tfrac{3}{4})V_1 - 2C$. (This is an approximation because it assumes that $p_1 = V_1$ while the price actually is slightly below V_1.) The previous result in (1) indicates that a price just below V_H would yield an expected profit of approximately $V_H - 2C$, so there will be no incentive to raise p_1 above V_H when $V_H - 2C \geq (\tfrac{3}{4})V_1 - 2C$, or equivalently, when

$$V_1 \leq (4/3)V_H. \tag{2}$$

In the experiments we shall choose parameter values to satisfy this weak inequality, and the buyer preferences for lower prices should then prevent p_1 from exceeding V_H. We shall also set the seller's average cost C to be at least as great as V_L:

$$C \geq V_L, \tag{3}$$

so we can expect to see p_1 exceed V_L. Thus our parameter choices will allow us to focus on the case where $V_L < p_1 < V_H$.

A buyer will reject all lows and accept all highs when $V_L < p_1 < V_H$, so the buyer's expected profit as a function of p_1, denoted $E^B(\pi(p_1))$, is computed:

$$(4) \qquad E^B(\pi(p_1)) = (1/2)(V_1 - p_1) + (1/4)(V_1 + V_H - 2p_1) = (3/4)V_1 + (1/4)V_H - p_1.$$

The expected total surplus, obtained by summing equations (1) and (4), is independent of the transfer price p_1:

$$(5) \qquad E^B(\pi(p_1)) + E^S(\pi(p_1)) = (3/4)V_1 + (1/4)V_H - 2C.$$

2. Sales in Bundles

When a seller's two units are sold in a bundle, the likelihood of rejection will depend on the level of the bundle price, which is denoted by p_2. A bundle is never rejected if $p_2 < 2V_L$. Only bundles composed of two low-quality units will be rejected if $2V_L < p_2 < V_1 + V_L$. Both a bundle with two lows and a mixed bundle with a high and a low will be rejected if $V_1 + V_L < p_2 < V_1 + V_H$. All bundles will be rejected if $p_2 > V_1 + V_H$, so p_2 will not be that high. If inequality (3) is satisfied, then we expect that p_2 will exceed the level $2V_L$. Thus there are two ranges of prices to be considered:

$$2V_L < p_2 < V_1 + V_L \text{ and } V_1 + V_L < p_2 < V_1 + V_H.$$

We first compute expected seller profit for a bundle price in the lower range, and then we use inequality (2) to show that the seller's expected profit cannot be increased by raising the bundle price above the upper limit of this lower range.

If $2V_L < p_2 < V_1 + V_L$, mixed bundles and bundles with two high-quality units will be accepted, and only bundles with two low-quality units will be rejected. Thus the probability that a bundle will be accepted at a price in the range being considered is 3/4, and the seller's expected profit as a function of p_2, denoted $E^S(\pi(p_2))$, is:

$$(6) \qquad E^S(\pi(p_2)) = (3/4)p_2 - 2C.$$

Remember that (6) holds for $2V_L < p_2 < V_1 + V_L$. An increase in p_2 above the level $V_1 + V_L$ would result in the rejection of mixed bundles, which occur with probability 1/2. A bundle with such a higher price would be accepted with probability 1/4 as long as $p_2 < V_1 + V_H$. The most favorable price for a seller in this higher price range is slightly below $V_1 + V_H$, and the seller's expected profit at such a price would be approximately $(1/4)(V_1 + V_H) - 2C$. It is straightforward to use inequality (2) to show that this expected profit is less than the expected profit determined by (6) when p_2 is approximately equal to the highest price in the lower range, $V_1 + V_L$. Thus a risk-neutral seller would not raise price above $V_1 + V_L$.

As long as $p_2 > 2V_L$, a buyer will only accept mixed bundles, which occur with probability 1/2, and bundles with two highs, which occur with probability 1/4. Thus in the feasible price range $2V_L < p_2 < V_1 + V_L$, the expected buyer's profit as a function of p_2, denoted $E^B(\pi(p_2))$, is:

$$\text{(7)} \qquad E^B(\pi(p_2)) = (1/2)(V_1 + V_L - p_2) + (1/4)(V_1 + V_H - p_2) = (3/4)V_1 + (1/2)V_L + (1/4)V_H - (3/4)p_2.$$

The expected value of total surplus, obtained by adding equations (6) and (7), is:

$$\text{(8)} \qquad E^B(\pi(p_2)) + E^S(\pi(p_2)) = (3/4)V_1 + (1/2)V_L + (1/4)V_H - 2C.$$

3. Comparison of Single-Unit and Bundle Sales

At this point it is possible to evaluate the efficiency advantage of bundled sales in this model. When inequalities (2) and (3) are satisfied, we expect that: (1) the singles price p_1 will be in a range (V_L, V_H) where only low-quality units will be rejected, and (2) the bundle price p_2 will be in a range $(V_L, V1)$), is computed:

$$\text{(1)} \qquad E^S(\pi(p_1)) = (1/2)p_1 + (1/4)2p_1 - 2C = p_1 - 2C$$

where C is the uniform unit cost. We assume that the cost C is incu in a bundle. A comparison of (5) and (8) shows that the total surplus in the case of the bundle sale is greater by an amount $V_L/2$. This is because half of the low-quality units are not rejected when sold in bundles, but are all rejected when sold singly.

Notice that only certain values of high and low quality units will cause bundling to be preferred by both buyers and sellers. In this example, if prices are at the upper end of price ranges $V_L < p_1 \leq V_H$ and $2V_L < p_2 \leq V_1 + V_L$, a comparison of expected profits will show that the buyers prefer bundling for $(3/4)V_1 + (1/4)V_L < V_H$ and the sellers prefer bundling for $(3/4)(V_1 + V_L) > V_H$. We normally would expect bundling to emerge more reliably in circumstances where it would be beneficial to both the buyers and the sellers.

This market situation will differ from the De Beers group diamond market institution in that there is more than one seller here and rejection by buyers is not punished as severely.[8] It is similar enough, however, to make interesting the question of whether bundling will survive when there is more than one seller. The market situation is similar also to the market for motion pictures involving film studios (as sellers) and exhibitors (as buyers). Sale in bundles, with rejection only of bundles, would compare to the contracts used (although rarely) for film exhibition rights between studios and smaller exhibitors. Sale of individual units would be approximated more by contracts between studios and the larger, first-run theaters, which might want to exhibit only higher quality films.

[8] In the DeBeers group diamond market, once an invited buyer rejects a box, or "sight," of diamonds, that buyer is not invited to buy again and so all future profits are forgone.

B. *A Model with Four Units per Trader*

We now consider a case in which each trader has four units. The traders need not specialize in singles or bundles, so we expect more mixed experimental results. The advantage of the four-unit design, however, is that traders have more complex choices, and a richer variety of behavior may be observed. For example, a seller in the two-unit setup cannot decide which units go into a bundle and which do not. But suppose that a seller in a four-unit setup sells two bundles and that only two of the units turn out to be of high quality. Then the seller must decide whether to deliver bundles of mixed quality. Buyer rejection decisions are also more complicated when four units are being purchased.

The analysis of the four-unit case is analogous to the preceding analysis of the two-unit case; the main lessons can be found by comparing the single-unit and bundle cases only. Thus, we calculate upper and lower limits on p_1 and p_2, and then we calculate expected profits and total surplus in each case and compare them. For cost and valuation parameters to be used, sales in bundles are shown to result in a higher expected total surplus because all low-quality units are rejected when sold singly in the anticipated price range for p_1, but some low-quality units are not rejected when bundled with high-quality units. Both the methods of analysis and the intuition behind the results in this section are similar to those for the two-unit market. Consequently, the reader may wish to skip to the beginning of section III on a first reading.

1. Single-Unit Sales with N = 4

We consider for illustration one seller who sells at most four units to one buyer who can redeem at most four units. Here, all four units will be sold as singles. There are 16 possible combinations of four units each and we shall draw probabilities of outcomes from these possibilities. We choose an average cost C which exceeds V_L, so we can expect p_1 to exceed V_L as long as there is a range of prices above V_L for which both buyer and seller earn positive expected profit. When $V_L < p_1 < V_H$, all low-quality units will be rejected. The seller's expected profit for a price in this range is $2p_1 - 4C$, because only half of the units, the highs, will be accepted. Notice that this profit is just twice the corresponding expected profit in equation (1) for the two-unit case.

Now consider more specific prices, and choices of buyers and sellers from among those prices. When the seller's price is a small amount below V_H, the upper end of the (V_L, V_H) price range, the seller's expected profit is approximately $2V_H - 4C$. Now consider whether the seller would be motivated to raise p_1 above V_H. Of course only the first high-quality unit will be accepted if $V_H < p_1 < V_1$, and the best price for the seller in this higher price range is just below V_1. There will be one high-quality unit in all of the 16 possible cases except for the one case of four low-quality units, so expected profit for the seller at price V_1 would be approximately $(15/16)V_1 - 4C$. It follows from these expected profit calculations for P_1 below and above V_H that a risk-neutral seller would not prefer to raise the price above V_H if the valuations satisfy $2V_H \geq (15/16)V_1$, or equivalently if:

$$V_H \geq (15/32)V_1. \tag{9}$$

This is analogous to inequality (2) for the two-unit setup.

When $V_L < p_1 < V_H$, so only low-quality units are rejected, a straightforward probability calculation can be carried out using the 16 possible outcomes of 4 units each. Summarizing the appropriate values indicates that the expected total surplus will be

$$(15/16)V_1 + (17/16)V_H - 4C. \tag{10}$$

2. Sales in Bundles with N = 4

A comparable evaluation can be made of prices set by a seller of four units who combines units into bundles of two units each. As in the case considered above, we assume that this seller deals with one representative buyer who can redeem at most four units. A seller of four units has 16 possible combinations of high and low quality units.

Consider combinations of two units into bundles. We should expect $p_2 < V_1 + V_H$ because at higher prices all bundles will be rejected. With $C \geq V_L$, we expect that $p_2 > 2V_L$. Within these upper and lower limits on p_2 there are three price ranges of interest: (i) $V_1 + V_L < p_2 < V_1 + V_H$, (ii) $V_H + V_L < p_2 < V_1 + V_L$, and (iii) $2V_L < p_2 < V_H + V_L$. The analysis of these cases is simpler if $V_1 + V_L < 2V_H$, and the parameters used in the subsequently reported experiments will satisfy this inequality. By tracing out possible sellers' combinations of bundles, expected profits can be determined for each of the three cases:

(i) If $p_2 > V_1 + V_L$, only bundles with two high-quality units will be accepted. At such high values of p_2, the seller will put two highs into a bundle whenever possible, and the resulting seller's expected profit is $(11/16)p_2 - 4C$ because eleven of the sixteen possible combinations of four units contain at least two high-quality units.

(ii) If $V_H + V_L < p_2 < V_1 + V_L$, and if $V_1 + V_L < 2V_H$, at most one bundle will be accepted by a buyer each period. That bundle could involve one high-quality unit and one low-quality unit or two high-quality units. This observation can be used to show that the resulting seller's expected profit is $(15/16)p_2 - 4C$, because the only way the seller can fail to sell one bundle is if all four units are of low quality, an event shown in section B.1 to have probability 1/16.

(iii) If $2V_L < p_2 < V_H + V_L$, all bundles with two high-quality units and all mixed bundles with a high and a low will be accepted. By constructing mixed bundles whenever possible, the seller can maximize the probability that a bundle sold at this price will be accepted. If the sixteen possible quality combinations of four units are again enumerated, it can be shown that the resulting seller's expected profit is $(13/8)p_2 - 4C$.

3. Comparison of Single Unit and Bundle Sales with N = 4

In the experiments with four units per trader, we shall select valuation parameters so that a seller's expected profit could not be increased by raising p_2 above the level $V_H + V_L$ which determines the upper limit of prices in case (iii). The downward pressure on prices exerted by buyers should then make case (iii) the relevant case. Recall that $p_2 < V_H + V_L$ in case (iii), so only bundles with two low-quality units will be rejected and all mixed bundles will be accepted. When only bundles with two low-quality units are rejected, a straightforward probability calculation can be carried out from the 16 equally likely 4-unit outcomes. The result shows that the expected total surplus for a buyer/seller pair (the expected benefit to both buyer and seller) with bundled sales in case (iii) is

$$(15/16)V_1 + (17/16)V_H + (20/16)V_L - 4C. \qquad (11)$$

This exceeds the corresponding expected surplus in (10) for sales of single units by an amount $(20/16)V_L$ which represents the value of low-quality units that are not rejected in mixed bundles, but would of course have been rejected if sold as single units.

Although for certain parameter values the bundle can be attractive as a means of avoiding rejection of lower valued individual units, its resulting advantage comes really from the way it can moderate the effect of uncertainty. The individual unit that would be rejected by a buyer would be a unit of lower value than the buyer had intended to accept, and would resemble a "bait and switch" tactic from which the buyer should be allowed an escape. But it is not an attempt to fool a buyer when a seller delivers a bundle of, say, one high-quality and one low-quality unit; the seller is creating a reputation, since buyers can see what it is doing. The seller is contending with uncertainty that is inevitable in its situation, and doing it by reasonably satisfying the average needs of the buyer while avoiding the waste that would follow from having rejected units.[9] Also,

[9] If we consider again the practice of block booking for motion pictures, the valuations included in this example would call for one high-quality film to be worth more than a second one ($V_1 > V_H$), and both of those to be worth more than a low-quality film ($V_H > V_L$). As noted above for the two-unit case, only certain valuations of high and low quality films would cause both film distributors and exhibitors to prefer bundling over single-unit transactions. In the example here, for prices at the upper end of the ranges $V_L < p_1 \leq V_H$ and $2V_L < p_2 \leq V_H + V_L$, it can be shown from expected profit calculations that bundling will be preferred by both buyers and sellers if a low-quality unit

(footnote cont'd)

bundling that reduces the probability of rejection makes the seller's profit less variable, and risk-averse sellers may prefer bundling for this reason.

III. Experimental Design and Method

We shall now examine this motive to bundle in an experimental setting with equal numbers of buyers and sellers. Two qualities of units will be identified as RED (low) and BLACK (high). The sellers of these units are required to make contracts before knowing which of the units that they have produced are REDs and which are BLACKs. As a consequence of this exogenous production uncertainty some buyers may not be satisfied with their purchases when they finally are able to inspect them. Buyers can reject a delivery but must reject the entire contract if more than one unit is involved. Once rejected, a unit cannot be resold; thus there are no inventories. Valuations can be chosen to satisfy the constraints noted above, so if they are allowed to do it sellers will create bundles in order to make rejection less likely.

The specific goods we employ are playing cards, which individually are red or black with 0.5 probability. The nominal values at which buyers can redeem their cards, which will motivate their purchases, can be chosen to make one type of card (BLACK) typically more valuable in that it generates more buyers' surplus than the other type (RED). The first BLACK that is redeemed brings more than others ($V_1 > V_H$), and other BLACKs bring more than any RED ($V_H > V_L$). Sellers incur a cost per period and receive only a limited number of cards each period. They learn how many are RED or BLACK only after they have made contracts with buyers. Sales contracts may be for either single units or for two-unit bundles.

Subjects for experiments according to this design were recruited from principles-of-economics classes at the University of Virginia "to earn money in a research project." The subjects were assigned to buyer and seller positions located far enough apart as to be unable to see one another's records. Instructions, which are contained in the Appendix, were read to the subjects. The Appendix also includes forms of records that were used by subjects. Before beginning the experiment, the subjects were told that the experiment would last from 2 to 2½ hours. Subjects were not permitted to speak to each other during the experiment, and they were paid in private immediately afterwards. Specific parameters for the two-unit and four-unit experiments are given in the two subsections which follow.

is worth at least a fourth, roughly, of a second high-quality unit. Something like this requirement might have been met at smaller theaters, where block booking was more common, but not at larger urban theaters where block booking was not used, perhaps because in the larger theaters a low-quality film could have had very little value relative to a high-quality film.

A. *Two Units per Trader*

It follows from a comparison of equations (4) and (7) that for risk-neutral buyers, bundles will be preferred if $p_1 > (3/4)p_2 - (1/2)V_L$. From (1) and (6) it is clear that for risk-neutral sellers bundles will be preferred if $p_1 < (3/4)p_2$. It is apparent that a higher value of V_L, the value of a RED, will increase the size of the region in the price space in which bundling is attractive to both buyers and sellers. We chose the following parameters:

$$V_1 = \$8.10$$
$$V_H = \$6.10$$
$$V_L = \$2.00 = C.$$

As indicated in section II, we anticipate the prices for single units to be in the range from C to V_H: \$2.00-\$6.10, and we anticipate the prices for bundles to be in the range from 2C to $V_L + V_H$: \$4.00-\$8.10. These ranges were derived from the observation that sellers would resist selling below cost and that for parameters chosen sellers would not profit from raising prices above the upper limits of these ranges, even if they could. But buyers are willing to pay any amount up to $V_1 + V_H$ (\$14.20) for a bundle and any amount up to V_1 (\$8.10) for a single, because they can always reject contracts with low-quality units. Thus the actual prices are indeterminate, although there is an efficiency motivation for bundling. Our previous analysis of seller bundling and buyer rejection decisions was for the case of one buyer and one seller. We use these calculations for buyer-seller pairs to provide predictions for a market with an equal number of buyers and sellers.

If gains for both buyer and seller together are greater for bundles, bundles also are more efficient than singles, and that will be true for the price ranges to be expected with these parameters. Substituting these parameters into (5), the sum of buyer's and seller's expected surplus is \$3.60 for single units (at prices for units between \$2.00 and \$6.10). Substituting the parameters into (8) yields expected total surplus of \$4.60 for bundles (at prices for bundles between \$4.00 and \$8.10), the higher value for bundles indicating that bundles are more efficient.

B. *Four Units per Trader*

The specific redemption values used for the four-unit case are:

$$V_1 = \$6.10$$
$$V_H = \$4.10$$
$$V_L = \$1.00 = C.$$

We anticipate the prices for single units to be in the range from C to V_H:

\$1.00-\$4.10, and we anticipate the prices for bundles to be in the range from 2C to $V_L + V_H$: \$2.00-\$5.10.

Substituting the above parameters into (10), we find the sum of buyer's and seller's expected profit is \$5.98 for single units (at prices for units between \$1.00 and \$4.10). Substituting into (11) we find combined expected profit is \$7.21 for bundles (at prices for bundles between \$2.00 and \$5.10). Thus bundles can result in significant efficiency gains for the parameters we have selected.

IV. Experimental Results

This section reports the results of two experiments with two units per trader and of two experiments with four units per trader.[10] An earlier pilot experiment is also discussed. In each experiment, simultaneous oral double auctions for both single-unit and bundle contracts were conducted for at least six periods. The four-unit experiments were conducted first, and the subjects in those experiments had no previous experience in oral double auctions. The subjects in the two-unit experiments had participated previously in oral double auctions, and all but one of these subjects had been a trader in one of the four-unit experiments. We begin with results of the four-unit experiments, since they were carried out first.

A. *Results for the Design with Four Units per Seller*

Recall that the parameters for the four-unit design discussed in the previous section are: V_1 = \$6.10, V_H = \$4.10, and C = V_L = \$1.00. The anticipated range of contract prices is \$1.00-\$4.10 for singles and \$2.00-\$5.10 for bundles.

Each of the four-unit experiments, Experiments I and II, was operated for exactly six market periods. Contracts for single units and bundles were permitted and both were observed in use every period. The average price per market period for a single or bundle is reported in Table 1, along with the number of units that were sold as singles or in bundles in each period. Figures 1 and 2 trace prices by contract within each period for Experiments I and II.

In Experiment I an interesting tendency is clear. Both Table 1 and Figure 1 show that bundle prices lie at the upper end of the bundle price range (\$2.00-\$5.10) and that many of the single-unit prices lie near the bottom of the single-unit price range (\$1.00-\$4.10). Also in Experiment I there is a slight tendency toward greater use of bundles, as evidenced in the last two market periods. In the first two market periods, single units traded at prices better than in bundles (two single units sold for a combined average value of \$5.50

[10] A detailed Data Appendix which contains all information on each contract (price, type of contract, identity of traders by subject number, realized quality of the unit(s), and buyer rejection decision) is available from the authors on request.

in period 1, while a bundle sold for an average of $4.90). As trading continued, the single units declined to less than their value in bundles (two single units were worth $4.10 together in period 6, while a bundle was worth $4.65). This price trend, together with the slight increase in the use of bundle contracts, suggests that bundles became more attractive to traders.

Bundling is more prevalent in Experiment II with two-thirds of all units being sold in bundles. The average prices and sellers' profits were much lower than was the case in Experiment I; bundle prices were lower than in Experiment I, and almost all of the single-unit prices were near the bottom of the anticipated price range ($1.00-$4.10). The tendency for a bundle to sell for more than twice the price of a single was apparent in both experiments but was stronger in the second. In both, sellers seemed to realize the advantage of delivering mixed bundles containing one RED and one BLACK. In the second experiment, every seller delivered mixed bundles each time the seller's cards permitted a bundle to be mixed.

An earlier pilot experiment was like the two experiments just reported with the exception that redemption values of REDs were $.50 higher—at $1.50—greater than the unit cost of $1.00. In addition, BLACKs after the second were valued exactly the same as REDs. There were four buyers and four sellers in the market. Although a price for single units that is just below V_H would also be optimal for a risk-neutral seller in that case based on the inequalities analyzed above, a price slightly below V_L actually turned out to be chosen most often. At this price ($1.40) every unit was always accepted, so the outcome was efficient. The price of $1.40 exceeded the sellers' unit cost of $1.00, so sellers' profits were positive and stable, although low. Sellers apparently were sufficiently risk averse to prefer that outcome.

At a per unit price of $1.40 with these parameters, sellers would be indifferent between bundles and single units. Bundles offer no efficiency advantage when every unit is accepted without them. Nevertheless, overall exactly half of the units were sold as single units and half were sold as bundles, although a slight trend toward single units developed over the four periods studied. The lesson we draw is that we must expect some bundles to be sold even if bundles offer no special benefit over single units, and the results of Experiments I and II should be judged in this context.

B. *Results for the Design with Two Units per Seller*

To force a clearer choice between single units and bundles, experiments were conducted in which subjects bought and sold only two units. This design forces each buyer and seller to deal in either single units or bundles and thus it allows a sharp test for the attractiveness of the bundle. Also, because of the greater simplicity of the experimental situation with only two units per trader, less time is needed for instruction and decision making, which permits more market periods to be conducted each session. The additional market periods enabled us to examine the role of the money-back guarantee in one experimental session.

The parameters for the two-unit design discussed in section III are: V_1 = \$8.10, V_H = \$6.10, and C = V_L = \$2.00. With those parameters, the anticipated price range for single-unit contracts is between C and V_H: \$2.00-\$6.10, and the anticipated price range for bundles is between 2C and $V_L + V_H$: \$4.00-\$8.10. As before, although efficient contracts can be described for particular parameters in this market,the price level is indeterminate. The prices were so low in Experiment III that sellers were losing money, and after period 7 they were each given an additional payment of \$10 to sustain a profit-seeking motivation. Earnings of sellers were higher in Experiment IV and no such added payment was needed.

The actual sequences of contract prices for these experiments are shown in Figures 3 and 4, and summary results are presented in Tables 2 and 3. In Experiment III, after 6 trials buyers and sellers were allowed to abandon the guarantee which results from the rejection option. The results of the first 6 trials for Experiment III are presented in Table 2, along with 11 trials from Experiment IV. A strong tendency is evident in both experimental sessions to rely eventually on bundles. After 3 periods in Experiment III, no further transaction occurred in single units; all transactions were in bundles. A more gradual movement to bundles occurred in Experiment IV but the trend is strong. The transition to bundles in this last experiment is interesting. As can be seen in Figure 4, no bundle was sold until period 5. Seller 4, who sold this bundle, received a RED and a BLACK, and the resulting mixed bundle was accepted by the buyer. None of the sellers sold bundles in the following period, period 6, and seller 4 received two REDs which were both rejected. The very first contract in period 7 was a bundle sold by seller 4, and bundles dominated in transactions after that.

Periods 7 through 10 in Experiment III were carried out with the guarantee as an option rather than a required condition of transactions. The instructions which were given to make guarantees optional are included in the Appendix. For a contract without a guarantee, the buyer does not have the option to inspect and reject, so buyers absorb the uncertainty for such a contract. With optional guarantees we essentially conducted four oral double auctions simultaneously; the auctions were for single units and bundles, each either with or without guarantees. Results are shown in Table 3 and Figure 3. Singles were used only in period 8, when one single was transferred with a guarantee and one was transferred without. All other sales involved bundles. There was a slight trend toward bundles with no guarantee, but not enough periods were carried out to determine whether the guarantee would be abandoned. When units are sold without a guarantee there is no possibility of unaccepted units and so the market outcome will be efficient whether bundles are involved or not. Further work is needed to determine whether buyers so prefer the guarantee that bundles are needed to reach an efficient outcome.

V. Conclusion

Transactions must often be arranged for goods of uncertain quality. When quality is uncertain and buyers have a right of inspection and rejection (a "money-back" guarantee), we show that bundles of units can be the most efficient basis for transactions. The advantage of the bundle is that an average quality can be provided more reliably for a bundle than for an individual unit. This efficiency advantage for bundling will arise only for certain conditions on costs and the valuations of unit qualities, but it can arise in a competitive market.

A preference for transacting in bundles was demonstrated in an experimental double oral auction market which was designed to satisfy conditions that make bundling efficient. A clear choice was required in a pair of experiments with three buyers and three sellers, where each buyer and seller handled only two units per period and so was forced either to bundle or not. There, bundles came to dominate as the basis for all transactions. In experiments where buyers and sellers each had four units to handle there was a slight preference for bundling. Although single-unit transactions did survive, they brought a relatively lower unit price, and trends had favored the use of bundles when the experiments were terminated. We conclude that the practice of bundling may be efficient in some circumstances, and that competition will then encourage its use.

The conditions we identify as appropriate for bundling to offer advantages seem roughly to have been satisfied in transactions where motion picture studios negotiated with smaller exhibitors, which is also where the so-called "block booking" practice was used. Block booking essentially is contracting for motion pictures in bundles, and it is now unlawful to use in distributing films either to theaters or to television. The role of bundling in this and other areas may deserve reconsideration in light of the efficiencies it has been shown to offer in competitive conditions.

Table 1

RESULTS FOR EXPERIMENTS WITH N = 4

EXPERIMENT I

MARKET PERIOD	AVERAGE PRICE		UNITS SOLD	
	SINGLE	BUNDLE	SINGLE	BUNDLE
1	$ 2.75	$ 4.90	6	6
2	2.68	5.03	6	6
3	2.32	5.00	6	6
4	1.82	4.97	6	6
5	1.75	4.82	4	8
6	2.05	4.65	4	8

Table 1 (cont'd)

EXPERIMENT II

Market Period	Average Price: Single	Average Price: Bundle	Units Sold: Single	Units Sold: Bundle
1	$ 1.13	$ 2.20	4	8
2	1.40	3.10	2	10
3	1.27	3.60	6	6
4	1.23	3.49	4	8
5	1.17	3.38	4	8
6	1.23	3.28	4	8

Table 2

RESULTS FOR EXPERIMENTS WITH N = 2

EXPERIMENT III (FIRST SIX PERIODS)

Market Period	Average Price: Single	Average Price: Bundle	Units Sold: Single	Units Sold: Bundle
1	$ 2.62	$ 5.30	4	2
2	--	5.23	0	6
3	2.35	5.10	2	4
4	--	5.07	0	6
5	--	5.10	0	6
6	--	4.93	0	6

EXPERIMENT IV

Market Period	Average Price: Single	Average Price: Bundle	Units Sold: Single	Units Sold: Bundle
1	$ 3.25	--	6	0
2	3.32	--	5	0
3	3.37	--	6	0
4	3.40	--	6	0
5	3.52	$ 6.10	4	2
6	3.40	--	5	0
7	2.90	6.00	2	4
8	2.40	6.00	1	4
9	3.60	5.95	1	4
10	--	5.97	0	6
11	--	5.93	0	6

Table 3

RESULTS WITH OPTIONAL GUARANTEE

EXPERIMENT III (PERIODS 7-10)

	With Guarantee			
	Singles		Bundles	
Market Period	Average Price	Units Sold	Average Price	Units Sold
7	--	0	$ 4.90	6
8	$ 2.30	1	4.80	2
9	--	0	4.75	4
10	--	0	5.00	2

	With No Guarantee			
	Singles		Bundles	
Market Period	Average Price	Units Sold	Average Price	Units Sold
7	--	0	--	0
8	$ 2.00	1	$ 4.00	2
9	--	0	4.00	2
10	--	0	3.95	4

Figure 1

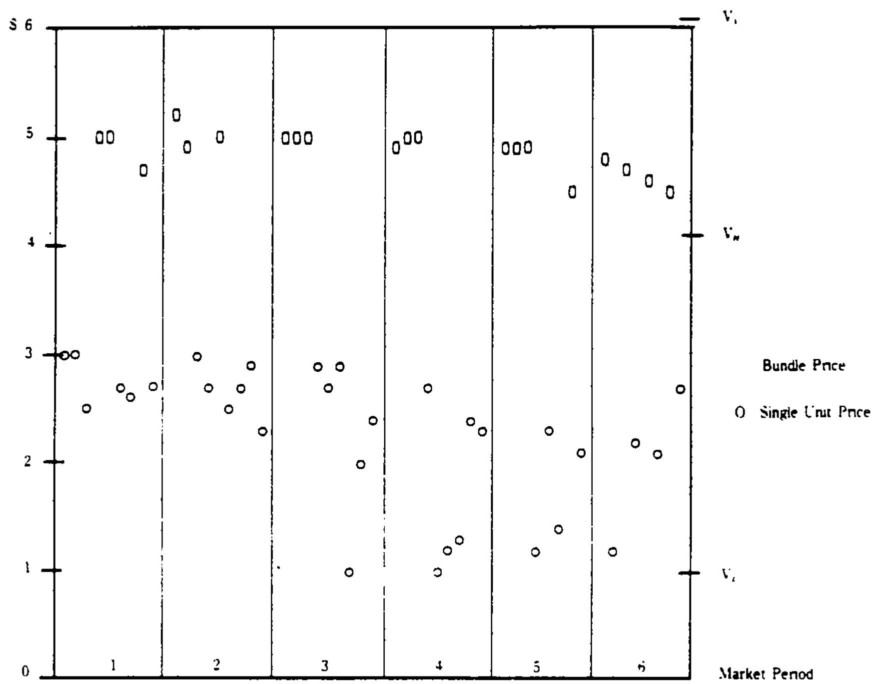

$ 6
5
4
3
2
1
0
1
2
3
4
5
6
Market Period
V
Bundle Price
O Single Unit Price

Figure 2

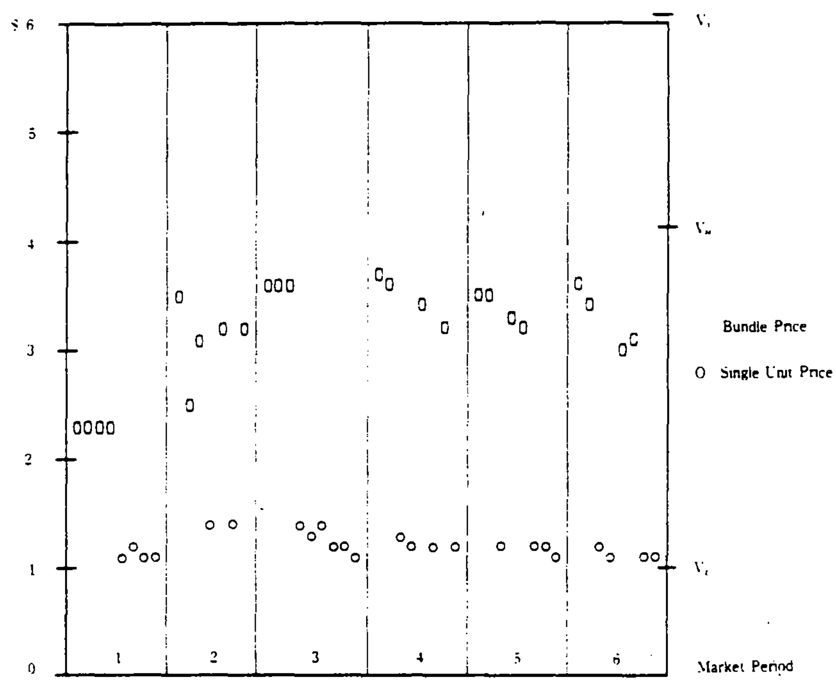

$ 6
5
4
3
2
1
0
1
2
3
4
5
6
Bundle Price
Single Unit Price
Market Period

Figure 3

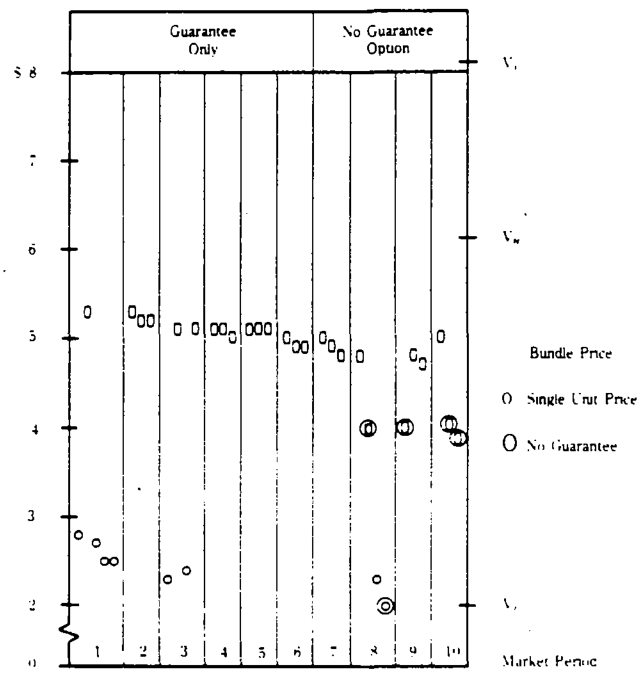
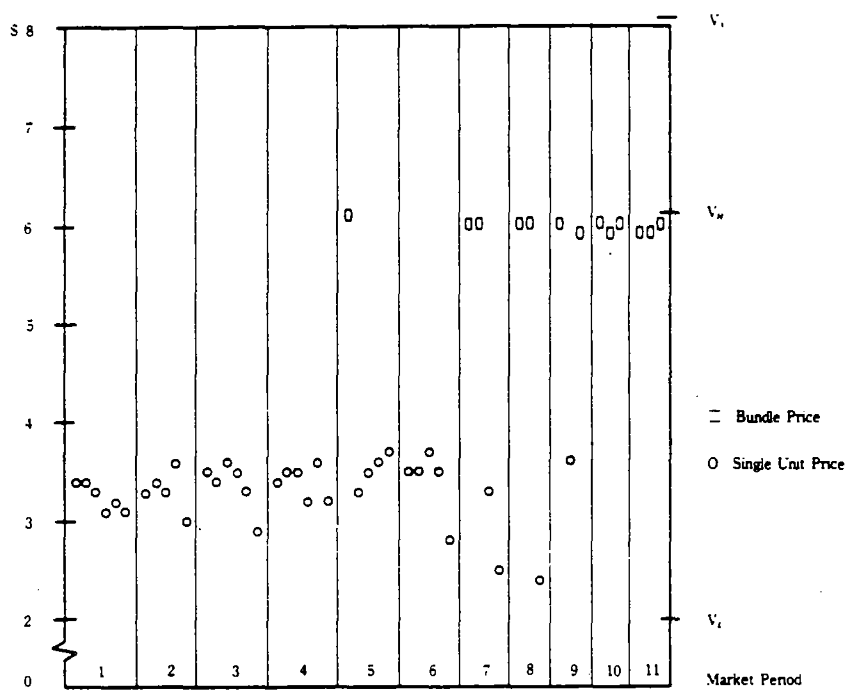

Guarantee Only
No Guarantee Option
$ 8
7
6
5
4
3
2
0
Bundle Price
Single Unit Price
No Guarantee
Market Period

Figure 4


$ 8
7
6
5
4
3
2
0
1
2
3
4
5
6
7
8
9
10
11
V_s
V_H
V_L
Bundle Price
O Single Unit Price
Market Period

Appendix

Instructions for Experiments with $N = 2$

General Instructions

This is an experiment in the economics of market decision making. Various organizations have provided funds for this research. If you follow the instructions carefully and make good decisions you might earn a considerable amount of money, which will be paid to you in cash at the end of the experiment. In this experiment we are going to conduct a market in which some of you will be buyers and some of you will be sellers in a sequence of "market days," or trading periods.

Attached to these Instructions you will find some sheets, labeled Record Sheet of Buyer or Record Sheet of Seller, depending on whether you are a buyer or a seller. Your Record Sheet contains information that determines the value to you of the decisions you might make. *You are not to reveal this information to anyone*. It is your own private information.

Specific Instructions to Buyers

During each market period you are free to contract for the purchase of up to 2 units from any seller or sellers. Units in the market will turn out to be of two kinds, called RED or BLACK. The units will be represented by ordinary playing cards. BLACKs tend to be more valuable than REDs. For the first BLACK unit that you redeem ***during a trading period*** you will receive the amount listed on Page One of your RECORD SHEET marked *lst BLACK unit redemption value*; if you redeem a second BLACK unit you will receive the additional amount marked *2nd BLACK unit redemption value*. The redemption values of the REDs you may purchase are similarly calculated from the information given on Page One of your Record Sheet.

The profits from each purchase (which are yours to keep) are computed by taking the difference between the redemption value and the purchase price of the unit bought. That is,

$$\text{your earnings} = \text{redemption value} - \text{purchase price}$$

At the time you make a purchase, neither you nor the seller will know whether the unit is RED or BLACK. Sellers will receive cards from a newly shuffled deck that has half RED cards and half BLACK cards, so each unit has a 50-50 chance of being RED or BLACK.

If you are not satisfied with what you receive from a seller you may reject the contract. A unit received counts among the two you are allowed, regardless whether you accept it or reject it. The contracts and deliveries of sellers will be made public, so you and other buyers will know each seller's deliveries. You must realize that since sellers do not know whether units are RED or BLACK when they make offers, they may sometimes have to deliver REDs which are less valuable than BLACKs. REDs or BLACKs also cannot be carried over from one period to the next; they have a "life" of only one period.

The values to you of REDs and BLACKs depend on the numbers of each you have to redeem. The example in the table below shows hypothetical redemption values of RED and BLACK units.

ILLUSTRATIVE REDEMPTION VALUES OF RED AND BLACK UNITS

	RED	BLACK
Redemption Value of 1st unit	$ 70.00	$ 200.00
Redemption Value of 2nd unit	$ 70.00	$ 150.00

To illustrate, if you had purchased two units at $50 each and when they were delivered you learned one was BLACK and one was RED, you would want to accept both the BLACK (profit would be $200-$50 = $150)and the RED (profit would be $70-$50 = $20). If you had purchased two units at $100 each and when they were delivered one was BLACK and one was RED, you would want to accept the BLACK (profit would be $200-$100 = $100) and reject the RED (profit would be $70-$100 = -$30). If you had purchased two units at $175 each and when they were delivered both were BLACK you would accept one (profit would be $200-$175 = $25) and reject the other (profit would be $150-$175 = -$25). The numbers used here are for illustrative purposes only (and they yielded profits per period far greater than you can expect). In the actual market periods all prices must be expressed in at least $.10 units.

Now please turn to the blanks in the table on Page Two of your RECORD SHEET. A new Page Two will be used each period. At the time of a purchase record the Contract Number, Units, Price and Seller Number on your RECORD SHEET. When delivery is received you can record the contents. If you accept the delivery you may also calculate the Redemption Value of the contract. At the end of the period, totals can be obtained for prices paid and for RED and BLACK units accepted and redeemed. The total obtained from redeeming RED units can be placed under the total obtained from redeeming BLACK units, as the short arrow indicates, and addition of the two sums will then indicate the total Redemption Value of all units. The total for Prices of Accepted Contracts can then be placed below the Total Redemption Value, as the arrow indicates, and subtraction will yield profit for the period.

Thus far we have considered bids and offers only of individual units. It will also be possible, however, for buyers to offer to buy two units as a package for a particular price, rather than one unit. Sellers also will have the option of offering two units for sale at a specified price. Of course it is possible also to bid on or offer a single unit for a single-unit price. The buyer's right of rejection always will apply to the total package. That is, a two-unit package must be accepted or rejected in total; parts of the package cannot be accepted or rejected separately. In recording any transaction involving a pair of units, you may indicate it is a pair by indicating "2" in the Units column and by reporting the delivery contents for the pair. It is advisable to use two rows if two units of the same color are delivered; that way their redemption values also may be entered separately, which can help to avoid making any mistake in determining total redemption values.

Specific Instructions to Sellers

During each market period you are free to make contracts to sell to any buyer or buyers as many units as you might want as long as the total does not exceed 2 units. Units in the market will turn out to be of two kinds, called RED and BLACK. Units will be represented by ordinary playing cards. When you offer to sell units during a period, you will not be certain how many of your units will be RED and how many will be BLACK.

Whether a unit is RED or BLACK will be determined randomly by the drawing of cards from a well shuffled deck of playing cards. Because it will be either a red or a black card, each of your units will have a 50-50 chance of being RED or BLACK. BLACK units will tend to be worth more to buyers than RED units, but RED units always have some value to buyers. After you are informed how many BLACK and RED units you have, you will make up deliveries to buyers who accepted your offers. Buyers may inspect the units you deliver and reject them if they wish. If your price per unit exceeds the value to the buyer of a unit, the buyer may reject it. Each separate transaction you make is subject to this possibility of rejection. The color of units actually delivered by each seller will be made public to all buyers at the end of the market period. At the end of each period sellers will also be informed of price offers, acceptances, deliveries and rejections for all sellers.

In each period, you will have 2 units to sell at a cost per unit which is listed at the top of your Record Sheet of Seller as *cost per unit*. You incur this cost whether a unit is accepted or not.

The profits on sales, which will be yours to keep, can be calculated by adding up the revenues you obtain for deliveries actually accepted, less your costs. A specific example will help illustrate the contracting process. Consider the offers indicated on the offer sheet below. (The numbers used are for illustrative purposes only.)

(Cost per unit is $80.00)

Contract No.	Units	Price	Buyer No.	Accepted	Revenue
1	1	$ 200.00	3	Yes	$ 200.00
2	1	$ 200.00	1	No	0
(*) Total revenue from accepted contracts:					$ 200.00
(**) Total cost:					- 160.00
(***) Profit [row (*) - row (**)]:					$ 40.00

In this example, your contract number 1 was to sell 1 unit to buyer 3 for $200, and your contract number 2 was to sell 1 unit to buyer 1 for $200. Buyer number 1 rejected your delivery, however, so you received no revenue on your second offer. The profit on your first offer would be revenue of $200 less cost of $80, or $120. On the second offer you would receive no revenue but have a cost of $80, so you would lose $80. Total profit would be $40. If the price of both units had been $150 rather than $200 and one unit still was rejected a loss would have been incurred, since revenue would have been $150 and cost $160. Notice that if cost per unit is $80 and only half the units are accepted, price per unit will have to exceed $160 to be profitable in this example.

The numbers above were used for illustrative purposes only (and they yielded profits per period greater than you can expect). In the actual experiment you must set prices in even $.10 increments.

Thus far we have considered bids and offers only of individual units. Sellers also will have the option of offering two units for sale at one price rather than individually, and buyers may offer to buy two units for one price. A seller may sell in one period two single units for a contract price in each case, or a pair of units for a two-unit price. The buyer's right of inspection and rejection will always apply to the entire package, however. That is, a two-unit package must be accepted or rejected in total; part of the package cannot be accepted or rejected separately. In recording a contract involving a pair of units you may indicate it is a pair by indicating "2" in the Units column and by reporting the delivery contents and the revenue for the pair.

Market Organization

The market will be conducted in a series of market periods. Each period will last for at most 4 minutes. Any buyer is free at any time during the period to raise his or her hand and make a verbal bid to buy either a single unit or a two-unit bundle at a specified price. Any seller is free at any time during the period to raise his or her hand and make a verbal offer to sell either a single unit or a two-unit bundle at a specified price. The bid (offer) should specify the Buyer number (Seller number), the number of units, and the bid (offer) price. The bid (offer) should be higher (lower) than the outstanding bid (offer) for the number of units that the bid (offer) pertains to, should such an outstanding bid (offer) exist. Any seller is free at any time to accept or not accept the bid of any buyer, and any buyer is free at any time to accept or not accept the offer of any seller. If a bid or offer is accepted, a contract has been closed for the number of units specified in the bid or offer, and the buyer and seller will record the contract price to be included in their earnings. Any ties in bids or acceptances will be resolved by random choice of buyer or seller. Except for the bids, offers, and their acceptance, you are not to speak to any other subject. There are likely to be many bids and offers that are not accepted, but you are free to keep trying.

After the market period ends, sellers will learn exactly what playing cards

(RED, BLACK) they have available to sell and will make up envelopes for accepted contracts. Each envelope will contain the playing cards (RED, BLACK) the seller is delivering against that contract. Delivery occurs when a seller transfers, via the experimenter, a contract to a buyer. The buyer then either accepts or rejects each contract (envelope). Buyers and sellers can record results for that period, and then the next period will begin. A large number of periods will be conducted.

You are not to speak to other participants in the experimental market. You will begin with a balance of $______, which is part of your payment for participating. You are free to make as much profit as you can.

Final Observations

1. Each individual has a large folder. All papers, instructions, records, etc. should be put into this folder. Leave the folder with us before leaving. Take nothing home with you.

2. We are able to advise you a little on making money. First, you should remember that small sums add up over many trades and a long period of time. Secondly, your earnings may not be steady. You may have some good periods and some bad periods.

3. We have noted that you are not to speak except as necessary to participate in contracts. You must certainly not mention anything about activities which might involve you and other participants after the experiment (*i.e.*, no physical threats, deals to split up afterwards, or leading questions).

4. Each individual will be paid in private. Your earnings are strictly your own business.

Added Instruction for Introducing No-Guarantee Condition

Thus far all packages (single units or bundles) have been transacted essentially with a money-back guarantee; buyers have been able to inspect the type of units they receive and reject any contract they wish. In future periods the option of forming such guaranteed contracts will continue to be available. But in addition it will be possible to form contracts without the guarantee. Under a no-guarantee contract an exchange takes place between buyer and seller at the agreed upon price, and no right of rejection is allowed the buyer. Separate bid-and-offer records will henceforth be maintained for Guarantee Singles, Guarantee Bundles, No-Guarantee Singles and No-Guarantee Bundles.

RECORD SHEET OF BUYER ________

Page 1—Redemption Values

RED	
1st RED unit redemption value	$ ________
2nd RED unit redemption value	$ ________
BLACK	
1st BLACK unit redemption value	$ ________
2nd BLACK unit redemption value	$ ________

Period ________

RECORD SHEET OF BUYER ________

Page 2—Contracts

(A total of only *two* units may be redeemed each period.)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Delivery Contents			Redemption Value	
Contract No.	Units	Price	Seller No.	RED	BLACK	Prices of Accepted Contracts	RED	BLACK
________	________	________	________	________	________	________	________	________
________	________	________	________	________	________	________	________	________
					Totals	________	________	________
								________
					Total Redemption Value			________
								________
					Total Profit for Period			________
					Cumulative Profit			________

Seller ________ Period ________

RECORD SHEET OF SELLER

Cost per unit $ ________

(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Delivery Contents		
Contract No.	Units	Price	Buyer No.	RED	BLACK	Revenue from Accepted Contracts
________	________	________	________	________	________	________
________	________	________	________	________	________	________

(*)	Total Revenue from Sales of Accepted Contracts	________
(**)	Total Cost of the Two Units	________
(**)	Total Profit for Period [row (*) - row (**)]	________
(***)	Cumulative Profit	________

Product Quality, Consumer Information and "Lemons" in Experimental Markets

Michael Lynch, *Federal Trade Commission*
Ross M. Miller, *Boston University*
Charles R. Plott, *California Institute of Technology*
Russell Porter, *Federal Trade Commission*

I. Introduction

This paper reports on the behavior of experimental markets wherein buyers were ignorant (unless truthfully informed by sellers) of the quality of the product purchased. True quality of the product was learned only after the sale. Sellers chose quality or "grade" and higher quality was more costly to produce. Our experimental markets were characterised by asymmetric information possessed by buyers and sellers who traded a pure "experience" good whose quality was endogenously determined.

Some theories predict that such markets will ultimately consist only of the lowest quality goods, that is, only "lemons" will be traded. Bad quality drives out good, in spite of the fact buyers are willing to pay the added cost of higher quality. This "lemons" outcome is clearly inefficient, because both buyers and sellers could be better off if higher quality goods were also produced. The inefficiency is due to the failure to effectively share the sellers' information on quality with the buyers. Thus the "lemons" equilibrium is one type of informational market failure, perhaps the simplest one that has been analytically modeled.

The "lemons" model itself is of more than academic interest to the FTC. The FTC staff (1978) explicitly referred to and used this model to argue the merits or lack thereof of (most appropriately, at least in an etymological sense) the proposed Used Car Rule. The model has also been explicitly invoked in various housing warranty cases. Less explicit but quite conscious use of the main theme of this model appeared in the Staff Report on Life Insurance (1979).

It is also of interest as a special case of a general problem. Under what conditions may market inefficiency be caused by a lack of consumer information or by seller provided misinformation? Alternatively, under what conditions are regulations penalizing deceptive or misleading seller claims, or government mandated disclosures or standards likely to improve market performance? This is clearly a central question for consumer protection policy—both for formulating the basic principles of policy and for allocating resources efficiently to implement them.

Some general answers have been given, and at this general level there seems to be considerable agreement, even among those who often disagree on specific policy issues. Both Posner (1973, 1979) and Pitofsky (1979), for example, agree that disclosure problems are most likely to arise for products or services that either have important "hidden characteristics" or that are infrequently purchased and expensive. A hidden or a "credence" characteristic is one that would not become apparent in normal use or consumption (*e.g.*, cholesterol content of butter or margarine). In either case, sellers may have little incentive to disclose a negative characteristic because failure to do so will not harm future sales, whereas disclosure may hurt present sales. Both also appear to agree that there is little reason to think that disclosure problems will ever arise for frequently purchased experience goods.

Agreement at this general level does not imply agreement on any specific issue, as a perusal of the Posner-Pitofsky exchange will quickly show. "Natural" markets, like that for used cars or new home warranties or new life insurance policies, are extremely complicated. Firms may pursue very diverse marketing strategies in different market segments. Consumers in different segments may have very different information on product quality. It is often very difficult to determine whether there has been any "failure" in such markets and even more difficult to test alternative theories of the cause of failure. Reputation effects could be very powerful, but extremely difficult to measure. The artificial experimental markets we have created have fewer complications and allow us to unambiguously identify "failure" when it occurs.

The study was designed with one primary objective and several secondary objectives. The primary objective was to investigate circumstances in which the "lemons" phenomena will arise in markets. The design of the markets, the market organization and institutions, were guided by models found in economic literature. Thus, the pursuit of the primary objective implicitly involved adding operational content to various theories as well as tests of the reliability of the resulting models. The secondary design objective, predicated upon the assumption that the first objective would be successfully attained, was to check the sensitivity of the "lemons" phenomena to parameters and regulations that some theories and policy arguments suggest will eliminate the phenomena and increase market efficiency.

The broad questions posed and answered by the research are as follows: (1) Can "lemons" problems occur in markets? Our answer is "yes" in the sense that we have designed markets in which it can be observed with substantial reliability. (2) Are express warranties an effective remedy if the lemons phenomena are viewed as a market failure? In our experimental markets express warranties and truthful advertising are the same thing. Regardless of the interpretation, the answer in the markets we studied is "yes." (3) Where quality is easily ascertainable after purchases (experience goods) will sellers form reputations that guarantee efficient market performance? Our answer is "not necessarily." We are then able to isolate some conditions that will help the reputation formation process. The "hidden" characteristics problem was not in the current experimental design, except in a very minor way. In studying and answering these three broad questions we pose several specific hypotheses that are suggested by both the data and existing ideas about the evolution of informational efficiency in markets.

The following section is an outline of the experimental design, procedures, and parameters. The third section is a discussion of models that might reasonably be expected to apply to the setting. The fourth section is a discussion of results, and the final section is a summary.

II. The Market and Regulatory Environments, Experimental Design, and Procudures

A. *Parameters*

A total of *twenty-one* markets were conducted plus some pilot experiments. Participants were students at Boston University (BU) California Institute of Technology (Caltech), and Pasadena City College (PCC). Some of these participants were involved in several markets as a control for experience.

All markets proceeded as a series of market days or trading periods. The number of periods was unknown to participants, but, because they knew roughly the maximum time of the experiment (three hours), they had some idea of when the last periods were approaching. Sellers remained sellers throughout an experiment and buyers remained buyers.

Sellers could supply units of grade Super or Regular. Each seller was limited to a total supply of two units per period. The units could be any combination of grades possible as long as each seller supplied a total of two or less units. Thus, the seller could sell two Rs, two Ss, one of each, one unit of some type, or nothing. The fact that Supers were more costly to sellers than were Regulars was public information. Both Supers and Regulars were supplied at constant marginal cost up to the limit of two units in total. For "high cost" experiments, which are all but selected periods of experiments 19 and 21, the (constant) marginal cost of Supers was 100 francs (one dollar) more than the (constant) marginal cost of Regulars. In the low-cost experiments, this difference in marginal cost was reduced to either 20 or 25 francs.

Buyers' redemption value of Supers was more than Regulars and this was public information. The redemption value for buyers is in Figure 1. As can be seen, the marginal valuation of a Super always dominates the marginal valuation of a Regular. Thus, given a choice of a Super or Regular, a buyer would always prefer a Super until a limit of three Supers is attained and the marginal valuation falls to zero. All buyers had identical redemption schedules.

For a typical experiment with eight buyers and six sellers the market demand and supply are presented in Figure 1. The values are in an experimental currency called "francs" that have a dollar conversion factor. As can be seen, the market supply is horizontal for twelve units and then becomes vertical.

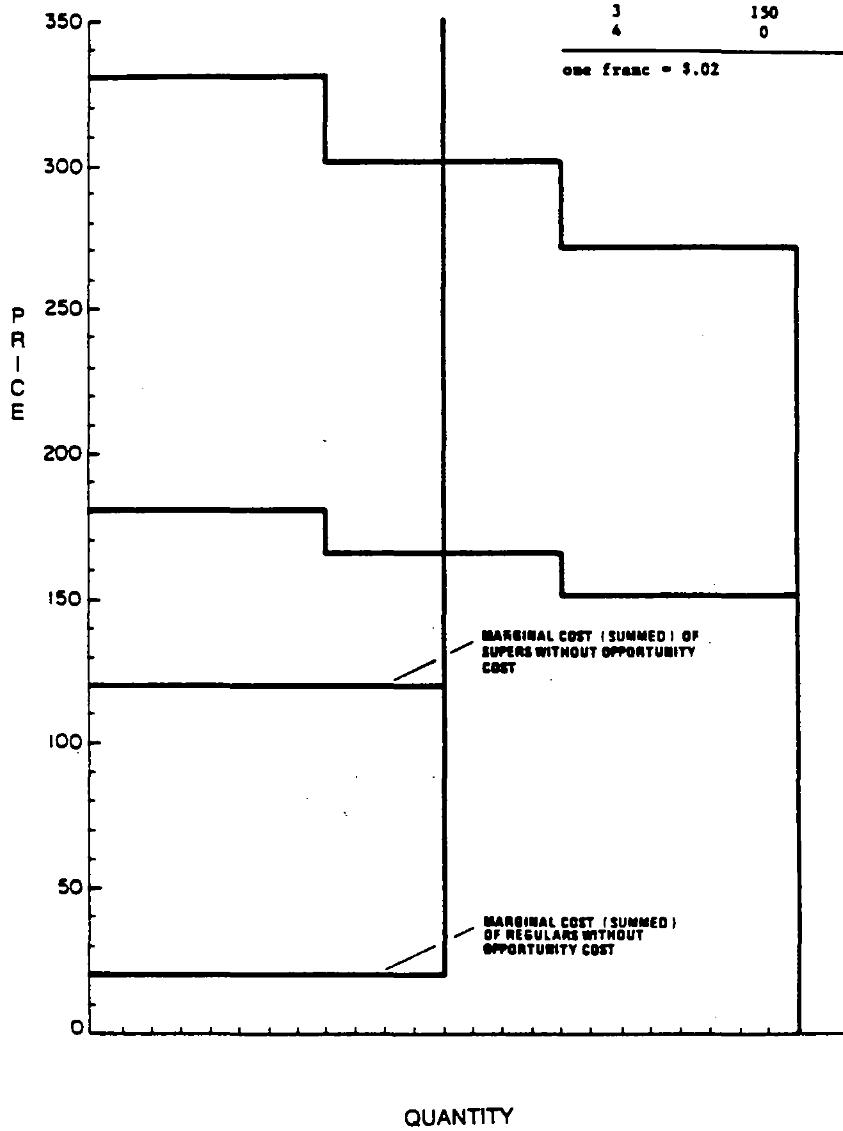
All transactions were in an experimental medium of exchange called francs. Francs could be converted to dollars at a predetermined rate known only to each individual. Prior to some markets, subjects were told that the dollar per franc conversion rate might be scaled upward after the experiment. In early experiments in which little was known about behavior and parameters, the value

Figure 1

BUYER REDEMPTION VALUES IN FRANCS

Unit Number	Regular	Super
1	180	330
2	165	300
3	150	270
4	0	0

one franc = $.02



of francs was increased so that on average participants earned about \$5 to \$7 per hour. This was thought necessary in order that the experienced subjects would be willing to participate again.

In addition to profits earned from purchases and sales, buyers were given a bonus of 50 francs each period and an unexpected one-time endowment of 200 francs at the end of the first period. Early pilot experiments demonstrated a potential problem of credibility and control, which the bonus helped to eliminate. During the first period inexperienced buyers would pay high prices for units on the expectation that sellers would deliver Supers. When Regulars were actually delivered, the buyers suffered substantial losses. Once operating at a loss, they seemed to suspect that the experimenter would not collect money from the subjects, so they had little to lose from further losses. With perceived downside risk gone, control over incentives was lost. The surprise bonus was sufficient to bring all buyers back to a profitable position. When the surprise bonus was given to buyers, they were told to expect no more bonuses. Of course we had no real control over expectations, so we were potentially trading one problem for another.

B. *Market and Regulatory Variables*

Institutional variables were those that deal with market organization, information, and the rights and guarantees afforded to participants. The institutional variables are the treatment variables. When and how did the grade of a unit become known to a buyer? What guarantees were available to buyers of Regulars who thought they were buying Supers? When and how did the sales record of individual sellers become known? Answers to these questions define the institutional structure of the markets. These institutional features will be discussed after the features common to all markets are outlined.

B.1 The Basic Market Organization

The basic market organization was the same for all markets. Buyers and sellers were located in different rooms. Communication between rooms was accomplished by citizen band (CB) radios. Each room had an experimenter in front of the room equipped with a large chalkboard and a CB radio. A long horizontal line scaled from zero to infinity francs was displayed on the chalkboard. Buyers submitted bids that were transmitted to the seller room over the CB by the experimenter. At the same time the experimenter in the buyer room entered the bid under the horizontal line at the franc value equal to the bid. When the bid transmission was received in the seller room, the experimenter repeated the bid and entered it under the horizontal line at the appropriate value. Similarly, when sellers tendered offers, the offer was entered above the line at the appropriate value and transmitted to the buyer room where it was verbally repeated and entered on the chalkboard. If two bids (offers) were tendered at the same price, the second one was listed below (above) the first one. Thus the time of tender is partially ordinally indexed by distance from the line.

Bids and offers remained open until accepted or canceled. Buyers or sellers accepted offers/bids verbally by indicating to the experimenter the one they wanted from those on the chalkboard. Traders were free to indicate the particular bid or offer they wanted independent of the temporal order of tender. An acceptance was immediately radioed to the other side of the market over the CB. Of course, since the CB transmitter and receiver were located in the room with agents, all transmissions over the radio were public. Once a trade was made the bid/offer was circled on the chalkboards and numbered. Aside from bids, offers, acceptances, and other necessary communications with experimenters, the participants were not allowed to say anything. No talking was permitted.

B.2 Regulatory Environment

The major treatment variables were warranties, warranty enforcement, identification of the seller of units, and the timing and public or private nature of grade revelation. These variables are discussed in order.

Warranties, when they existed, were express warranties[1] generated by a claim or grade advertisement by the seller prior to the buyer's purchase. In some cases sellers and buyers could do nothing other than make bids and offers with no reference at all to the grade of the unit. This condition is designated as "N" because no warranties of any sort existed or could exist. Under a different condition, condition "O," sellers had the option of advertising a unit as a Regular or Super at the time an offer was tendered to the market. The offer was then tagged on the chalkboard as an S or R according to the seller advertisement. Likewise, under the "O" condition buyers had the option of indicating along with a bid the grade of the unit desired. A third condition, "R," required sellers to advertise or disclose units as either a Regular or a Super at the time of an offer and required buyers to indicate with all bids, the grade of the unit desired. Thus, the regulatory environment governing warranties could be any of the conditions (N,O,R).

Warranties could be unenforceable (condition U) or enforceable (condition E). If warranties were unenforceable, no regulations existed governing the cases in which sellers failed to deliver the grade that was promised in the advertisement or requested by the buyer. That is, sellers could advertise a unit as a Super but deliver a Regular and the buyer could do nothing about it. In essence, false advertising was permitted. If warranties were enforceable (condition E) buyers were granted "specific performance."[2] That is, the seller was required

[1] Section 2-313 of the uniform commercial code requires: (a) Any affirmation of fact or promise made by the seller to the buyer which relates to the goods and becomes part of the basis of the bargain creates an express warranty that the goods shall conform to the affirmation or promise. (b) Any description of the goods which is made part of the basis of the bargain creates an express warranty that the goods shall conform to the description.

[2] Consistency of this regulation with the uniform commercial code is covered in Section 2-716. (footnote cont'd)

to deliver a Super to the buyer if the unit had been so advertised. Thus, the enforcement condition could take two values (U,E).

In some markets sellers' identification numbers accompanied all offers and bids transmitted over the CB. Furthermore, under such conditions buyers were able to direct bids to individual sellers and such tagged bids could only be accepted by the requested sellers. This condition is designated as K to indicate that sellers'' (but not buyers'') identifications were known at the time of a contract. In the alternative condition U, neither buyer nor seller ever knew the identity of a trading partner. Thus, the identification variable took two values (U,K).

Unless grades were covered by an enforced express warranty, buyers became aware of grade either immediately after the purchase (condition A) or at the end of a period (condition E). Under condition A the seller held up a card immediately after the sale with letter S or R indicating the unit as Super or Regular. The information was then transmitted by the experimenter to the buyer. Under condition E the seller would submit a slip of paper indicating the grade for each trade in which the seller was involved. Trades were numbered on the chalkboard and sellers and buyers would record the number attached to each trade along with the price, etc.

The case in which the enforced warranty is provided is a little hard to describe notationally. If a grade was advertised, which need not be the case under condition "O," the buyer was aware of the grade prior to purchase. Thus the notation B is used. The actual announcement, however, could have been "A" or "E."

Some interpretations are in order. An enforced warranty can be interpreted as a case in which all characteristics of the product can be fully identified and evaluated by the customer prior to purchase. If the grade becomes known immediately after the sale, the customer has no recourse from unfulfilled expectations except alterations in future purchase patterns. Since the information becomes available immediately after a purchase, the consumer can react through modifications of purchasing behavior for the remainder of the period as can other buyers if the information is public. If the information becomes available only at the end of a period, the consumer is faced with a type of "credence" problem. During a period the consumers are unable to evaluate purchases. The information that permits evaluation becomes available only after a delay.

Information about grade was either publicly revealed (condition Pub.) or privately revealed (condition Pvt.). In the case of public revelation the informa-

---

(1) Specific performance may be decreed where the goods are unique or in other proper circumstances. (2) The decree for specific performance may include such terms and conditions as to payment of the price, damages or other relief as the court may deem just. (3) The buyer has a right of replevin for goods identified to the contract if after reasonable effort he is unable to effect cover for such goods or the circumstances reasonably indicate that such effort will be unavailing or if the goods have been shipped under reservation and satisfaction of the security interest in them has been made or tendered.

tion regarding grade communicated to the experimenter was then announced over the CB for all to hear including the buyer. If the information was privately revealed, the slip indicating grade was passed along to the buyer or a cipher was used to privately transmit the grade over the CB. The latter procedure was useful if the rooms were so far apart that physical delivery of slips slowed the process excessively.

C. *Experimental Design*

A total of twenty-one markets was studied. The treatment variables included experience on the part of buyers and sellers, the relative cost of Supers, and the regulatory variables listed above. Obviously, with the large number of potential treatment variables not all possible experiments could be conducted. The strategy was to follow the sequential process outlined in the introductory statements. The choice of a particular experiment depended in part upon the availability of subjects and the pattern of previous results.

The treatments chosen for each of the twenty-one markets are listed in Table 1. The conditions of an experiment are indicated by an 8-tuple.

1	2	3	4
Warranty Offered (N,R,O)	Warranty Enforcement (U,E)	Trader I.D.'s (U,K)	Time of Grade Revelation (B,A,E)
5	**6**	**7**	**8**
Method of Grade Revelation (Pub,Pvt)	Experience (N,E,VE)	Relative Super Cost (L,H)	Location (BU,CIT,PCC)

For example, the index (O,E,U,B,Pub,VE,H,PCC) is a market in which warranties were optional but enforced if provided; trader I.D.'s were unknown; grades were known before purchase because warranties were enforced; the grades were publicly announced; traders were very experienced; the cost of Supers was in the relatively high condition; and the experiment was conducted at Pasadena City College.

Subjects with no experience (N) had participated in no experiments of the type under examination here, but some subjects from Caltech had participated in market experiments of a different type and were thus somewhat familiar with a market experimental environment. Experienced (E), subjects had participated in at least one previous experiment in this series. In almost all cases of new subjects the first market experience involved at least two different treatment variables that resulted in different patterns of market price so that afterwards subjects were all somewhat familiar with aspects of the parameters. Very experienced subjects (VE) had participated in at least two previous experiments.

The description of the other variables must proceed with the discussion in the section above. The easiest way to understand the variables is to notice that advertising and the warranty are tied together in interpretation. If a grade is specified along with a bid or offer, it is viewed as both advertising and a warranty. The two are equivalent, because if a grade specification is available to any buyer, it is available to all. The interesting additional variable is whether or not the warranty is enforced or, equivalently, whether or not the advertising must necessarily be truthful. A "defective unit" backed by an unenforceable warranty is equivalent in these markets to an advertisement about grade that is false. As will be discussed in the parameter section below, the cost to sellers of delivering Supers was always higher than Regulars. If the difference was 100 francs ($1) per unit, the cost was in the high (H) condition. If the cost difference was 20 francs ($.20) or 25 francs ($.25) per unit, the condition was low (L).

D. *Experimental Procedures*

Subjects were recruited from BU undergraduate business and PCC undergraduate economics classes and from Caltech dorms. The "sales pitch" included with the instructions in Appendix A contains the essence of the information given subjects when they were recruited. All were told that the experiment would take approximately three hours. They were told that we could not guarantee an amount, but that they would have an opportunity to make "more than they would likely make in a comparable hourly period," that "we have never had a dissatisfied customer," and that "we were interested in studying situations in which people make decisions that matter, so we provided incentives accordingly." Such statements were intended as assurances that the stakes could at least cover their opportunity cost. Of those that signed up at PCC, approximately 65 percent actually showed up. The rates were higher at BU and Caltech.

At the assigned time and location the number of subjects present were counted and a decision was made about the number of buyers and sellers.[3] Subjects were randomly assigned instruction sheets as buyers or sellers. Buyers were on one side of the room and sellers were on the other side. Forms in the instructions were reproduced on the chalkboard. Instructions were then read, questions were answered. The market process was explained, including the bids and offers process, the chalkboard, and the determination of Supers and Regulars. If warranties or advertising were involved, special instructions regarding these were included.[4] After all questions were answered, sellers were then accompanied to another room.

[3] We preferred to have two more buyers than sellers. This would assure unique price predictions by certain models.

[4] *See* appendix.

Table 1
A List of Conditions, Subject Pools and Parameters Used in all Experiments

Key to Abbreviations

Column 3	Column 4	Column 5
N = no	E = yes	U = unknown
R = required	U = no	K = known
O = optional		

1	2	3	4	5	6	7	8
					Information about Grade		
Exp. No.	Location	Grade Warranty Requirement	Grade Warranty Enforcement	Seller* Identification	Time	Method	Subject Experience
1	BU	N	-	U 1-6 K 7.8	E	pvt	N
2	BU	N	-	U	E	pvt	N
3	BU	R 1-6 N 7-9	E -	U U	B E	pub pvt	N
4	BU	R 9-8 N 9.11	E -	U U	B E	pub pvt	N
5	BU	N	-	K	E	pub	N
6	BU	N	-	K	E	pub	E
7	BU	N	-	K	E	pub	E
8	CIT	N 1-6.9.10 R 7.8	- E	U U	E B	pvt pub	N
9	PCC	N 1-8 R 9-11	-	U U	E B	pvt pub	N
10	PCC	N 1-6 O 7.8	- E	K K	E B	pub pub	N
11	CIT	N	-	K 1-8 U 9-10	E	pub 1-8 pvt 9.10	8-E 6-N
12	PCC	N 1-6 O 7	- E	K	A	pub	N
13	PCC	N 1-6 O 7	- E	K	A A	pub pub	N
14	PCC	N 1-7 O 8.9	- U	K K	A	pub	N
15	CIT	N	-	K	E	pvt	5-VE 6-E 3-N
16	PCC	N 1-7 O 8.9	- E	K	A	pub	E
17	PCC	N 1-8 R 9	- E	K	A	pub	3-VE 8-E 2-N
18	PCC	N 1-6 O 7.8.9 R 10.11	- U E	K	A	pub	5-VE 9-E
19	PCC	N 1-3 O 4-14	- U	K	A	pub	2-VE 5-E
20	CIT	N	-	K	E	pvt	N
21	PCC	N	-	K	A	pub	VE

* Buyer identifications were always unknown except periods 7 and 8 of experiment 1.

Key to Abbreviations

Column 6	Column 7	Column 8
B = before purchase	pub = public	N = none
A = after purchase	pvt = private	E = experienced
E = period end		VE = very experienced

1	9	10	11	12	13	14	15
	Cost		Number		Full Information Competitive Equilibrium Price**		
Exp. No.	Supers	Regulars	Buyers	Sellers	Supers	Regulars	Total Number Periods
1	120	20	7	6	300	165	8
2	120	20	7	6	300	165	7
3	120	20	5	4	300	165	9
4	120	20	8	6	300	165	11
5	120	20	8	6	300	165	7
6	125	25	8	6	305	170	12
7	125	25	8	6	305	170	12
8	120	20	6	6	300	165	10
9	120	20	7	6	300	165	11
10	120	20	8	6	300	165	8
11	120	20	8	6	300	165	10
12	120	20	8	6	300	165	7
13	120	20	8	6	300	165	7
14	120	20	8	6	300	165	9
15	120	20	8	6	305	170	10
16	120	20	8	6	305	170	9
17	120	20	7	6	305	170	9
18	120	20	8	6	305	170	11
19	120 (7-14)40	20 20	5	1(6)***	300	165	14
20	120	20	8	6	300	165	9
21	125 (5-14)45	25 20	8	6	305	170	14

** $/F was .02 for buyers and .01 for sellers. Buyers received 50F per period endowment plus a one-time unexpected payment of 200F after period 1.

*** The monopoust has the capacity of eight sellers.

When buyers and sellers were in separate rooms, questions were again answered. Buyers completed a period zero.[5] They were also warned that they must keep accurate records and note the transaction numbers. If we found anyone who "mistakenly" recorded Regulars as Supers, we would need to terminate the experiment.[6] The market opened for period one and it remained open for seven minutes as opposed to the usual five. After period one the extra bonus of 200 francs was given to buyers in addition to the 50 franc per period endowment. Buyer and seller record sheets were checked after the first, second, and third periods and occasionally after that.

III. Models and Ideas

Ideas and models are outlined in five different categories. We have applied the models to generate a prediction, but the reader should notice that with all of the ideas outlined in this section some latitude exists regarding how a model might best be applied to the markets we created. For example, some models found in the literature are supported by analysis that involves the reasoning process that agents undertake, what they observe, and how they process these observations. Since we did not have access to such data, theories that rest on such ideas remain untested. Instead we applied the models using those operational concepts and measurements that were available and seemed reasonable.

A. *The Full Information Model*

This idea rests on the hypothesis that the markets will behave as if all information about the underlying state of nature available to any agent will be revealed to all through the market process. A natural assumption would be that this model could only be applicable in cases where the buyer knows the seller, or some form of direct communication is possible. However, it is conceivable that the predictions of the model be borne out even when such special conveniences are absent. Sequences of bids, special prices, special offers, etc. could all serve as some sort of signal. Any market is filled with such possibilities, so the model could generate good predications even in cases where buyers and sellers have far less than full information.

The idea is as follows. Each seller presumably knows the quality of a unit to be sold at the time an offer is tendered.[7] The state of nature is thus the pattern of Supers and Regulars offered on the market. The hypothesis is that buyers

[5] *See* instructions. Each buyer was required to list the redemption values in the practice record sheet assuming a sequence of purchases S, R, R, S in period 1 and RSR in period 2. This exercise removed certain confusions about the redemption values.

[6] Fortunately the only cheating problems we detected were in the pilot experiments that caused us to add this statement.

[7] In most markets the seller need not commit to a grade until after a sale. We assume, however, that the decision about gradeis made before an offer is tendered.

will behave as if they can distinguish between offers of Supers and Regulars. Sellers will develop a profit maximizing response to buyer decisions. Application of the laws of supply and demand yield a prediction that only Supers will be sold at a price of P_S (*see* Table 1).

B. *Null Expectations Model*

This idea rests on the hypothesis that buyers without prior instruction on the likelihood of Supers and Regulars will treat them as equally likely. The rational expectations postulate is not applied and neither is a substitute learning axiom. So expectations are postulated to be unchanging. Sellers will adopt a profit maximizing response to this behavior. If Supers and Regulars are expected to be equally likely, application of the laws of supply yields a prediction that all Regulars will be sold at a price equal to the average of P_S and P_R (*see* Table 1).

Clearly a null expectations model could involve any probability at all. The choice of 50:50 is arbitrary. The model is used primarily as a point of reference.

C. *Lemons Model*

Sellers, faced by buyers who behave as if they cannot distinguish Regulars from Supers will adopt a short-term maximizing strategy and sell only Regulars. Buyers seeing only Regulars delivered will develop rational expectations and behave as if they expect only Regulars. Application of the laws of supply and demand yields predictions of all Regulars at a price P_R (*see* Table 1).

D. *Signaling Models*

If firms have a means of adding some distinguishable feature to units, that feature can sometimes be used as a signal that distinguishes offers of Supers from offers of Regulars. If the cost of adding this feature is sufficiently lower for Super units as opposed to the cost of adding the feature to Regular units, then signaling models predict a ***signaling equilibrium***. The feature will be added to Supers only, and its presence will serve as a signal that lets buyers differentiate the underlying grades of units. *See* Spence (1977), Rothschild and Stiglitz (1976), Miller and Plott (1983).

Signaling models have an obvious application when warranty instruments exist. If warranties exist and are costlessly enforced, the cost of adding a warranty of Super to a Super unit is zero and the cost of adding a warranty of Super to Regular units is the difference between the cost of providing a Super and the cost of providing a Regular. The warranty guarantees specific performance, so a seller advertising a Super must deliver one and therefore loses the cost advantage of delivering a Regular. If warranties are required or are optional, then the signaling model becomes the full information model and therefore has the same predictions. The results will be volume that is all Super units sold with a warranty[8] and the price will be P_S (*see* Table 1).

[8] Grossman (1981) develops the notion that the warranty will be added. His model differs from the signaling model, but in this narrow case the predictions are the same.

A model developed by Grossman (1981) leads to the same conclusions (Leland, 1981), but the Grossman model is based on different principles. Grossman applies a perfect equilibria principle from game theory[9] and a rational expectations principle.[10]

If warranties are not enforced, then the cost differential between adding the special feature to Supers and to Regulars disappears. Regulars can be advertised as Supers. The signaling model then predicts that no separation will occur because Super units and Regular units will both add the special feature. Regular units will be offered along with an unenforceable warranty that the unit is a Super. Buyers will adopt expectations accordingly and anticipate that all units are regulars. The final result will be all Regulars at a price of P_R (*see* Table 1).

The lemons model can also be interpreted as a degenerate case of the signaling model. In Akerlof's (1970) model, price serves the dual role of equilibrating supply and demand and signaling the quality of the product sold. Because of the one shot nature of trades and the absence of any cost associated with signaling high quality with high price, price cannot effectively signal quality and therefore only lemons are traded.

E. *Reputation Models*

Models of reputation formation tend to be motivated by the theory of dynamic games. Buyers behave as if they are aware of seller identities and adopt dynamic strategies of rewarding and punishing sellers. Sellers who perform as the buyer desires are rewarded with future business, and sellers who do not perform are avoided. Sellers recognize buyer behavior in developing their own dynamic strategies.

A model developed by Klein and Leffler (1981) postulated a *quality guaranteeing price (weak version)*. Buyers who observe a Regular delivered on terms that buyers would ordinarily expect a Super act as if that seller will always deliver Regulars in the future. A seller who has once "fooled" buyers will sell only Regulars at P_R (*see* Table 1). If sellers anticipate this buyer reaction and if sellers expect one full period more in the market, then, given the parameters in these markets, sellers have an incentive to deliver Supers at any price above P_S - 10. Rational expectations and the law of supply and demand yield a model that predicts only Supers will be sold in the market and these will be delivered at a price of P_S. As the end of the experiment approaches, sellers will sell Regulars at P_S and thereafter sell Regulars at P_R.

A natural extension of the theory to a *quality guaranteeing price (strong version)* can be applied even when buyers do not know seller identities. Buyers, once seeing a regular delivered to the market in the "high" price range, will anticipate that all future deliveries will be regulars. The resulting demand function will be that for Regulars. Price will immediately fall to the regular com-

[9] The principle is imbedded in equations (A4) and (A5) on page 481, Grossman (1981).

[10] Statement (A6) on page 481, Grossman (1981).

petitive equilibrium. Sellers know that a single regular sale will "spoil" the market for all. Thus, if the price is high enough, sellers will sell only supers.

Other reputation models can be found in the literature (Rogerson 1982; Shapiro 1982a,c; Nelson 1974; Schmalensee 1978). The thrust of these models is that sellers who feel that buyers can tailor their reactions to individual sellers by refusing to buy from them or by paying a premium to certain sellers will in turn modify their behavior in anticipation. According to the model, buyers will patronize sellers who have a history of offering good grades and sellers will respond by offering good grades. The result in the parameters of our experimental markets will be that only Supers will be sold. Premiums, prices above P_S, might be paid to sellers who consistently sell Supers.

IV. Results

The time series of all markets are in Figures 2 through 22. Each contract is shown according to price and the ordinal time at which it occurred. Market efficiencies, summary statistics for each period, and the regime of treatment variables is also shown. Tables 2, 3 and 4 show average efficiencies for various periods under each regime.[11] Comparisons of efficiency between different periods or regimes should be made cautiously because the periods may have occurred at different stages in a given market, there may be a different number of periods in the intervals compared, etc. In spite of these difficulties, the reader may find the tables useful in gaining an overview of our results.

Conclusion 1. When disclosures (if made) must be truthful, the full information model works well.

Argument. The relevant markets (periods) are shown in Table 3. The full information model predicts all Super units, 100 percent efficiency and prices equal to P_S in Table 1. Of the 308 units sold during the relevant periods, 275 (89 percent) were Supers. On eight occasions an enforced warranty was imposed after the market had previously been operating under an alternative regulation and in all eight cases efficiency increased immediately. In three cases truthful disclosure was removed and in all three cases efficiency fell immediately. The absolute levels of efficiency are "near" the predicted 100 percent by the second period of enforced warranties. More precisely, the levels are above 90 percent in seven of seven second periods and at 100 percent in three of the seven. By the second period of enforced warranties prices in all cases are within 10 francs of the price predicted by the full information model. Average efficiency for all periods was above 95%.

[11] We are indebted to Richard Craswell for suggesting the classification scheme embodied in Tables 2-4. Market efficiency as developed by Plott and Smith (1978) refers to actual earnings as a percentage of the maximum possible earnings.

Figure 2: Experiment 1

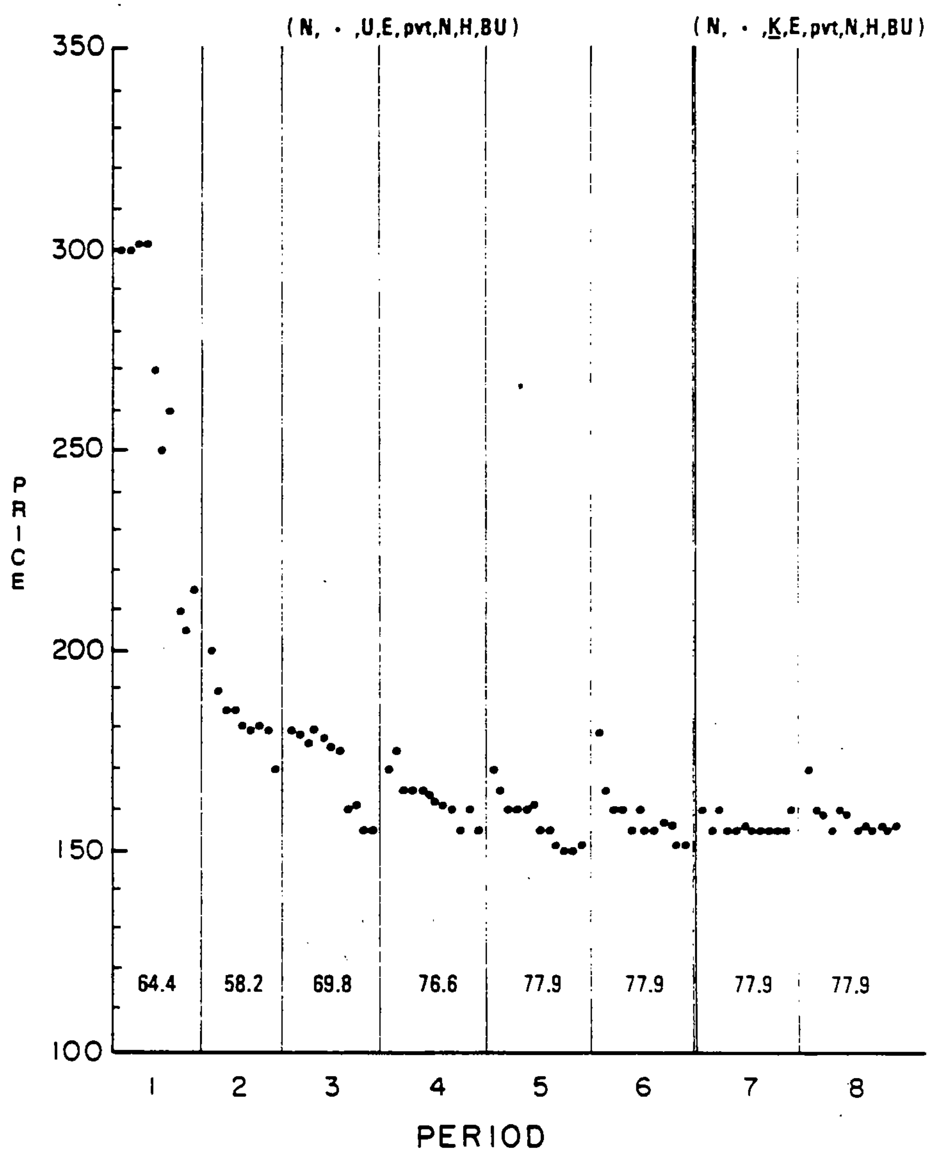


Figure 3: Experiment 2

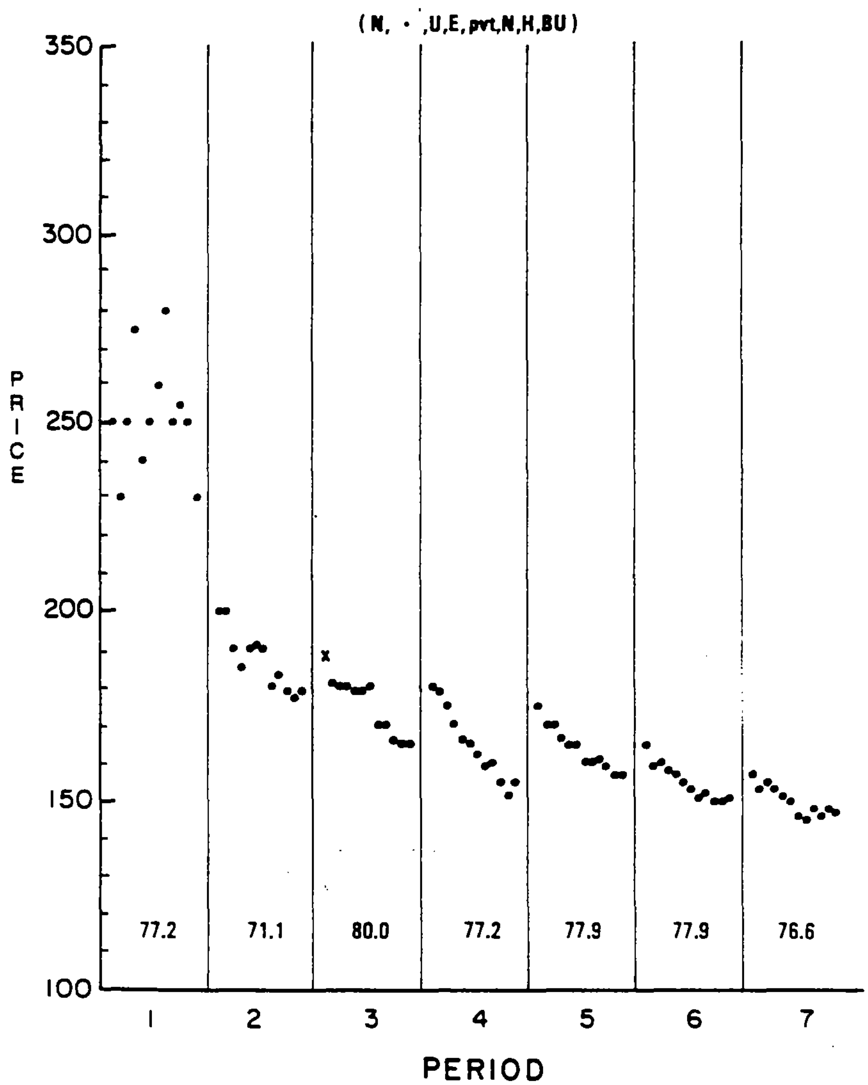


Figure 4: Experiment 3

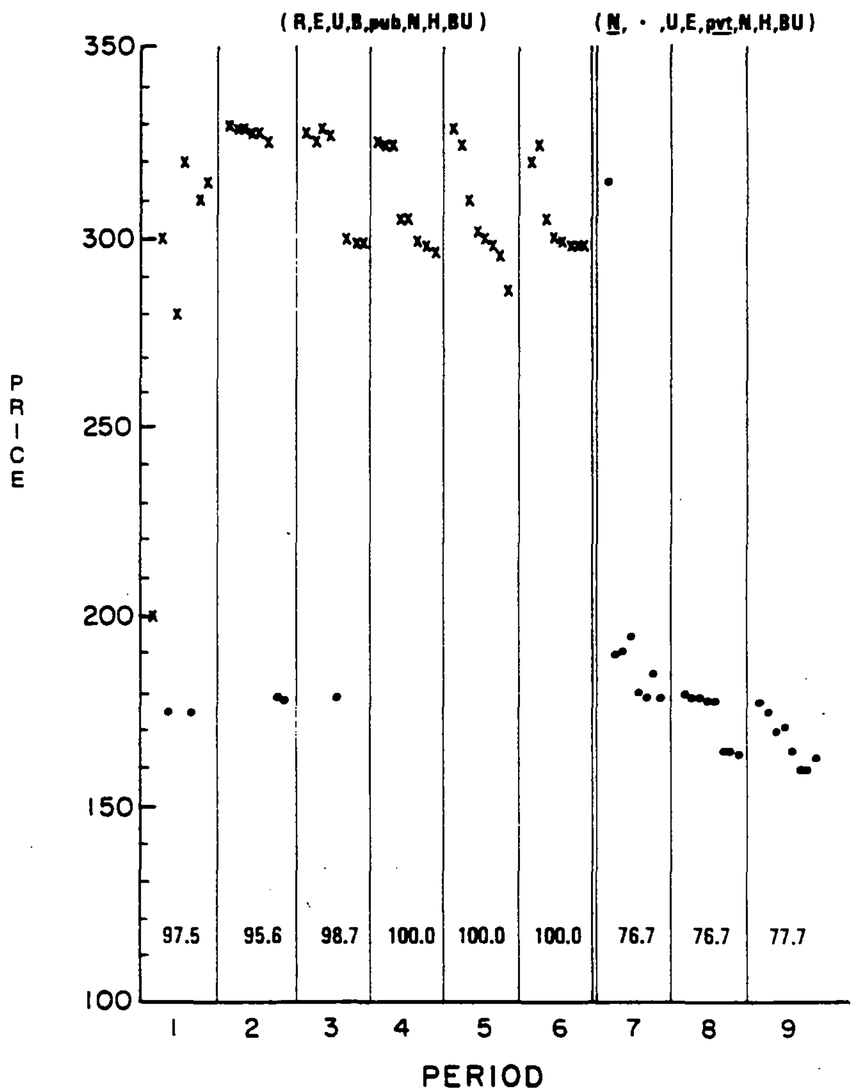


Figure 5: Experiment 4

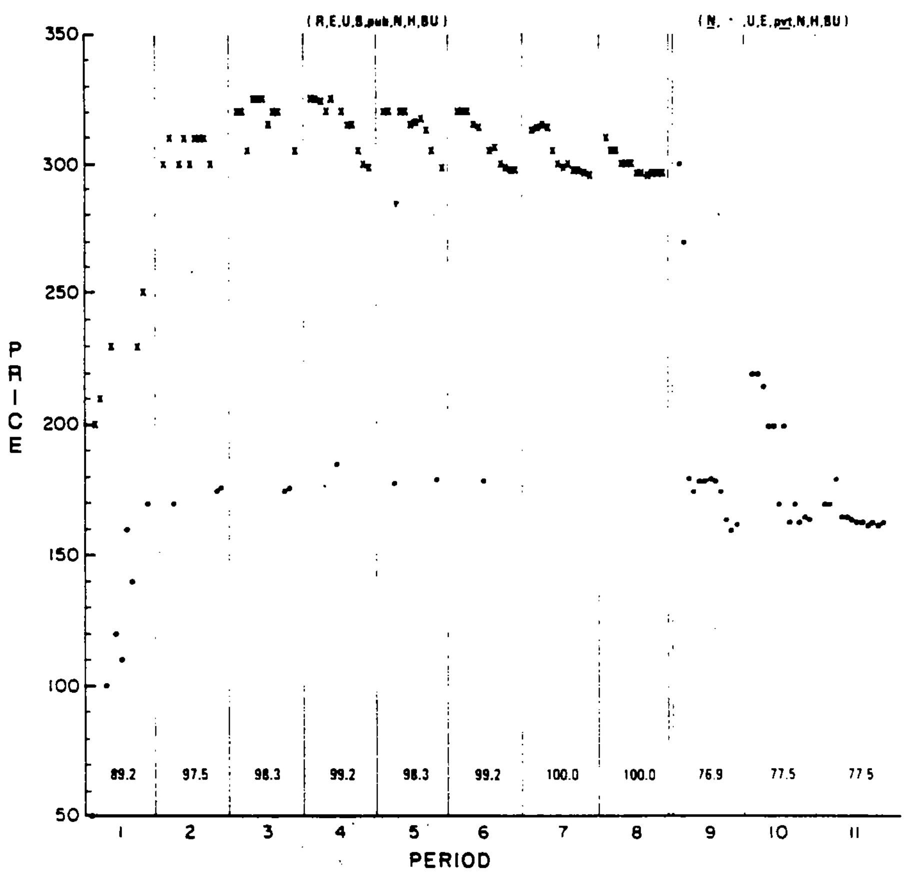


Figure 6: Experiment 5

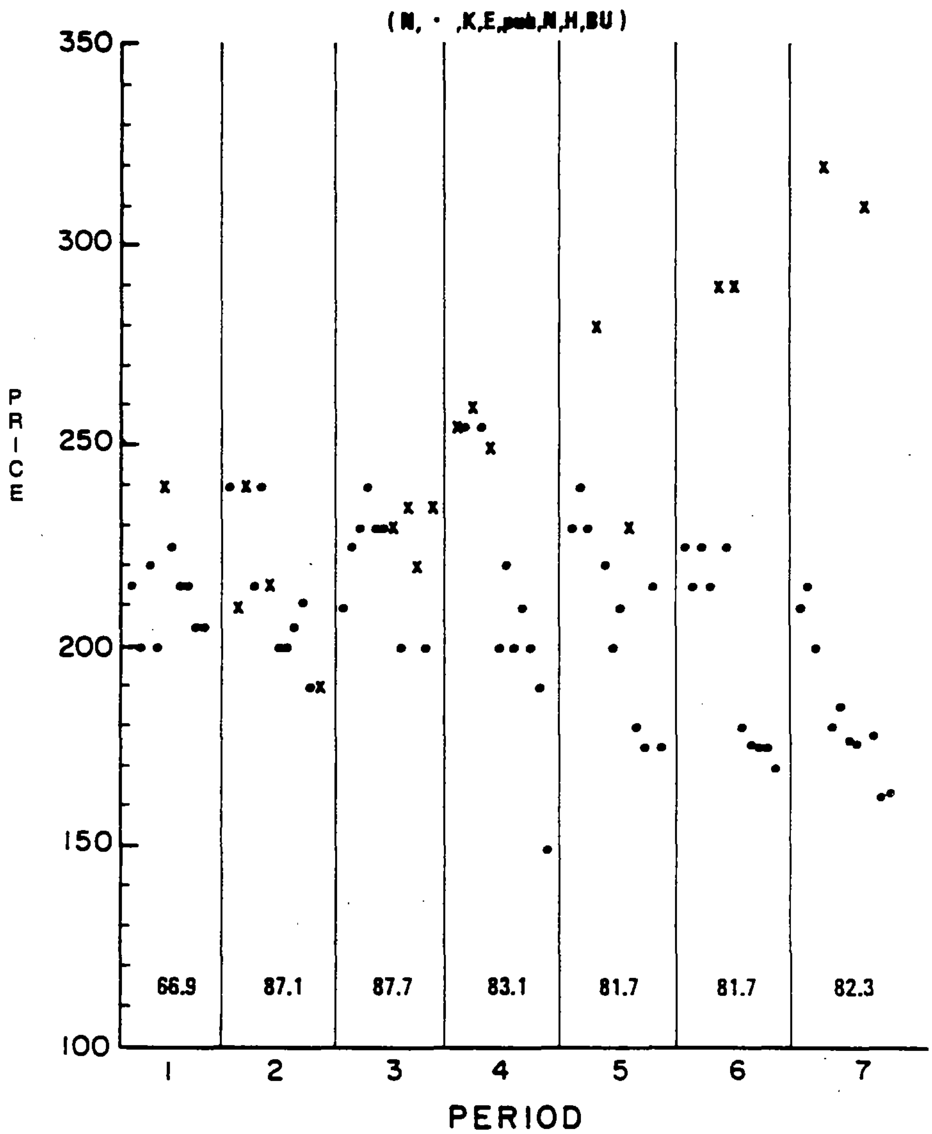


Figure 7: Experiment 6

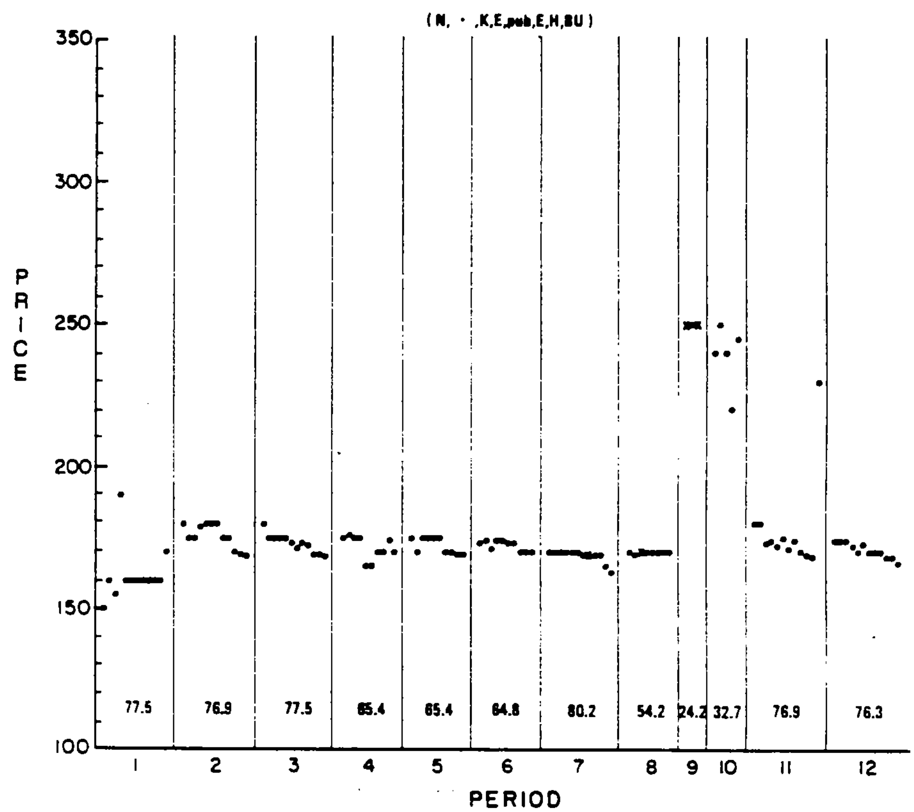


Figure 8: Experiment 7

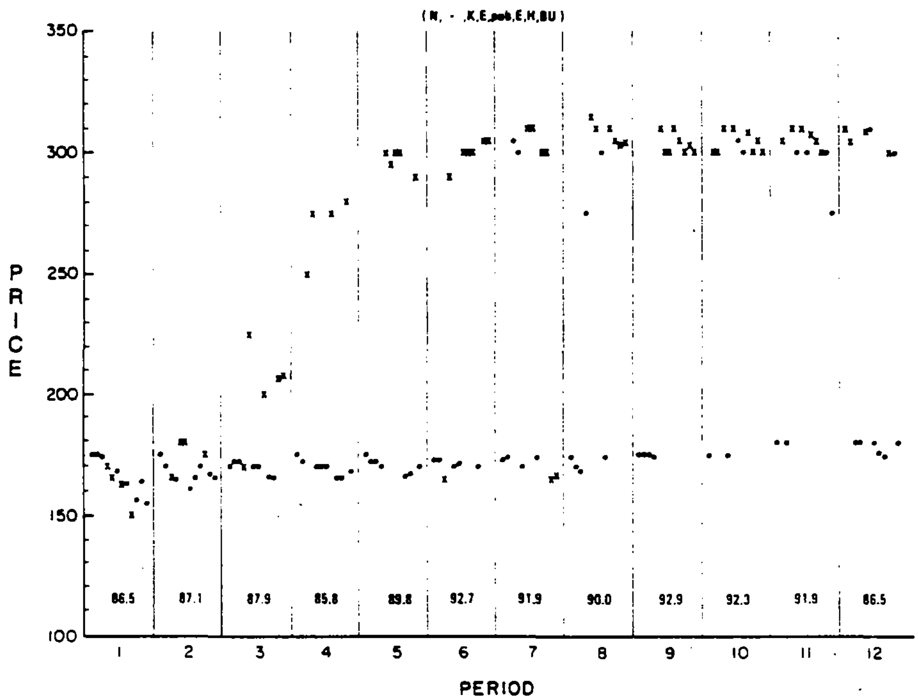


Figure 9: Experiment 8

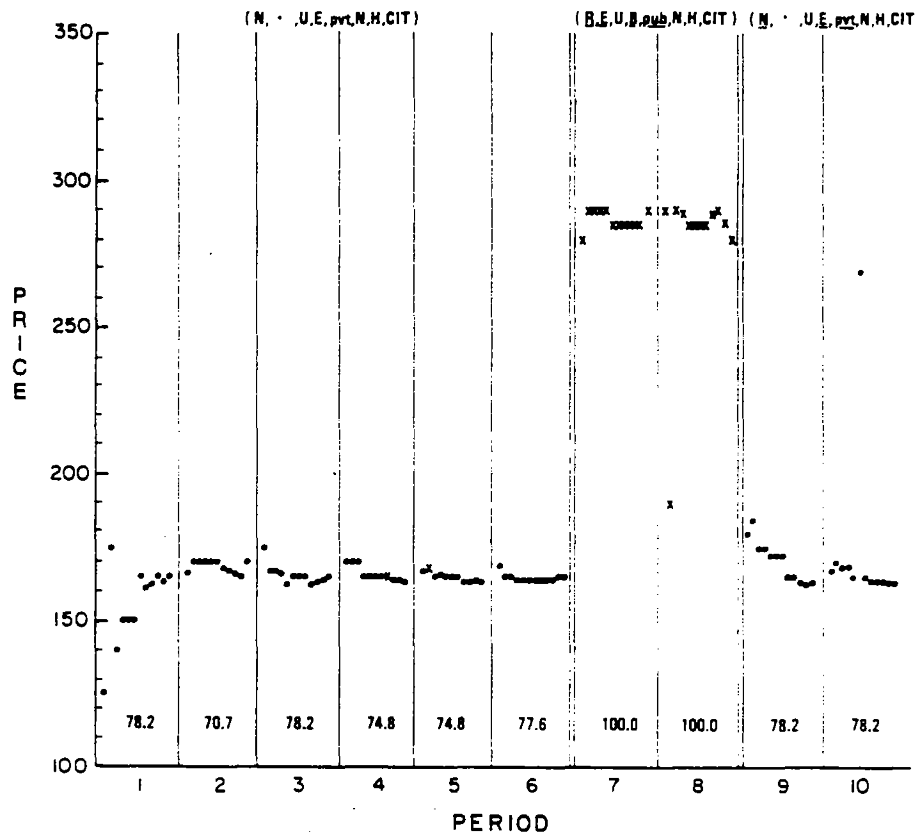


Figure 10: Experiment 9

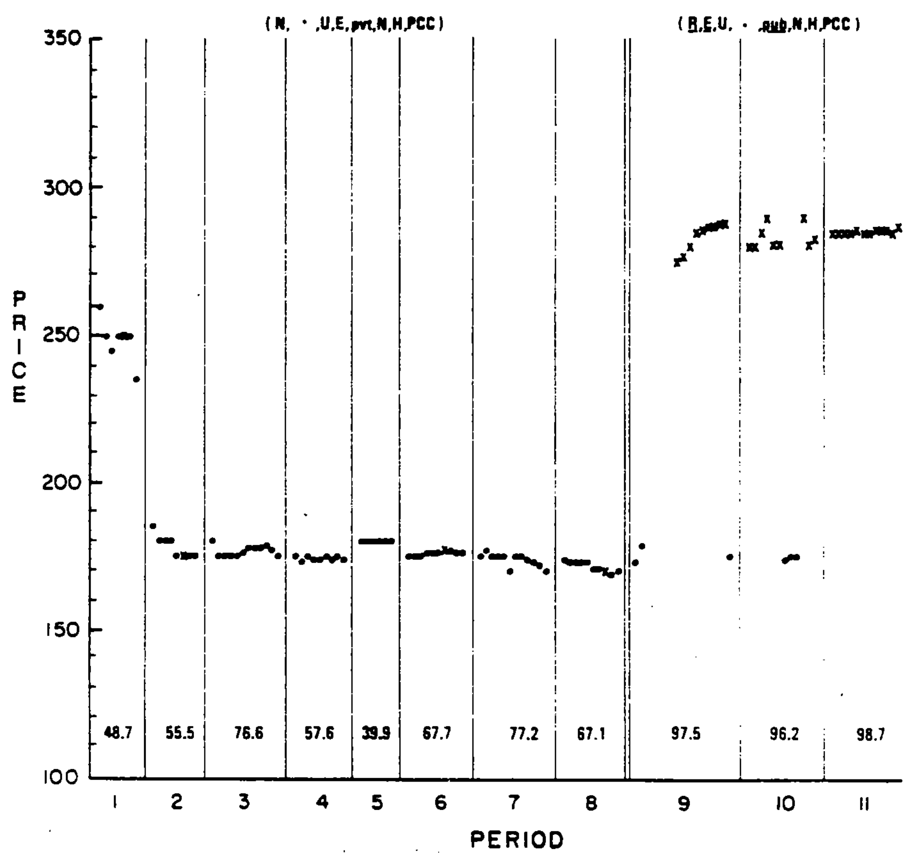


Figure 11: Experiment 10

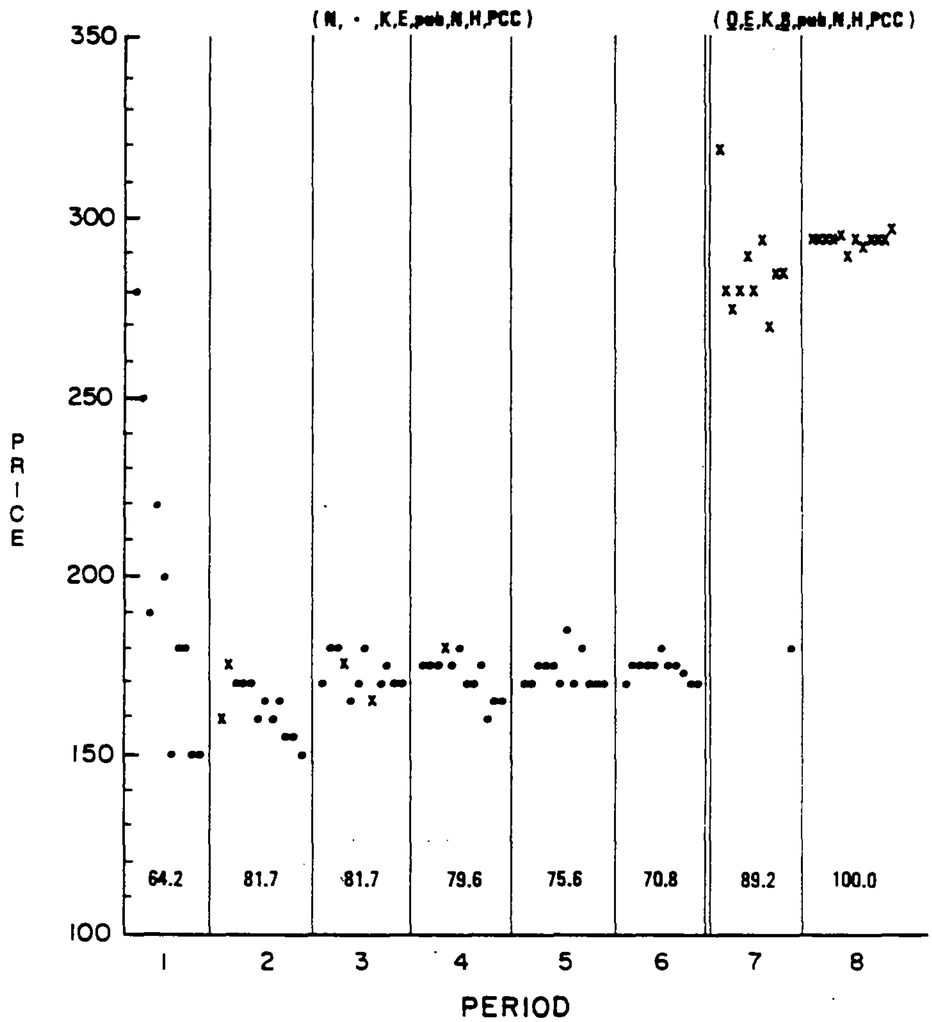


Figure 12: Experiment 11

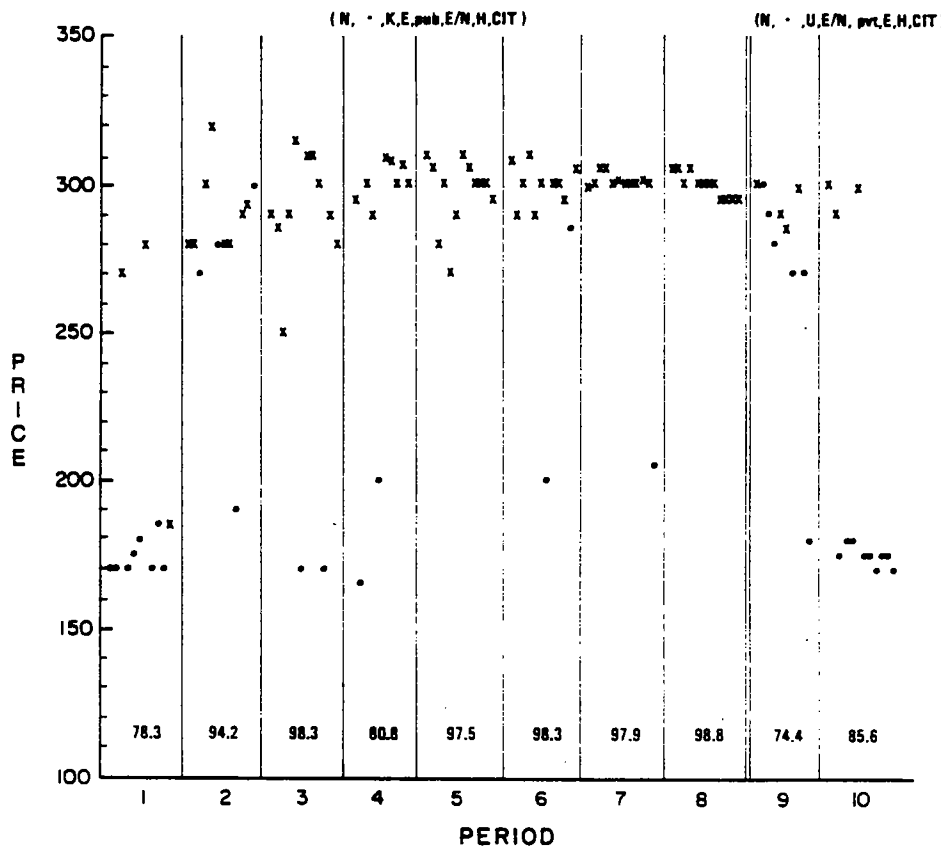


Figure 13: Experiment 12

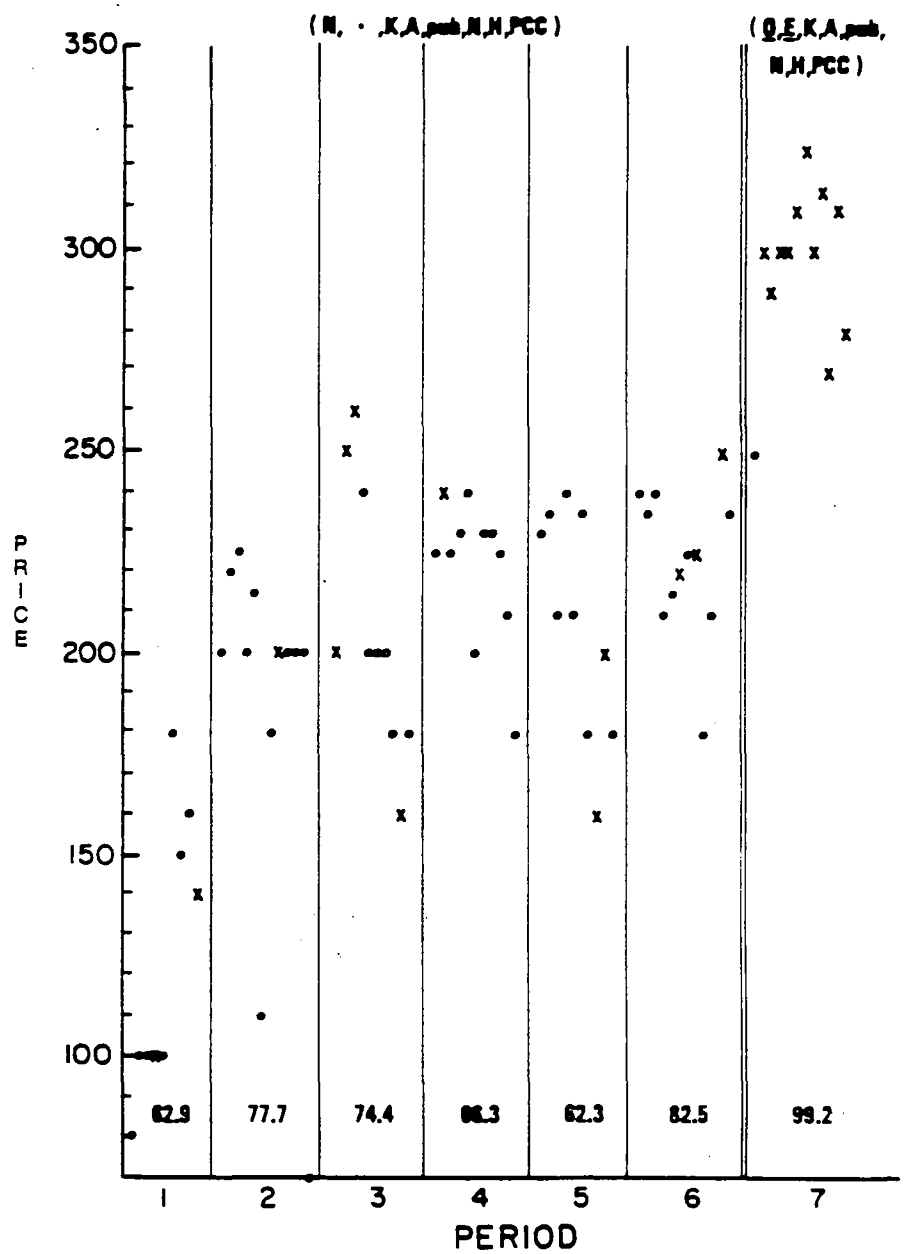


Figure 14: Experiment 13

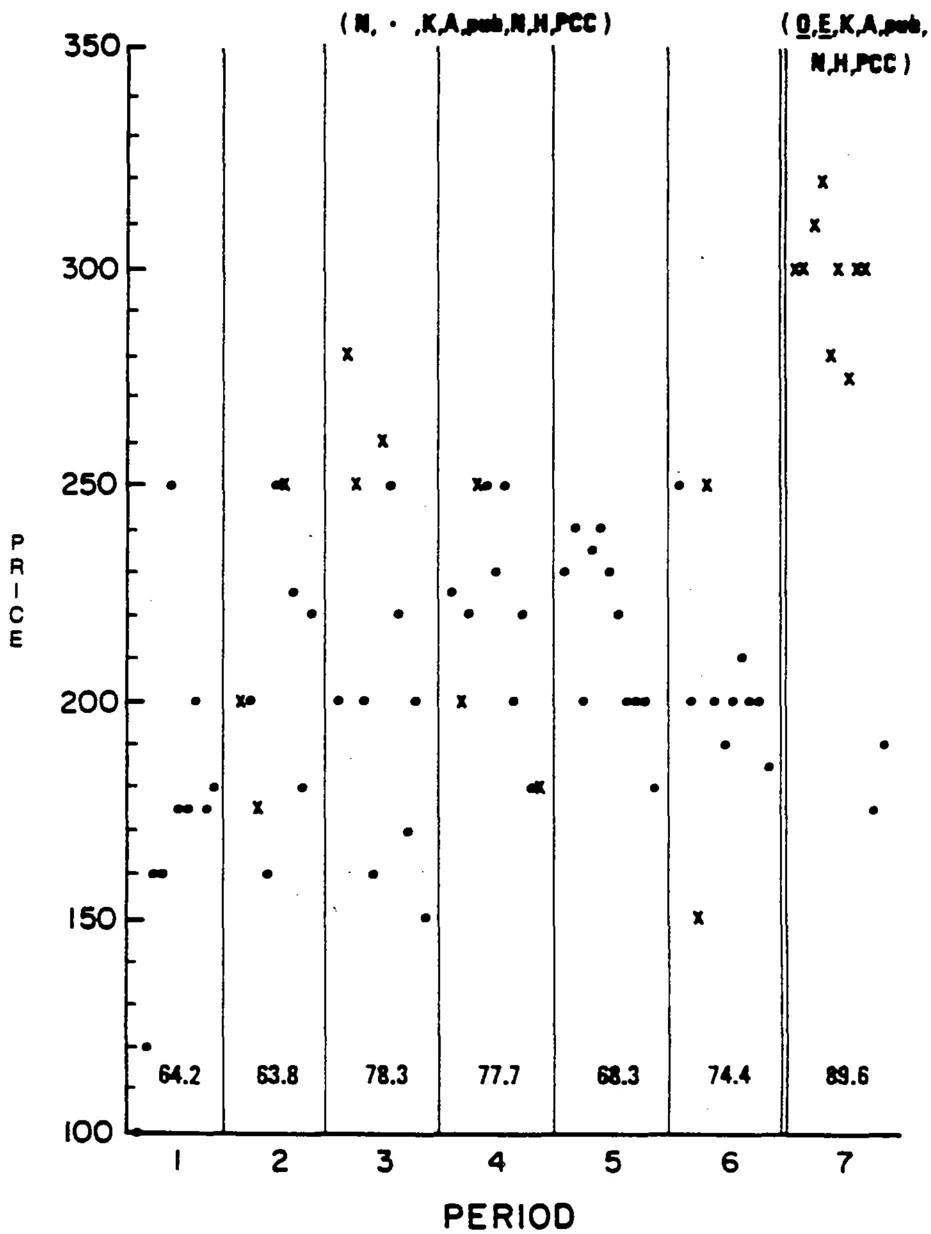


Figure 15: Experiment 14

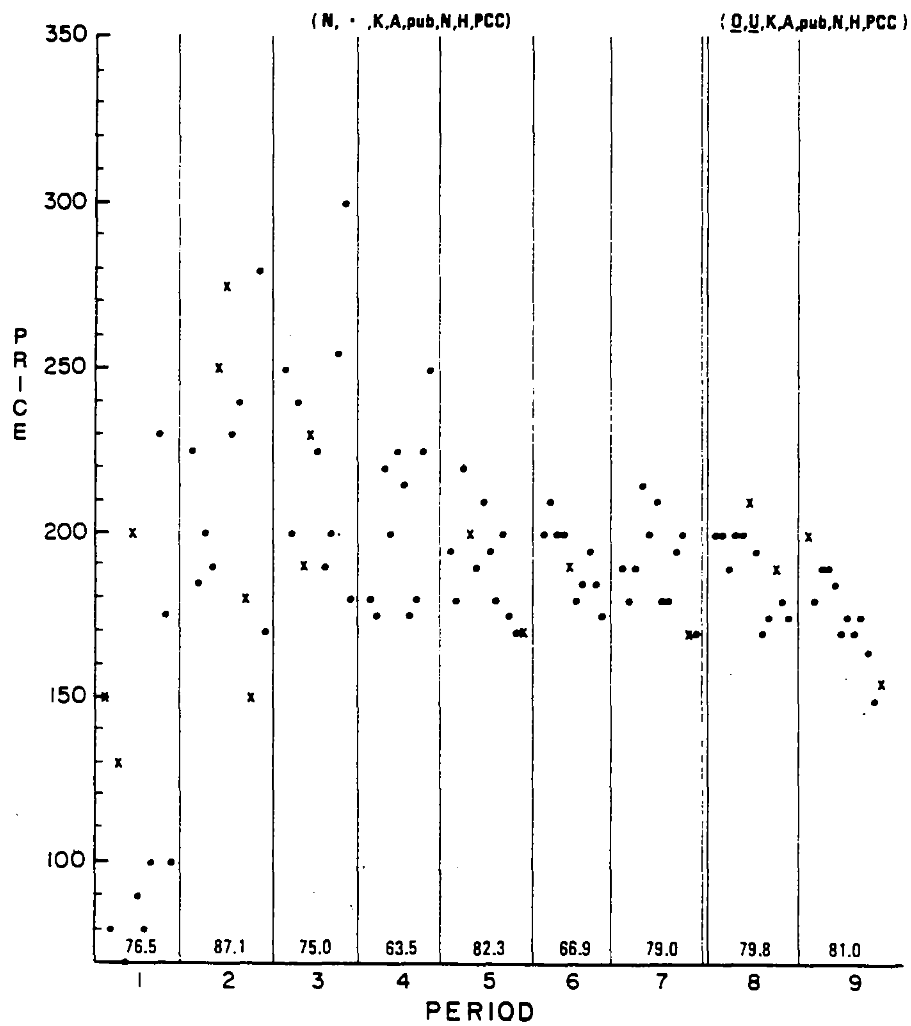


Figure 16: Experiment 15

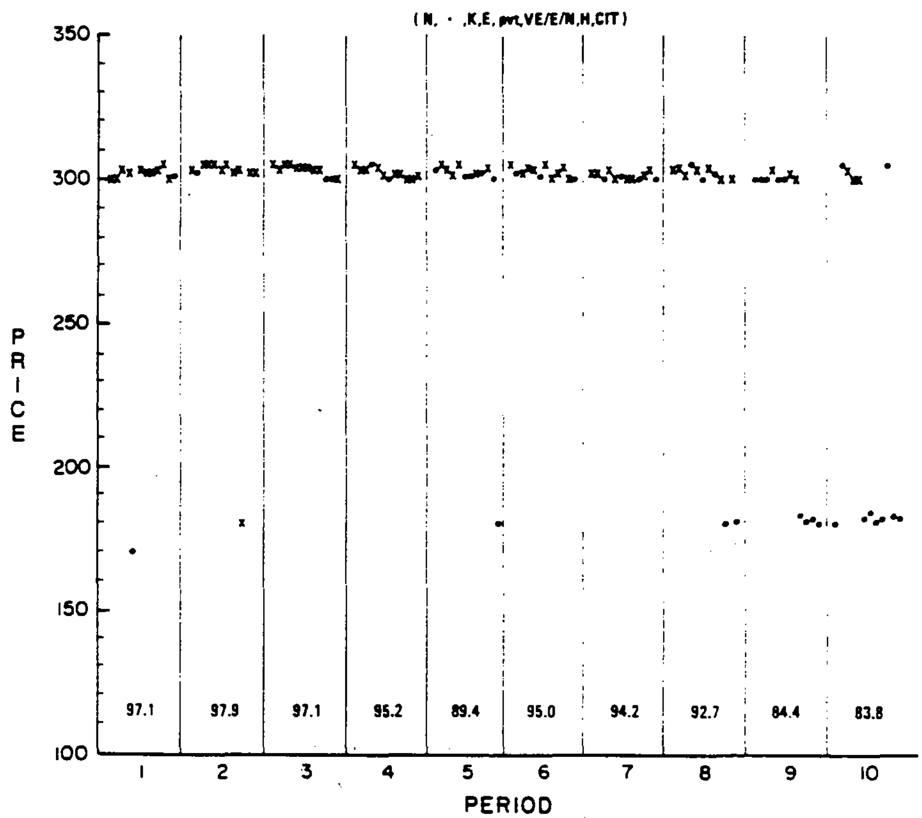


Figure 17: Experiment 16

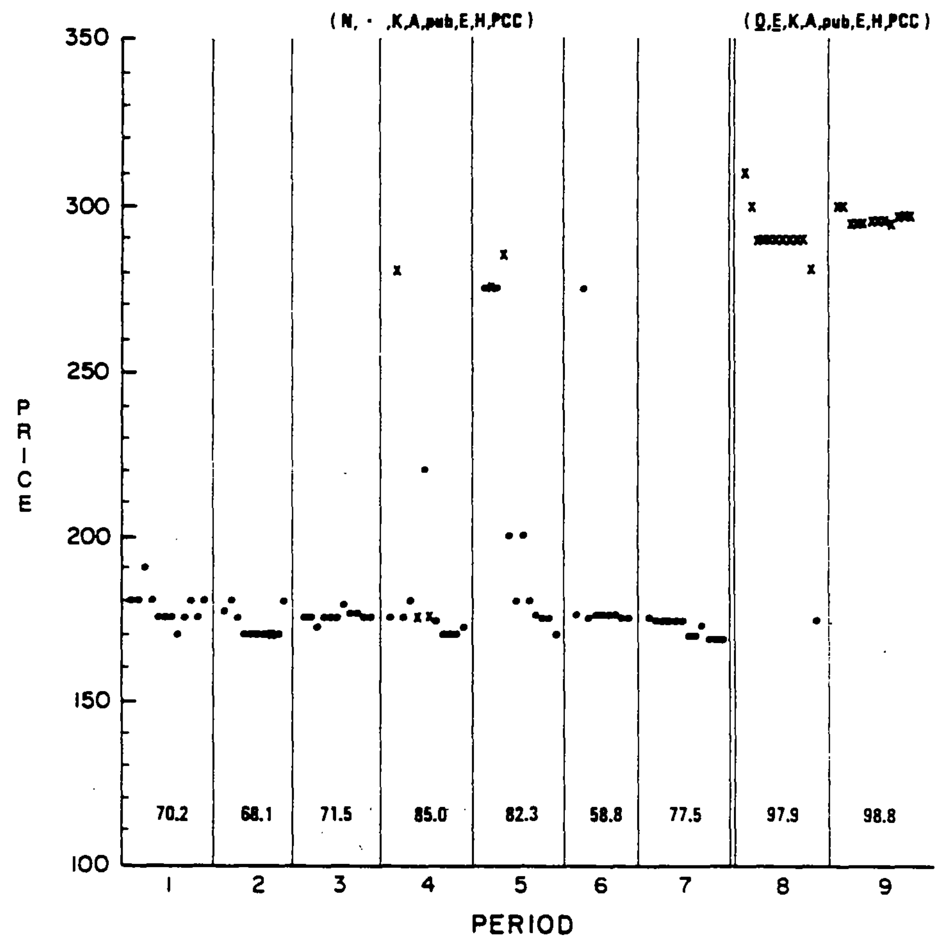


Figure 18: Experiment 17

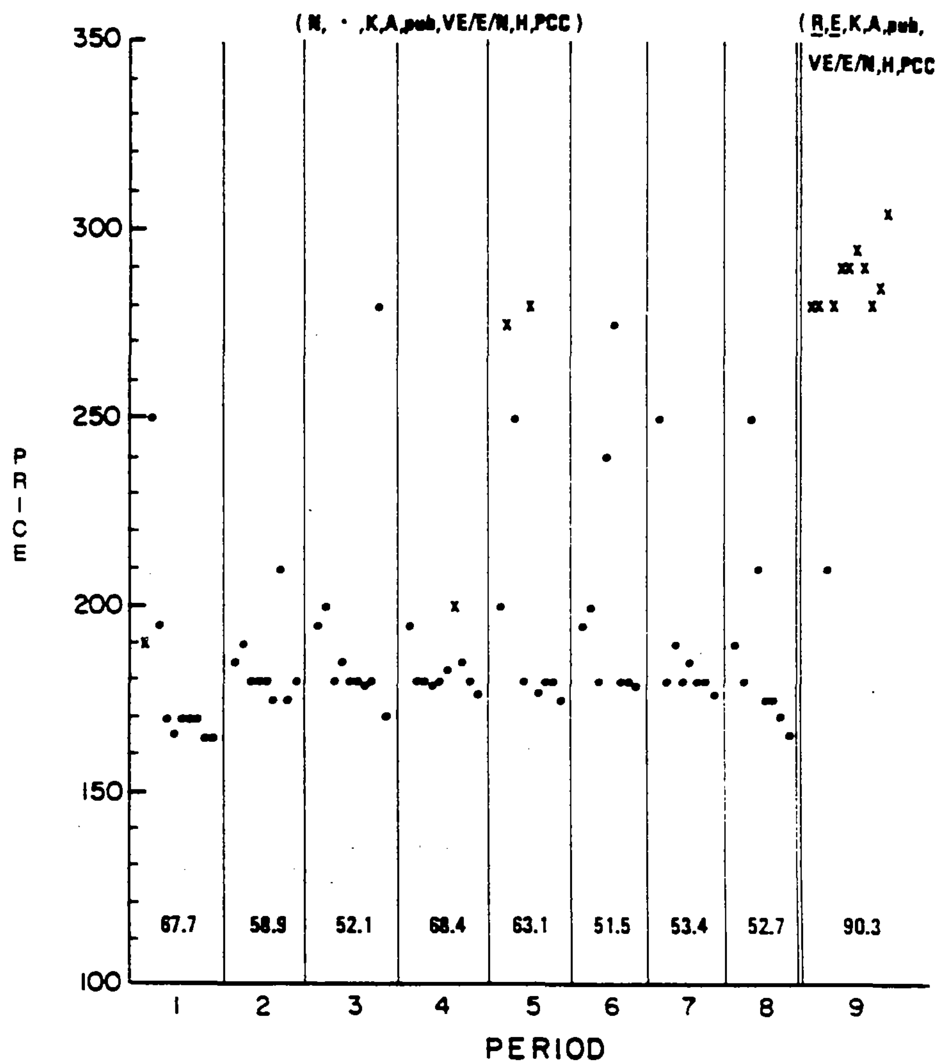


Figure 19: Experiment 18

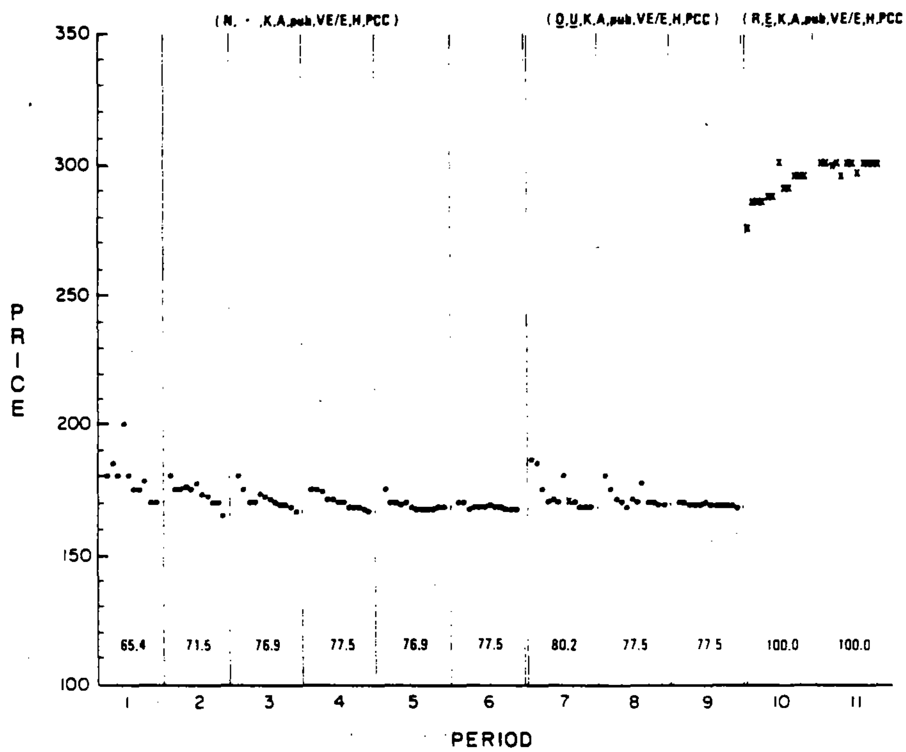


Figure 20: Experiment 19

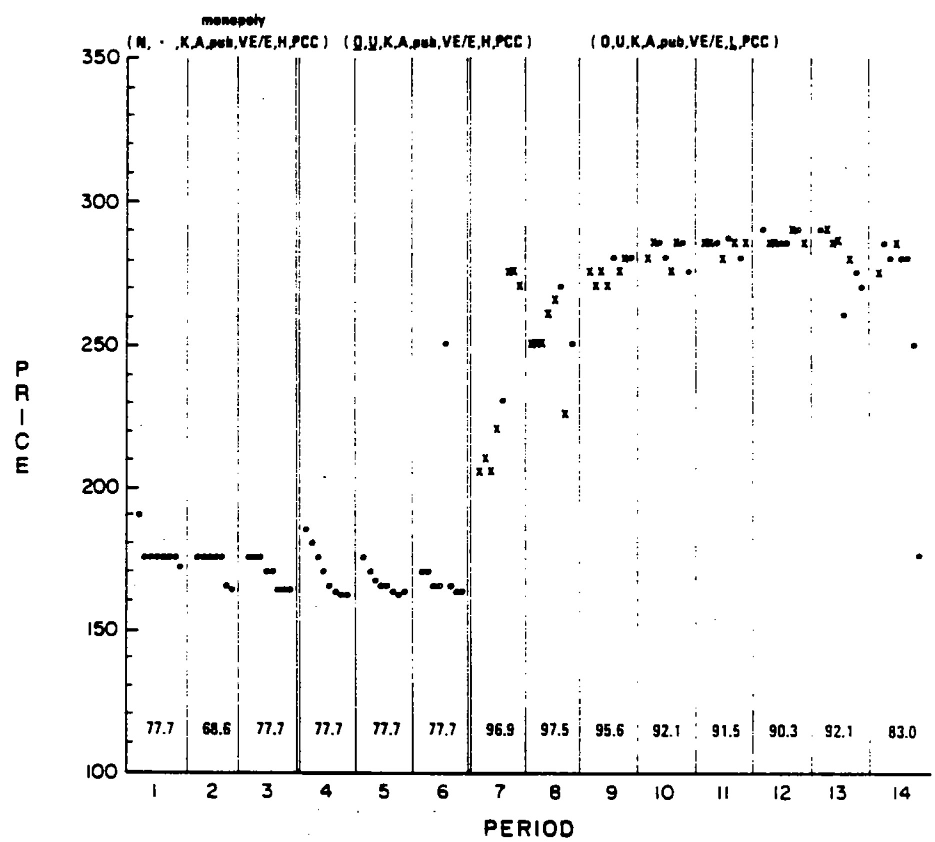


Figure 21: Experiment 20

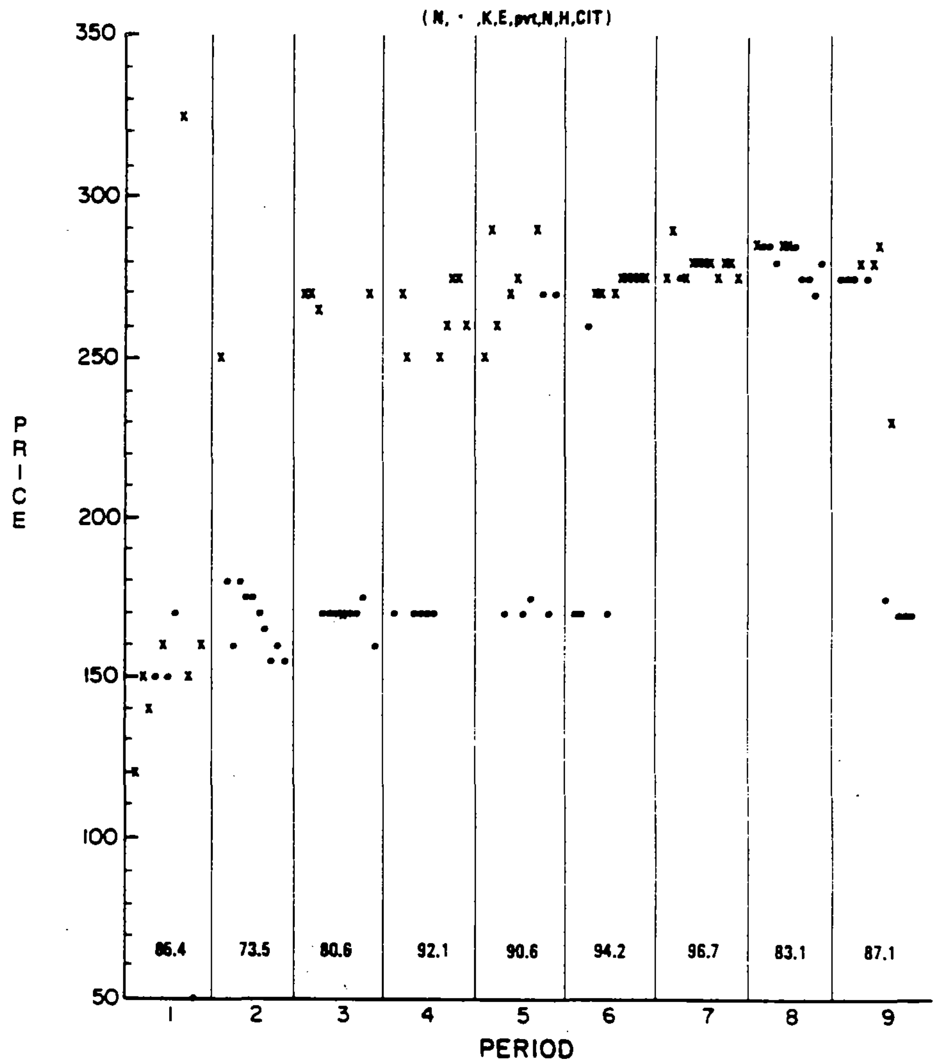


Figure 22: Experiment 21

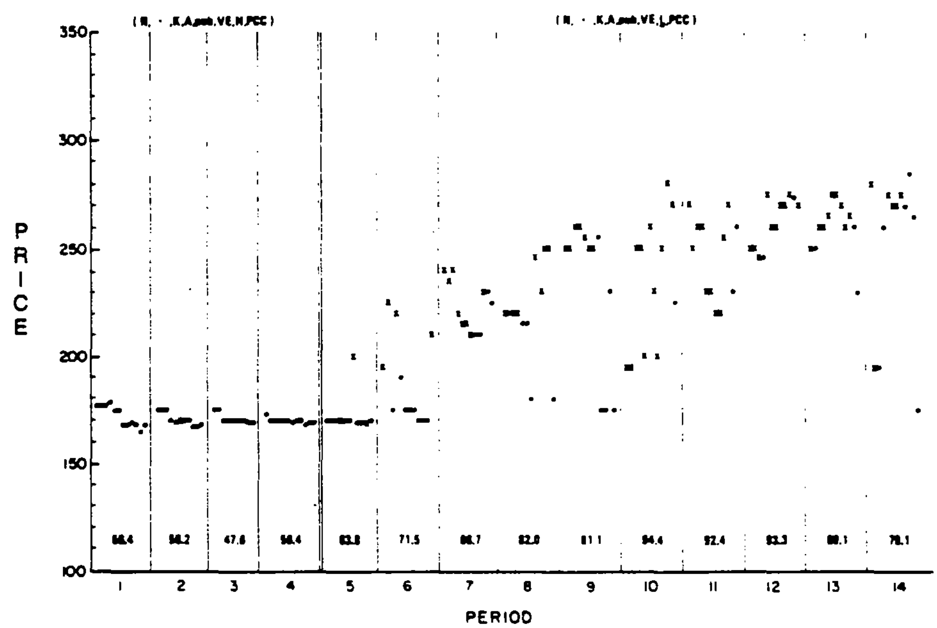


Table 2

EFFICIENCY IN MARKETS WHERE DISCLOSURES ARE PROHIBITED

	Seller ID Known			Seller ID Unknown		
	Experiment &Periods	School	Efficiency	Experiment &Periods	School	Efficiency
Buyers Have:						
	1:(7,8)	BU	78%	1:(1-6)	BU	71%
	15:(1-10)	CIT	93%	2:(1-7)	BU	77%
Private Information Only	20:(1-9)	CIT	87%	3:(7-9)	BU	77%
	Average	CIT	90%	4:(9-11)	BU	77%
	Average	All	89%	8:(1-6,9,10)	CIT	76%
				9:(1-8)	PCC	61%
				11:(9,10)	CIT	80%
				Average	BU	75%
				Average	CIT	77%
				Average	PCC	61%
					All	72%
	5:(1-7)	BU	82%			
	6:(1-12)	BU	64%			
	7:(1-12)	BU	90%			
	10:(1-6)	PCC	76%			
	11:(1-8)	CIT	93%			
	12:(1-6)	PCC	71%			
Public Information	13:(1-6)	PCC	71%	NOTE: An all Regulars "Lemons" Equilibrium is 78% of the Maximum Possible Surplus		
	14:(1-7)	PCC	76%			
	16:(1-7)	PCC	73%			
	17:(1-8)	PCC	58%			
	18:(1-6)	PCC	74%			
	19:(1-3)	PCC	75%			
	21:(1-4)	PCC	73%			
	21:(5-14)	PCC	83%			
	Average	BU	78%			
	Average	CIT	93%			
	Average*	PCC	71%			
	Average*	All	75%			

*Excludes 21:(5-14), periods with low cost supers. For these periods a "lemons" equilibrium is only 56% of maximum surplus.

Table 3

EFFICIENCIES IN MARKETS WHERE TRUTHFUL DISCLOSURES ARE PERMITTED

Advertising (or Labeling or Warranty Provision)

Advertising (Labeling, or Warranty Provision)	Must Be Truthful					
	Seller Known			Seller Unknown		
	Experiments & Periods	School	Efficiency	Experiments & Periods	School	Efficiency
Is Required	17:(9)	PCC	90%	3(1-6)	BU	99%
	18:(10-11)	PCC	100%	4:(1-8)	BU	98%
				8:(7,8)	CIT	100%
				9:(9-11)	PCC	97%
	Average	All	97%		All	98%
Is Optional	10:(7,8)	PCC	95%	NOTE: An all Regulars "Lemons" Equilibrium is 78% efficient		
	12:(7)	PCC	99%			
	13:(7)	PCC	90%			
	16:(8,9)	PCC	98%			
	Average	All	96%	NOTE: Public Information in all periods.		

Table 4

EFFICIENCIES IN MARKETS WHERE DISCLOSURES ARE PERMITTED AND MAY BE FALSE

Advertising (or Labeling, Warranty Provision	Seller Known			Seller Unknown	
	Experiment & Periods	School	Efficiency	Experiment & Periods	School
Is Required	None			None	
Is Optional	14:(8,9)	PCC	80%	None	
	18:(7-9)	PCC	78%		
	19:(4-6)	PCC	78%		
	19:(7-14)	PCC	83%		
	Average*	All	79%		

*Average excludes 19:(7-14), *i.e.*, those periods where the cost of supers was reduced. For these periods the all regulars "lemons" equilibrium is only 56% of the maximum surplus.

Conclusion 2. The lemons model works well when seller identification is unknown and disclosures are prohibited.

Argument. The relevant markets (periods) are shown in the right hand columns of Table 2. The lemons model predicts that only regular units will be sold. Of the 399 units sold in the periods in which seller identification was unknown and warranties were unenforced 384 (96 percent) were regulars. Efficiency predicted by the lemons model is approximately 78 percent in all markets. In the next to last periods of the sequences listed above, the efficiencies are within 1 percent (relative to full efficiency) of the lemons equilibrium efficiency. Average actual prices are within 5 francs of the prices predicted by the lemons model by the fourth period of all markets except market 9 where they are from 10-15 francs too high and only slowly converging. Average efficiency for all relevant periods was actually less than the lemons equilibrium at 72%.

Conclusion 3. The Grossman/signaling hypothesis that enforceable warranties will be voluntarily added to units (or that if disclosures must be truthful sellers will voluntarily make them) is supported in the data.

Argument. The relevant periods are shown in the lower left hand corner of Table 3.[12] Of the 72 offers in the relevant periods, 53 indicated a Super, and of the 83 bids, 74 requested a Super. Of the 65 Supers sold, 64 were supported by an express warranty. Average efficiency is 96%, about the same as when disclosures are mandatory and truthful.

Conclusion 4. Sellers will nontruthfully advertise when it is possible. The "pooling" or regulars posing as supers phenomena predicted by the signaling models are observable when advertising need not be truthful (express warranties are not enforced).

Argument. The relevant periods are shown in the lower half of Table 4 in which advertising was optional but the implied express warranty was not enforceable. A total of 147 trades were made during these periods of which 105 were regulars and 42 were supers. A total of 61 of the regulars sold were falsely advertised as Supers (58 percent) with the other 44 advertised as Regulars or unadvertised. These misrepresentations are not random mistakes, because all of the 42 supers sold were also advertised as Supers. There were no "mistakes" at all.

Conclusion 5. Knowledge of seller indentification in the absence of truthful but voluntary disclosure:

i) does not guarantee efficiency improvements over the lemons" equilibrium, but in some markets such knowledge increased efficiency.

ii) can continue to have an influence if grade is only privately disclosed.

[12] Some ambiguity exists about whether or not warranties were optimal or were required in 17(9) and 18(10,11). Of the 43 offers in these periods, 37 indicated a Super, and of the 39 bids 36 requested a Super. All of the 34 Super sales were supported by a warranty.

Argument. The relevant markets (periods) for part i are given in the left half of Table 2 and the lower left half of Table 4. The relevant periods of experiments 1, 6, 10, 12, 13, 14, 16, 17, 18, 19(1-6), 21(1-4) have efficiency levels insignificantly above that of the lemons equilibrium. For the most part these efficiencies are close to the lemons model predictions, even though the price data suggest that buyers (especially inexperienced PCC buyers) are more optimistic than the lemons model suggests they would be. Experiment 5 is on the borderline but the relevant periods of 7, 11, 15, 19(7-14), 21(5-14) have efficiencies substantially above the lemons model values. In addition, these markets show evidence of improved efficiency over time. The data in parts of 15, 20, and 21 compare favorably with truthful disclosure markets.

Controlling for possible subject pool differences by comparing only experiments drawing from the same subject pool and in which seller identifications are known (for more than three periods) versus those in which they are unknown, the conclusion still emerges. Efficiencies in all but one of the nineteen periods of 5 and 7 are higher than the comparable experiments in the BU subject pool of 1 and 2. In the CIT pool, of the twenty-five periods of 11, 15, and 20 all but two are higher than the relevant periods of 8. Finally, in 11 efficiency goes down when seller I.D.'s are removed. The data in the PCC experiments are less clear unless the cost of Supers is lowered.

The relevant data for part ii come from the CIT markets 15 and 20 in comparison with 8 (upper half of Table 2). In all of these markets grade was only privately revealed. Efficiencies in all but two periods of 15 and 20 are higher than in all periods of 8. Average efficiency is about 90% for 15 and 20, compared to 76% for 8.

Conclusion 6. The signaling model works except where seller identification is known. Where seller identifications were known, the predictions of the model were less reliable and in some cases inaccurate relative to the full information model.

Argument. The signaling model predictions coincide with the full information model when warranties are enforceable and with the lemons model when warranties are not enforceable. Both predictions are supported. When warranties are optional but enforced, the signaling model predicts they will be used.

The predictions of the signaling model remain the same when seller identifications are known. In fact, the signaling model is basically a static model and, unless it is reinterpreted to designate "reputation" as a "signal," predicts that the market will be insensitive to the revelation of seller identification. Conclusion 4 demonstrates that the model fails at this point.

Conclusion 7. Buyer reaction to "ripoffs" is not that postulated by either the strong or weak versions of the quality guaranteeing price model.

Argument. The reaction postulated by the model has buyers boycotting sellers who deliver Regulars at a price that is unprofitable for buyers. Buyers necessarily lose money on any purchase of a regular at a price of 180 or more. Consider all experiments where unenforceable warranties existed, seller identifications were known, and there was public revelation of grade. On 25 occasions sellers delivered a Regular at a price above 180 and were then able to make the very next sale at a price above 180. In 10 instances a seller sold a Regular to a buyer at a price above 180 and then sold to the same buyer within the next period at a price above 180 and without delivering a Super during the intervening time. The models predict that this will never occur.

Conclusion 8. A seller's demand depends not only upon his/her own "reputation" for delivering Supers, but also upon the market 'reputation." The Shapiro model (1982a), if it is to be generalized to multiple firms, must be changed to add a "market reputation" term.

Argument. The following model was estimated.

$$P_{iT} = C_1 + C_2 \cdot \sum_{t=1}^{T-1} (S_{i,T-t} \cdot A^{t-1}) + C_3 \sum_{t=1}^{T-1} (S_{\cdot,T-t} \cdot B^{t-1})$$

$$S_{\cdot,T-t} = \sum_{i=1}^{N} Si,t\text{-}t$$

P_{iT} = price received by seller i in period T

N = number of sellers

C_k,A,B = constants to be estimated

$S_{i,T-t}$ = number of Supers sold by seller i in period T - t so

$\sum_{t=1}^{T-1} (S_{i,T-t} \cdot A^{t-1})$ is *own reputation* = a weighted sum of all past Super sales by i.

$S_{\cdot,T-t}$ = number of Supers sold in the entire market in the period T - t, so $\sum^{T-1} (S_{\cdot,T-t} \cdot B^{t-1})$ is a *market reputation* = a weighted sum of all past Super sales.

The estimated coefficients are in Table 5. Data from experiments 19 (monopoly) and 21 (low cost Supers) were not included. The conclusion stated above is supported by the fact that five of the twelve experiments have a significant C_3 term.

Conclusion 9. The markets where individual sellers prices are not influenced by either their own reputation for selling Supers or the market reputation for Super sales do not exhibit either lemons behavior or full information behavior. Furthermore, significant influence of individual seller's own reputation on own prices is not a sufficient condition for reputation-induced efficiency gains.

Argument. In markets 12 and 13 neither C_2 nor C_3 is significant. Market behavior is not captured well by either of the two models. Both experiments 16 and 17 exhibited sensitivity to individual reputations, but neither exhibited substantial efficiency gains.

The A and B parameters measure the "discount" rate over time—whether past or most current Supers sales are most important. Those experiments for which A is small and C_2 significant suggest the importance is on the most recent individual behavior. Where B is small and C_3 is significant, the most recent market behavior seems to be the most important.

Conclusion 10. A reduction of buyer information about the grade deliveries of individual sellers from public information to private information decreases market efficiency.

Argument. Because of subject pool differences in market behavior, the only opportunity to reject the proposition occurred in the CIT experiments (left half of Table 2). Because efficiencies at both BU and PCC tended to be low, little opportunity existed for further efficiency losses. The public revelation of individual seller's decisions in market 11(1-8) at CIT produced efficiencies at near the 100 percent level. Efficiencies in experiment 11 (with public information and after period 4) dominate the efficiencies (after period 4) in both 15(1-10) and 20(1-9), where only private revelation of grade existed. The pattern of trades in the two private revelation markets is that suggested by the theory. Define a "ripoff" as a contract in which the price indicates that the (risk neutral) buyer was at least 90 percent confident that the unit would be a Super (270F), but a regular was delivered. In this case a Regular is delivered at a price of 270 francs or more. In market 11 a total of four ripoffs occurred while twenty-eight and fifteen ripoffs occurred in markets 15 and 20 respectively. While no tests are provided, both prices and efficiencies appear to be drifting downward in 15 and 20 and upward in 11.

Conclusion 11. A subject pool difference exists and subject experience makes a difference in market behavior.

Argument. The best example is between CIT 11 and PCC 16 and 17. These experiments had experienced participants, yet the market behavior of the PCC

Table 5

Regression Results

Seller Price as a Function of Own Reputation and Market Reputation

EXP	C_1	C_2	A	C_3	B
5	215.16	13.16	0.958	-1.34	1.257
(N=72)	(34.50)**	(4.62)**	(7.232)**	(1.45)	(4.503)**
6	170.15	5.44	0.101	26.01	0.220
(N=106)	(109.23)*	(0.74)	(0.083)	(8.40)**	(2.955)
7	72.54	37.58	0.72	15.37	0.300
(N=136)	(5.53)**	(6.22)**	(0.499)	(4.33)**	(2.278)
10	164.39	0.07	3.675	2.62	0.741
(N=59)	(103.99)**	(0.13)	(0.383)	(3.42)**	(4.964)**
11	283.49	13.46	0.547	-3.75	-0.741
(N=82)	(22.93)**	(2.46)**	(3.041)**	(1.78)	(0.173)
12	185.81	-1.85	1.314	3.34	1.093
(N=55)	(13.66)**	(0.29)	(0.729)	(1.09)	(2.714)**
13	205.35	-7.48	-0.221	2.55	0.145
(N=53)	(16.51)	(1.01)	(0.211)	(0.84)	(0.147)
14	186.97	-0.61	-1.318	5.37	0.205
(N=67)	(13.74)**	(0.14)	(0.555)	(2.25)*	(0.482)
15	206.92	22.36	0.367	4.55	-0.200
(N=108)	(11.71)**	(3.48)**	(2.190)*	(2.12)*	(-0.760)
16	177.87	23.17	-0.501	1.81	-0.023
(N=66)	(28.57)**	(2.79)**	(1.185)	(0.75)	(0.012)
17	191.21	38.97	-0.200	-2.02	-0.600
(N=62)	(47.70)**	(5.04)**	(0.955)	(0.63)	(0.501)
20	162.78	17.37	-0.054	3.79	0.728
(N=94)	(10.82)**	(2.46)**	(0.132)	(1.65)	(3.346)**

Numbers in parentheses are t-statistics.

* significant at .95 level

** significant at .99 level

group is lemons, while the behavior of CIT is full information. Again an interesting comparison exists between the inexperienced PCC participants of 12, 13, and 14, which is difficult to describe in terms other than overly optimistic buyers, with the behavior of 20, which used inexperienced CIT participants. Notice, however, that with some institutional arrangements, such as markets with unknown sellers and markets in which express warranties were enforced, the differences between subject pools is almost nonexistent.

Experience seems to be important in the PCC subject pool when seller identifications are known, and enforced warranties are absent. Compare 12, 13, and 14 with 16, 17, 18, 21. The purchases that can be characterized as made by overly optimistic buyers substantially disappear with experience. In brief, the models seem to work better as participant experience increases.

Conclusion 12. The time of revelation, whether revelation was made at the end of the period or immediately after the sale, made no difference.

Argument. Markets 5(BU) and 10(PCC) had inexperienced subjects and the revelation came at the end of a period. Markets 12(PCC), 13(PCC), 14(PCC) had inexperienced subjects and the revelation was made immediately after a sale. Market efficiencies are indistinguishable.

Experienced participants in 6(BU), 7(BU) with revelation at the end of the period can be compared to participants in 16(PCC), 17(PCC), 18(PCC) and 21(PCC) when revelation was made immediately after sale. Market 7 with the credence property had the highest efficiencies. The others are indistinguishable.

Conclusion 13. A reduction in the relative cost of Supers switched market behavior from that of the lemons model to that of the full information model.

Argument. On two occasions the relative cost of Supers was lowered, markets 19(7-14) and 21(5-14). Prior to the lowering of cost the markets were essentially at a lemons equilibrium. After the cost was lowered the number of Supers delivered increased significantly as did efficiencies and prices.

Conclusion 14. Aside from a possible small increase in price at first, nontruthful advertising had no effect on average price.

Argument. Two of the three cases, in which warranties were optional but unenforced, experienced a slight upward movement in price at first, 18(7), 19(4), but prices then returned to previous levels. The third case, 14(8) experienced no upward movement at all.

V. Summary and Interpretation of Results

The lemons phenomena can occur (conclusion 2). We are aware of no other clear documentation of its existence. Markets will not necessarily allocate information to the agent that values it most. Informational failure in a market can be observed. Of course this result alone says nothing at all about the likelihood of informational failure in naturally occurring markets. The result is important because it demonstrates that the tools and theories used to analyze naturally occurring markets were not rejected when put to an important test.

With the existence of the lemons problem documented, the analysis turns to an examination of the conditions that generate it. The lemons phenomena do not automatically go away when firms have an incentive and opportunity to establish a reputation for good quality (conclusion 5). Reputation and brand names are not sufficient devices to guarantee efficient market operation even in the case of experience goods and repeat purchases. Several factors can operate to frustrate the competitive development of reputations. First, the cost of developing a reputation is evident in several markets. Supers must sell at regular prices in sufficient quantity to attract buyers' attention and develop their confidence in the seller's reliability. Of course, this can generate substantial temporary losses. The problem can occur because the market price response must be sufficiently rapid to reward sellers who adopt a strategy of delivering high-grade units, and this price response is not well understood. In fact, the positive responsiveness axiom that states that super deliveries will be rewarded by higher prices or increased demand is not always reliable. This axiom is at the heart of many models as in Klein and Leffler (1981), Nelson (1974), Peltzman (1981), Schmalensee (1978), Shapiro (1980, 1982abc).Buyers might not even respond positively to high-grade deliveries (conclusion 9). Instead of understanding seller motivations or believing that sellers have an interest in reputation development, buyers might regard sellers as being totally random or buyers might even avoid sellers who deliver Supers on the belief that the sellers were attempting to trap the buyer or lure the buyer into paying a high price and then delivering a regular. While we cannot actually document the existence of such extreme buyer skepticism, some of the markets seemed to have that characteristic (*i.e.*, markets 12, 13, 14), and in some cases it might even be justified. In summary, buyer reactions to poor quality deliveries are not as uniformly predictable or as punishing and rewarding as presupposed by some dynamic models such as the quality guaranteeing price (conclusion 7).

How a policy might alter belief and learning processes or even buyer reactions to seller strategies is an open question. Marketing programs or regulatory policies that "properly frame" the problem that buyers face might be important. Conceivably the very existence of some sort of regulation, even if unenforced, is a type of public information that might foster buyer confidence in seller intentions and also foster seller beliefs about buyer reactions to "ripoffs." With such changed beliefs the market would possibly provide the proper rewards for quality such that further regulation would be unnecessary. Because the participants have incomplete information, multiple equilibria might exist, and the existence of multiple equilibria might be the source of confirming results. A publicly stated regulation might serve as a focal point that coordinates actions toward one of the equilibria. At this point theory provides very few hints and the issue is appropriate for more experimentation.

Buyer confidence and learning is just part of the reputation cost problem. The confidence must be translated into price increases. Once buyers recognize a seller who reliably delivers Supers, the price of that seller's units must adjust sufficiently rapidly to reduce the reputation cost. Data from experiments (Plott, JEL 1982) leads one to suspect that this price adjustment property is sensitive to market organizational features independent of any learning properties of buyers. The cost of reputation development depends upon the speed of price adjustment in response to changed buyer beliefs. Price adjustment speed appears to be related to market organization. Therefore reputation costs and the resulting evolution of quality products might be sensitive to market organization. Thus, empirical reasons exist for economists to have some interest in market organization, in addition to the theoretical propositions about the relationship between quality and market organization developed by Wilson (1980).

A second factor that can prevent reputation development from guaranteeing market efficiency is a type of externality that seems to exist in some markets. Individual seller success can be related to a market reputation for delivering Supers as well as to the individual seller's own reputation (conclusion 8). The externality can work negatively in two ways. First, individual sellers have an incentive to free ride on the reputations and markets developed by others. After one or two sellers have incurred the cost of reputation development and are successfully selling supers at a high price, an entrant can coat-tail on their reputations. Buyers will test units of entrants priced just below the price at which Supers are being sold (price is a signal) and if the entrant delivers Supers, its reputation is almost costlessly established. The free rider aspect can dampen the development of reputations and the resulting market efficiency. The externality also can work negatively on a seller that has an established reputation. If other sellers decide to destroy their reputation by dumping regulars and thereby make a profit on the ripoffs, buyer reaction can be negative toward all firms. Even sellers that continue to deliver Supers can experience a drop in demand as buyers appear to become suspicious of all firms. This negative externality can depress the returns from reputation development. Whether or not alternative market organization, public announcements, or regulations can effectively promote the development of quality by reducing such externalities awaits further experimentation.

A third potential problem is structural and derives from the problems discussed above. If quality improvements can only be achieved by large and discrete increases in cost, markets might equilibrate at local equilibria that have the lemons property. The large discrete increases in cost mean that the cost of quality improvements can be covered only by large changes in price. Either the buyer must be willing to take a risk and pay a premium in hope that the seller will deliver a Super or the seller must incur large losses by selling Supers at regular prices in order to build buyer confidence. Risk aversion on both sides will make reputation development and the resulting high quality difficult. In markets in which the relative cost of Supers is lowered, the instances of super sales and resulting reputation-like behavior in the market becomes much more pronounced (conclusion 13).

A final problem also derives from the others listed above when supplemented by the fact that multiple markets are involved. If buyers are optimistic and bid prices high even in the face of many regular deliveries, sellers have no incentive to develop a reputation for delivering Supers. The difference between the going prices of units that are being delivered as Regulars and the maximum value that one might get from a Super is not enough to cover the cost differential (*see* markets 12, 13). Before a reputation is worthwhile, buyer optimism must be dampened and the prices must fall to a point that makes reputation development profitable. Complete market quality deterioration, all lemons, might be a necessary condition for automatic market recovery. Commentators with a taste for paradoxical statements could say that things cannot get better until they get worse; or regulation is needed least when market performance is at its worst.

Market reputation development may be difficult in some circumstances, but it is certainly not impossible. In some of the markets knowledge of seller identification alone (brand names) was sufficient to guarantee behavior consistent with the full information model (conclusions 5 and 13). This opens a possible role for third party actions that facilitate such reputation development. Reputation development is clearly a tool but we do not know its exact limitations. Enforced warranties also will induce the market behavior that is captured by the full information model (conclusion 1). Markets that are otherwise behaving in a confusing and inefficient manner recover almost immediately when enforced warranties are introduced. The power of the instrument in fostering market efficiency is remarkable.

Legal instruments or practices that have the effect of a costlessly enforced warranty will be voluntarily offered by sellers. Such warranties, if they exist, will also be voluntarily demanded by buyers (conclusion 3). Such instruments require that any disclosures made are truthful. Competition, in turn, forces disclosures. The data in these experiments suggest that the Grossman/signaling models that predict the voluntary use of such instruments (when their availability is publicly known) are reliable in this respect as models of warranty-like instruments. We are thus not too far from an understanding of the process through which the warranty-like instruments have an effect on markets. Further support for this type of theory has substantial ramifications for regulatory policy because a direct implication of the theory, when applied to experience goods as opposed to credence goods, is that mandatory disclosure is unnecessary.

Markets need not be characterized by either the full information model or the lemons model. The reasons for such confusing behavior are not understood. Of course one can speculate that it reflects a lack of sophistication on the part of market participants or a lack of experience, or a number of things idiosyncratic to the population (conclusion 11). The problem could be due to the existence of multiple Bayes equilibria as was mentioned above. These are just speculations that call for more detailed investigation. Precisely because the behavior of such markets is not understood, it is necessary for policy analysts to know when standard principles can only be applied with substantial precautions. Markets that behave in understood patterns are characterized by either private reputation formation or market reputation formation or both (conclusion 9).

Our final observation is related to advertising. False advertising exists in our markets (conclusion 4) even when buyers quickly and easily detect the deception. Thus policy analyses (Posner 1973, 1979) or models (Nelson 1970, 1974) that imply that false advertising cannot be sustained or will be beneficial are not supported by our results. We hasten to add that conditions relied upon by Nelson were not present in our markets and invites further experimentation.[13] Though false advertising occurred and the effects were not beneficial, the effects (for experience goods) are not as deleterious as presupposed by some advocates of advertising regulations. People are not misled. They simply dismiss all sellers' claims so that advertising fails to provide effective information which could enhance efficiency. This last finding may provide some insight into the advertising industry's strongly voiced support for the FTC's advertising substantiation program.[14]

Appendix

Experiment Instructions

GENERAL

This is an experiment in the economics of market decision making. Various research foundations have provided funds for this research. The instructions are simple and if you follow them carefully and make good decisions you might earn a considerable amount of money, which will be paid to you in cash at the end of the experiment.

In this experiment we are going to conduct a market in which some of you will be buyers and some of you will be sellers in a sequence of market days or trading periods. Attached to the instructions you will find some sheets, labeled Buyer or Seller, which describe the value to you of any decisions you might make. *You are not to reveal this information to anyone.* It is your own private information.

The type of currency used in this market is francs. All trading and earnings will be in terms of francs. Each franc is worth ______ dollars to you. Do not reveal this number to anyone. At the end of the experiment your francs will be converted to dollars at this rate, and you will be paid in dollars. Notice that the more francs you earn, the more dollars you earn.

[13] The conditions are that advertising is costly and sellers can increase market share.

[14] *See Advertising Age*, November 1, 1980, p. 40; *Television/Radio Age*, November 29, 1982, p. 35.

SPECIFIC INSTRUCTIONS TO BUYERS

During each market period you are free to purchase from any seller or sellers as many units as you might want. The value of a unit depends upon its grade. There are two grades (Regular and Super) and the value of a Super is much greater than the value of a Regular. At the time you buy a unit you will not know the grade but *(at the end of a trading period)* *(after the purchase)* you will be told the grade of each unit you bought.

The attached information and record sheet will help you determine the value to you of any decisions you make. Page ____ of your information and record sheets contains two schedules. The schedule in the left column identifies the redemption values of Regulars and the schedule in the right hand column contains the redemption values for the Supers. The redemption value of the first Regular you purchase is in the row marked First Units and the column marked Regular. The redemption value of the first Super you purchase is found on the same row only, under the column marked Supers. The redemption value of second units are found in the second row, etc. The profits from each purchase (which are yours to keep) are computed by taking the difference between the redemption value and the purchase price of the unit bought. That is,

your earnings = (redemption value) - purchase price.

In addition to these earnings you will receive a capital payment of ________ francs each period.

Suppose, for example, the redemption value for your first Regular is 1000 and the redemption value for your first Super is 4000. If you buy two units at 1200 and one is a Regular and one is a Super your profits are

1000 - 1200 =	-200
4000 - 1200 =	2800
TOTAL	2600

Turn now to the second page of the information and record sheet. The purchase price of the first unit you purchase should be listed in row two for the first unit purchased. The purchase price of the second unit should be listed in row 2 of the second unit, etc. When the grades of units become known you should enter the redemption values in rows 1 for each unit. If, for example, your first unit purchased is a Super and if your second purchase is a Regular, you record the redemption value for the first Regular because even though the unit is the second purchase it is only your first Regular. Profits at the end of the period should be recorded at the bottom of the page.

SPECIFIC INSTRUCTIONS TO SELLERS

During each market period you may sell to any buyer or buyers as many as ________ units. There are two types of units, Supers and Regulars. Each Super will cost you ________ and each Regular will cost ________. Notice that the cost of Supers is more than the cost of Regulars. The profits or losses on each sale (which are yours to keep) are computed by taking the difference between the price at which you sold the unit and its cost.

Your total profits for a market period are computed by adding the profits or losses on each sale during the period. The attached record sheet will help you keep track of your profits or losses. Enter the price of the first unit you sell in the appropriate column (Super or Regular) in row 1 at the time of sale. Then record the profit or loss as directed in row 3. The sale price of the second unit should be listed in the appropriate Super or Regular column in row 4. Profits should be similarly calculated and the total for the period recorded in row 16. All profits over ________ are yours to keep.

MARKET ORGANIZATION

The market for this commodity is organized as follows. The market will be conducted in a series of trading periods. Each period period lasts for at most ________ minutes. Any buyer is free at any time during the period to make a bid to buy the commodity at a specified price, and any seller with units to sell is free to accept or not accept the bid. Likewise, anyone wishing to sell a unit is free to make an offer to sell one unit at a specified price. All bids and offers are entered on the blackboard and remain there until accepted or canceled. If a bid or offer is accepted, a binding contract has been closed for a single unit at the specified price and the contracting parties will record the contract price. Any ties in bids or acceptances will be resolved by random choice. Except for the bids and their acceptance or cancellation you are not to speak to any other subject. There are likely to be many bids that are not accepted, but you are free to keep trying. You are free to make as much profit as you can.

Trading period 0 will be a trial period to familiarize you with the procedure, and will not count toward your cash earnings.

FINAL OBSERVATIONS

1. (*At the end of the period*) (*After each sale*) sellers indicate to the experimenter those trades that involved Regulars and those that involved Supers. This information will be transmitted to the buyers who participated in those transactions. Buyers can then calculate their profits.

2. Each individual has a large folder. All papers, instructions, records, etc, should be put into this folder. Leave the folder with us before leaving. *Take nothing home with you.*

3. We are able to advise you a little on making money. First, you should remember that pennies add up. Over many trades and a long period of time very small amounts earned on individual trades can add up to a great deal of money. Secondly, you should not expect your earnings to be steady. You will have some good periods and some bad periods. During bad times try not to become frustrated. Just stay in there and keep trying to earn what you can.

It all adds up in the end.

Some people rush to trade. Others find it advantageous to "shop" or spread their trading over the period. We are unaware of any particular "best" strategies and suggest that you adapt accordingly.

4. Under no circumstances may you mention anything about activities which might involve you and other participants after the experiment (*i.e.*, no physical threats, deals to split up afterwards, or leading questions).

5. Each individual will be paid in private. Your earnings are strictly your own business.

6. Buyers tender bids verbally by indicating in sequence "(buyer number) BIDS (amount)."

7. Sellers tender offers verbally by indicating in sequence "(seller number) OFFERS (amount)."

8. Each trade in a period will be numbered. (*At the end of the period*) (*after each sale*) each seller will (*submit a slip of paper*) (*hold up a card*) for each trade specifying a Super or a Regular. The seller is free to determine the grade of the units he sells and may mix grades within or between periods.

Trader No. ________

R S

Buyer ______ Seller______

INFORMATION AND RECORD SHEET

RECORD SHEET FOR SELLER NO._____

Unit Sold			Market Period						
			1		2			18	
			Super	Regular	Super	Regular		Super	Regular
1st	1	sales price							
	2	cost					. . .		
	3	profit/loss (row 1-row 2)							
2nd	4	sales price							
	5	cost							
	6	profit/loss (row 4-row 5)							
			⋮					⋮	
	13	total profit/ loss per item							
	14	total profit/ loss					. . .		
	15	deduction per period							
	16	net profit							

INFORMATION AND RECORD SHEET

RECORD SHEET FOR BUYER NO.______

Unit Purchased			Market Period							
			1	2	3	4	5		17	18
1st	1	redemption value								
	2	purchase price								
	3	profit/loss (row 1 - row 2)						. . .		
2nd	1	redemption value								
	2	purchase price								
	3	profit/loss (row 1 - row 2)								
				⋮					⋮	
5th	1	redemption value								
	2	purchase price								
	3	profit/loss (row 1 - row 2)						. . .		
	1	period profit								
	2	capital payment								
	3	total								

Name______________________ Soc. Sec. No.______________ Total Payment_______

Address____

INFORMATION AND RECORD SHEET

	Regular	Super
1st units redemption value		
2nd units redemption value		
3rd units redemption value		
⋮	⋮	⋮
8th units redemption value		
Capital payment per period_________		

References

Advertising Age, November 1, 1980, p. 40.

Akerlof, George A. "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* August 1970, 84, 488-500.

Federal Trade Commission Staff, *Sale of Used Motor Vehicles*, Bureau of Consumer Protection 1978.

_____________, *Life Insurance Cost Disclosure*, Bureaus of Consumer Protection and Economics, July 1979.

Grossman, Sanford J. 1981, "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics* December 1981, 24, 461-484.

Klein, Benjamin and Leffler, Keith B., "Non-Governmental Enforcement of Contracts: The Role of Market Forces in Guaranteeing Quality." *Journal of Political Economy* August 1981, 89, 615-641.

Leland, Hayne, "Comments on Grossman", *Journal of Law & Economics* December 1981, 24, 485-490.

Miller, Ross M. and Plott, Charles R., "Product Quality Signaling in Experimental Markets," California Institute of Technology Social Science Working Paper no. 447, 1982.

Nelson, Phillip, "Information and Consumer Behavior", *Journal of Political Economy*, March/April 1970, 78, 311-329.

_____________, 1974. "Advertising as Information", *Journal of Political Economy* July/August 1974, 82, 729-754.

Peltzman, Sam, "The Effects of FTC Advertising Regulation," *Journal of Law and Economics* December 1981, 24, 403-448.

Pitofsky, Robert, 1979. "Mandated Disclosure in the Advertising of Consumer Products", "Commentary" (on Posner) and "Dialogue," in *Business Disclosure: Government's Need to Know*, edited by Harvey Goldschmid, New York: McGraw-Hill, 1979, 311-330, 360-366, 367-376.

Plott, Charles R., "Industrial Organization Theory and Experimental Economics." *Journal of Economic Literature* December 1982, 20, 1485-1527.

_____________ and Smith, Vernon L., "An Experimental Examination of Two Exchange Institutions," *Review of Economic Studies* February 1978, 45, 133-153.

Posner, Richard, "Regulation of Advertising by the FTC." American Enterprise Institute Evaluative Studies no. 11, 1973.

_____________, "The Federal Trade Commission's Mandated-Disclosure Program: A Critical Analysis," and "Dialogue," in *Business Disclosure: Government's Need to Know*, edited by Harvey Goldschmid, New York: McGraw-Hill, 1979, 331-366, 367-376.

Rogerson, William, "Reputation and Product Quality." Stanford University Institute for Mathematical Studies in the Social Sciences, Economic Series, Technical Report no. 384, revised October, 1982.

Rothschild, Michael and Stiglitz, Joseph, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economic* August 1976, 90, 629-649.

Schmalensee, R., "Advertising and Product Quality," *Journal of Political Economy* June 1978, 86, 485-503.

Shapiro, Carl, "Advertising and Welfare: Comment," *Bell Journal of Economics* Autumn 1980, 11, 749-752.

__________, "Consumer Information, Product Quality and Seller Reputation," *Bell Journal of Economics* Spring 1982a, 13, 20-35.

__________, "Product Differentiation and Imperfect Information: Policy Perspectives." Federal Trade Commission, Economics Bureau Working Paper no. 70, July 1982b.

__________, "Premiums for High Quality Products as Rents to Reputation." Federal Trade Commission, Economics Bureau Working Paper no. 43, 1982c.

Smith, Vernon L., "Microeconomic Systems as an Experimental Science," *American Economic Review* December 1982, 72, 923-955.

Spence, A. Michael, 1977. "Consumer Misperceptions, Product Failure, and Producer Liability," *Review of Economic Studies* October 1977, 44, 561-562.

Television/Radio Age November 29. 1982, p. 35.

Wilson, Charles, "The Nature of Equilibrium in Markets with Adverse Selection", *Bell Journal of Economics* Autumn 1980, 11, 108-113.

An Experimental Study of Warranty Coverage and Dispute Resolution in Competitive Markets

Thomas R. Palfrey and *Thomas Romer*
Graduate School of Industrial Administration
Carnegie-Mellon University

1. Introduction

In service and product markets where warranties are offered, disputes over warranty performance frequently occur between buyer and seller. Resolving such disputes in a fair and effective way has become an increasingly important and controversial question in recent years. Some observers have gone so far as to argue that the pervasiveness of such disputes and the inability to resolve them effectively is having a corrosive effect on society.[1] This is probably somewhat extreme, but even a less excited perspective suggests that the design of procedures to handle consumer disputes is a matter for serious concern.

The design of appeals mechanisms for resolving consumer disputes has centered largely on procedural issues and features of internal functioning. Mechanisms are evaluated on the basis of comparisons that involve the relative cost of the procedures, how the burden of costs is distributed, and the ability of the procedure to make the "right" decision in specific cases. Consumer advocates, for example, have suggested that more desirable outcomes can be achieved if sellers are forced to bear *all* litigation costs relating to a dispute and its subsequent adjudication by a "dispute resolution mechanism" (DRM). Much attention has focused on the need for having a DRM evaluate the facts accurately and neutrally, and the importance of procedural fairness.[2] The overall tenor of discussion has generally been that the primary consideration is with the welfare of the disputants: Given that a dispute has occurred, what is the best way to resolve it?

We have argued elsewhere [Palfrey and Romer (1983)], that, at least in service and product markets with warranties, the welfare of consumers and sellers who are *not* involved in disputes should be considered as well. Specifically, we claim that the choice of rules for a particular DRM can affect prices, the range of warranties offered, and the likelihood of disputes in these markets, as well as their direct effect on outcomes when disputes do actually occur. Thus, the design of DRMs will not only affect disputants, but will also indirectly affect non-disputants.

Unfortunately, data are not available to assess the effects of DRMs on these markets and on the welfare of buyers (both disputants and nondisputants) in these markets in a systematic way. The purpose of this study is to provide some

[1] *See, e.g.*, Nader (1980).
[2] *See* FTC (1980) and Ray (1983).

empirical evidence about the effects of DRMs on the equilibrium in markets where warranties may be offered. In order to do this, it was necessary first to develop a technique for conducting laboratory markets with uncertain product quality and warranties, which allowed for the possiblity of disputes and which could later be adapted to include appeals procedures to deal with disputes over warranties. The type of uncertainty in our markets and the combination of a "market" phase and a "warranty" phase in these experiments represent significant departures from standard experimental methodology. Our approach can also be useful in studying other problems in the economics of uncertainty. For example, DeJong, Forsythe, and Uecker (1984) have been able to combine our procedures with some novel features of their own to study the effects of auditing rules in a principal-agent setting.

Using our methods, we designed and carried out a series of laboratory market experiments to study the menu of prices and warranties offered in a competitive market for a product (or service) of random quality. Since our ultimate goal is to investigate the effects of DRMs on markets outcomes and buyer welfare, we included several experiments in which buyers could appeal under different rules. These rules varied in an important component: how the cost of the DRM was shared between the disputants (seller and buyer). Some hypotheses about the experimental outcomes—*e.g.*, prices, warranties, buyer welfare and frequency of disputes—are obtained from a simple analytical model of warranties, and are evaluated using the data from these laboratory markets.

In the next section, we give an overview of the experiment setting. In Section 3, we provide details of the experimental design. Section 4 presents an analytical framework for deriving hypotheses about experimental results. In Section 5, we enumerate some hypotheses. Section 6 presents the data from the experiments and discusses the results. Our concluding remarks are in Section 7.

2. Overview of Experimental Setting

In Palfrey and Romer (1983), we developed a theoretical framework for analyzing a market in which warranties are offered and disputes over warranty performance occur. We argued that the key feature of such markets is that buyers and sellers have differential and imperfect information. In order for warranties to exist, there must be some uncertainty about (future) product quality. In order for disputes to occur, it must be possible, at least in some circumstances, for buyers and sellers to disagree about whether performance by the seller is called for under the warranty covering the product. There is a wide variety of ways that such disagreements can occur, but they reduce to two essential cases:

- Buyer and seller may agree about warranty terms, but disagree about whether observed product characteristics trigger warranty performance.
- Buyer and seller may agree about product characteristics, but disagree about whether the claimed product defect is covered by the warranty.

In Palfrey and Romer (1983), we noted that, to a large extent, these two causes of disputes can be formally equivalent in their impact on warranty markets. The important point is that disputes occur when buyer and seller *perceive* (or *claim to perceive*) the situation differently. For concreteness, however, our experiments are structured so that potential disagreements concern perceived product quality, not what the warranty covers. In our laboratory markets, at the time a sale is made, buyers and sellers are uncertain about product quality. After purchase, a buyer gets a noisy signal about product quality. Depending on the type of warranty covering the product, the buyer may then make a claim to the seller for warranty performance. Again, depending on the type of warranty, the seller may automatically honor the buyer's claims, or may condition his performance on a noisy signal that he obtains about product quality. If the market structure has no DRM, then the seller's decision is the final one. If there is a DRM in place, and there is a disagreement between buyer and seller, the buyer may choose to appeal to the DRM, whose decision is final and binding. The DRM, too, may be subject to error. In the experiments, its decisions are based on a noisy signal about product quality.

In many markets, transaction costs may be quite important. This is particularly relevant in considering DRMs. Our experiments allow for transaction costs both in making warranty claims and in appealing to the DRM (if there is one). Moreover, these costs maybe different for buyers and sellers.

One of the questions of interest is whether the availability (or mandating) of a DRM has an effect on the type of warranty offered in the market. For example, sellers may wish to avoid costs associated with a DRM by offering a warranty under which disputes are extremely unlikely or do not occur at all. (*E.g.*, sell only "as is" or, at the other extreme, adopt a "customer is always right" policy.) Or, the availability of a DRM may make buyers more willing to purchase products with intermediate or "limited" warranties, knowing that, in the event of a dispute, the seller's word is not the final one. To capture such effects, our laboratory market allows for the possibility of more than one warranty type.

While a variety of warranties are conceivable, we restrict our attention to replacement policies; that is, the buyer is given a replacement unit if a claim is made *and* the seller agrees that the claim is a valid one. We distinguish between an unconditional (which we loosely refer to as "full") warranty and a conditional (which we loosely refer to as "limited") warranty. With an unconditional warranty, the seller always agrees that the buyer's claim is valid; there are "no questions asked." We call this a *Type 1* warranty.

With a conditional warranty, a seller promises to perform only if the unit is of low quality. Under such a warranty, a seller "inspects" the unit if a claim has been made, and then honors the claim if the "inspection" tells the seller it is a low quality unit. He refuses the claim if the "inspection" indicates it is not a low quality unit. Such warranties are *Type 2* warranties.

The seller may also sell a unit "As Is." This null warranty is called *Type 3*.

More specifically, our laboratory markets involved the buying and selling of a "commodity." The quality of a given unit of this commodity is either high (a Good unit) or low (a Defective unit). A high quality unit yields a value V_h to the buyer, while a low quality unit generates a value $V_l < V_h$. Before a unit is sold, neither buyer nor seller knows the true quality of the unit, but the probability that a unit is good, t, is common knowledge to all parties. Each unit is drawn independently from the same distribution, regardless of the seller from whom it is bought. The cost to the seller of a unit is c.

After purchasing a unit, a buyer receives a signal about the quality of the unit. This signal is either 0 or 1. The probability, q, of receiving the 0 signal, conditional on the unit being Defective, equals the probability of receiving the 1 signal, conditional on the unit being Good. We set $q > 0.5$, so that 0 is a (noisy) signal of low quality, and 1 is a (noisy) signal of high quality. Thus q is the probability of a correct signal.

If the unit was purchased under a Type 1 or Type 2 warranty, then the buyer may return the unit to the seller from whom it was purchased. If he does so, he is charged a return fee c_b, and the seller is charged a return fee c_s.

Under a Type 2 warranty, if a unit is returned, the seller obtains a signal about the unit's quality. All buyer and seller clues are drawn from the same distribution and are independent, conditional on true quality. A seller is obligated to honor the warranty if it is a Type 1 or if he receives a 0 signal under a Type 2 warranty. An important and limiting feature of our experimental design is that, under a Type 2 warranty, sellers are forced to act strictly on the outcome of their signal, which is public knowledge. Thus, our formulation rules out the possibility that a seller might offer a warranty and then always refuse claims, regardless of his signal about product quality. Such fraudulent behavior may underlie the calls for dispute resolution in some markets.[3] Our primary intent, however, is to focus on disputes that may emerge due to imperfections of information, rather than disputes resulting from purely fraudulent behavior. In future work, we intend to investigate DRMs in markets where sellers have the option of reneging.[4]

If there is a DRM available, then a buyer whose claim under a Type 2 warranty is refused has the option of appealing. If he appeals, he is charged an appeal fee αc_m and the seller is charged an appeal fee βc_m. Our "mechanism" is, in fact, quite mechanical. Rather than incorporating all the intricacies of decision-making by a judge, arbitrator, jury, or panel, we collapse the decision-making process into yet another noisy signal. In the event a dispute is brought to the DRM, a signal is drawn, independently of buyer and seller signals. The probability $\hat{q}$ that the signal is 0, conditional on the unit being Defective, equals

[3] Indeed, the FTC has mandated DRMs as part of the settlement in a variety of cases involving systematic failure to honor warranties.

[4] Lynch *et al.* (1984) in their study of product quality also have warranties that are "perfectly enforceable." We believe the possibility of fraudulent behavior is important in markets with warranties, and may be significantly affected by DRMs. This issue is beyond the scope of this paper.

the probability that the signal is 1, conditional on the unit being Good. If the DRM signal is 0, the seller must perform under the warranty. If it is 1, the buyer must keep the original unit.

3. Experimental Design: Details

We conducted eight experiments. Four of the experiments had no DRM, and four experiments included a DRM.[5] Subjects were undergraduate students at Carnegie-Mellon University, none of whom had any previous experience as research subjects in experimental markets. Each subject participated in only one experiment. Each experiment consisted of nine or ten subjects, four of whom were designated sellers, with the remaining subjects designated as buyers. Each experiment lasted approximately three hours. The average amount a subject earned in these experiments was approximately $30, and there was considerable variation in earnings across subjects.

At the beginning of each experiment, after the instructions were read, a demonstration of how the random variables in the model worked was given. This was followed by a short test, to make sure that all subjects understood the rules and how to keep records. The instructions and the test differed slightly depending upon whether or not the experiment used a DRM. Both sets of instructions are included in the Appendices.

The parameters of a given experiment without a DRM consist of a specification of $\{t,q\}$ for the market, a specification of $\{V_h,V_l,c_b\}$ for each buyer, a specification of $\{c,c_s\}$ for each seller, and a specification of a maximum number of units each buyer could purchase in each market period. In all experiments, buyers were limited to two purchases per period. (This was public information.)

The values of t and q were publicly announced at the beginning of the experiment. All other parameters were private information to the particular subject. No subject was given any information about other subjects' values of these parameters, and direct communication of private information was prohibited. In fact, all sellers' parameters were identical and all buyers' parameters were identical.

In the experiments with a DRM, the precision parameter of the DRM, $\hat{q}$, was publicly announced at the beginning. Each buyer was privately informed of αc_m and each seller was privately informed of βc_m.

In all eight experiments, we set $t = 0.5$ and $q = 0.75$, and in the four DRM experiments, we set $\hat{q}(= q) = 0.75$. The complete parameter structures for the eight experiments are given in Table 1 (Experiments 1-4) and Table 2 (Experiments 5-8). Note that experiments 1 and 2 share the same parameters as do experiments 3 and 4, and, except for α, β, and c_m, experiments 5-8 have the same parameters.

[5] In addition, we conducted two pilot experiments in which we tested the clarity of the instructions and the manageability of our procedures.

Table 1

PARAMETERS FOR EXPERIMENTS 1-4 (NO DRM)

	Exp. 1	Exp. 2	Exp. 3	Exp. 4
t	0.50	0.50	0.50	0.50
q	0.75	0.75	0.75	0.75
V_h	6.00	6.00	8.00	8.00
V_l	0.00	0.00	0.00	0.00
c	1.20	1.20	2.40	2.40
c_b	0.10	0.10	0.10	0.10
c_s	0.10	0.10	0.10	0.10
Buyer Reservation Price (assuming buyer uses optimal return policy)				
Type 1 ("Full")	3.70	3.70	4.95	4.95
Type 2 ("Limited")	3.70	3.70	4.95	4.95
Type 3 ("As Is")	3.00	3.00	N.A.	4.00
Seller Expected Cost (assuming buyer uses optimal return policy)				
Type 1 ("Full")	1.85	1.85	3.65	3.65
Type 2 ("Limited")	1.62	1.62	3.20	3.20
Type 3 ("As Is")	1.20	1.20	N.A.	2.40
No. of Buyers	5	6	6	6
No. of Sellers	4	4	4	4

Table 2

PARAMETERS FOR EXPERIMENTS 5-8 (DRM)

	Exp. 5	Exp. 6	Exp. 7	Exp. 8
t	0.50	0.50	0.50	0.50
q	0.75	0.75	0.75	0.75
V_h	6.00	6.00	6.00	6.00
V_l	0.00	0.00	0.00	0.00
c	1.20	1.20	1.20	1.20
c_b	0.10	0.10	0.10	0.10
$\hat{q}$	0.75	0.75	0.75	0.75
c_m	0.00	0.80	0.80	0.80
α	0	0	1	0
β	0	1	0	1
	Buyer Expected Value			
Type 1 (Strategy S_8)	3.70	3.70	3.70	3.70
Type 2 (Strategy S_6)	3.70	3.70	3.70	3.70
(Strategy S_5)	3.79	3.79	3.64	3.79
(Strategy S_7)	3.84	3.84	3.69	3.84
	Seller Expected Cost			
Type 1 (Strategy S_8)	1.80	1.80	1.80	1.80
Type 2 (Strategy S_6)	1.58	1.58	1.58	1.58
(Strategy S_5)	1.91	2.06	1.91	2.06
(Strategy S_7)	1.69	1.84	1.69	1.84

Each experiment was divided into a number of trading periods, with the parameters of the market held constant across all periods of the experiment. Between eight and eleven periods were run, within a three hour time limit.

Experiments differed in the types of warranties available for sellers to offer buyers. In Experiments 1 and 2, sellers could offer all three warranty types simultaneously. After running these experiments, we were concerned that our design did not permit subjects sufficient experience with the stochastic structure of Type 1 and 2 warranties. Consequently, in Experiment 3, only Types 1 and 2 were permitted. In Experiment 4, only Types 1 and 2 were allowed in the first four periods. Beginning with period five, Type 3 warranties were introduced, and subjects were told that sellers could now offer these in addition to Types 1 and 2. Because the stochastic structure in the DRM experiments was so complicated, in all four of the DRM experiments (Experiments 5-8), only Types 1 and 2 warranties were permitted, thereby allowing subjects as much experience with these warranties and the DRM as possible.

Market Organization in Experiments 1-4 (no DRM)

The first four markets were organized as follows. Each experimental period was divided into two stages. *Stage 1* was an oral auction in which sellers offered units for sale by announcing price-warranty combinations. Sellers were permitted to offer units for sale *simultaneously* at different prices and with different warranties. As a result, there were multiple markets (one for each warranty type) operating simultaneously. Buyers did not make bids. They simply indicated willingness to accept an offer. Although each buyer was permitted to purchase at most two units per period, there was no restriction on how many units a seller could sell. When an offer was accepted, the buyer and seller were asked to record on their record sheets (*see* Appendix A) the price, warranty, and identity of the other party to the transaction. Stage 1 ended when (a) all buyers bought two units, or (b) the preannounced time limit (usually 7 minutes) ran out.

At the beginning of *Stage 2*, each buyer was given a "quality clue" for each unit he or she had purchased in stage 1 of that period. These clues were private information. Buyers with units that had Type 1 or 2 warranties were then asked, in turn, whether they wished to return.[6] Whenever a warranty claim was made, the buyer and seller of that unit were privately charged c_b and c_s, respectively. If a claim was made under a Type 1 warranty, the buyer was given a replacement unit and gave up the original unit. The seller was charged c for the replacement (in addition to the cost c of the original unit). When a claim was made under a Type 2 warranty, the seller of the unit was given a "quality clue." This clue mandated the seller's behavior. If the clue was "0" ("low quality"), the seller had to replace the unit. If the clue was "1" ("high quality"), the buyer had to keep the original unit. Units purchased "As Is" (Type 3) could not

[6] These return decisions were public information.

be returned. Replacement units could not be returned, either, regardless of the original warranty type.

At the end of Stage 2, buyers were informed of the true quality of each unit (original or replacement, as the case may be). Buyers' payoffs were determined by the true quality of the units they held at the end of each period, *less* the price they paid, *less* return costs. Sellers' payoffs were given by their revenues, *less* "production costs," *less* return costs.

Figure 1 shows all possible sequences of decisions and events that can occur between the time a unit is purchased and the time true quality is announced.

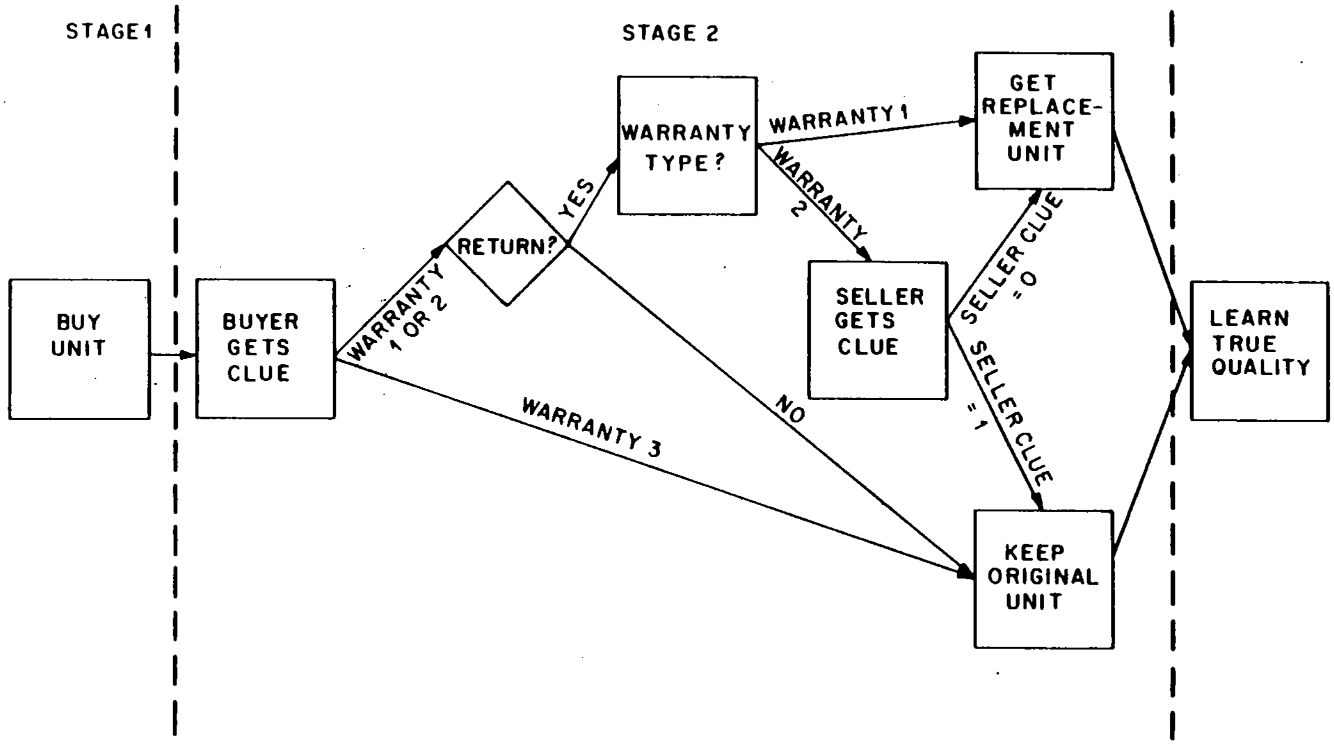

STAGE 1
STAGE 2
BUY UNIT
BUYER GETS CLUE
WARRANTY 1 OR 2
WARRANTY 3
RETURN?
YES
NO
WARRANTY TYPE?
WARRANTY 1
WARRANTY 2
SELLER GETS CLUE
SELLER CLUE = 0
SELLER CLUE = 1
GET REPLACEMENT UNIT
KEEP ORIGINAL UNIT
LEARN TRUE QUALITY

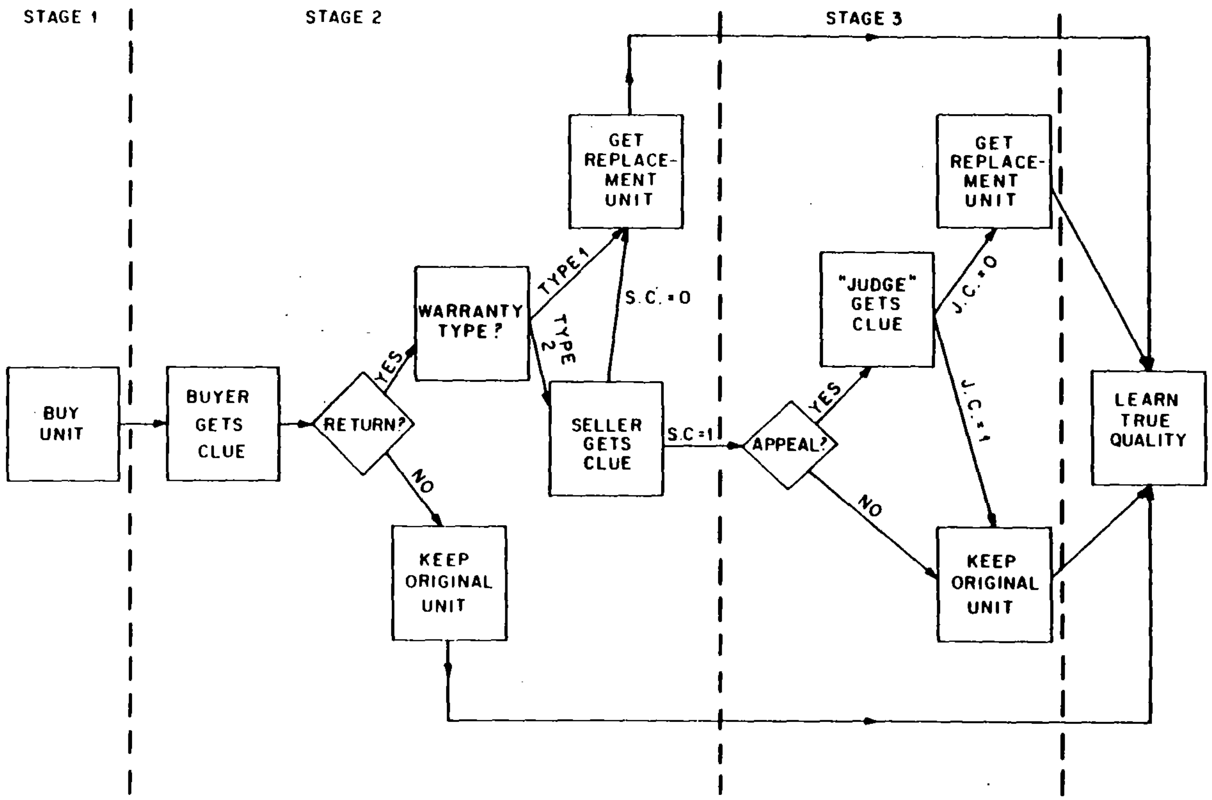
Figure 1

Market Organization in Experiments 5-8 (with DRM).

In these experiments, each trading period consisted of *three* stages. The first two stages were identical to the first stages in each period of experiments 1-4. The third stage permitted buyers who had returned units under a Type 2 warranty, but failed to receive a replacement in stage two (*i.e.*, a dispute occurred) to appeal to the dispute resolution mechanism. In case such a buyer chose to appeal, he or she had to pay an appeal fee (*i.e.*, αc_m) and the seller had to pay an appeal fee (*i.e.*, βc_m). The decision of the mechanism was then determined by a random draw from the appropriate conditional probability distribution.[7] If the DRM ruled in favor of the buyer, the seller had to replace the unit, at an additional cost of c. If the DRM ruled in favor of the seller, no replacement was made, and the buyer kept the original unit. In either case, no further appeal was allowed. Buyer and seller payoffs were determined as in markets with no DRM, except that appeal costs were subtracted where appropriate. (Appendix B contains instructions and record forms for experiments 5-8.)

Figure 2 shows all possible sequences of decisions and events that could occur in experiments 5-8 between the time a unit is purchased and the time true quality is announced.

[7] In order to save time, all of the random draws (including true quality, and buyer and seller clues) were done in advance.



STAGE 1
STAGE 2
STAGE 3
BUY UNIT
BUYER GETS CLUE
RETURN?
YES
NO
WARRANTY TYPE?
TYPE 1
TYPE 2
GET REPLACEMENT UNIT
SELLER GETS CLUE
S.C.=0
S.C.=1
KEEP ORIGINAL UNIT
APPEAL?
YES
NO
"JUDGE" GETS CLUE
J.C.=0
J.C.=1
GET REPLACEMENT UNIT
KEEP ORIGINAL UNIT
LEARN TRUE QUALITY

Figure 2

4. A Formal Model of the Market

As a baseline for formulating hypotheses about outcomes in our laboratory, we present in this section a model in which we assume competitive behavior by risk neutral agents. This model is similar to that developed by Palfrey and Romer (1983), with important modifications to capture features of our experimental design (competitive markets rather than monopoly; and only one replacement permitted for each unit under warranty). In the model, average product quality, buyer preferences, production and transaction costs, and the extent to which "true" quality can be observed by buyer and seller interact to determine warranties, product price, and the likelihood of disputes.

An equilibrium of the market is defined as follows. A buyer takes the price of a unit of the commodity and the warranty policy (Type 1, 2, or 3) offered by the seller as given. He chooses the best response; *i.e.*, for a risk neutral buyer, the action that maximizes expected value, net of transaction costs. This involves whether or not to buy and when to return under the warranty. These decisions generate a response function for the buyer. Sellers take these response functions as given, and offer prices and warranties to maximize expected profit. We call the outcome of this interaction between buyers and sellers the equilibrium in this market.

We develop this model in the absence of a DRM. This provides the basis for Experiments 1-4. We then incorporate DRMs into the framework.

4.1 Market with no DRM

All buyers are identical and all sellers are identical. Since each unit is drawn independently from the same distribution, it is sufficient to calculate expected costs, values, and profits for a typical unit, for a representative buyer and a representative seller. Recall that if a seller does not replace a unit under a warranty, the buyer keeps the original unit, and the buyer's *ex post* payoff is determined by the true quality of the unit. If the seller makes a replacement, the original unit is discarded, and the replacement unit may not be returned. The buyer's *ex post* payoff is determined by the true quality of the replacement unit. The true quality of a unit is not known until all relevant transactions (buying and selling, warranty claims, and replacements) are completed.

We summarize the notation:

V_h	Buyer value if unit is high quality
$V_l \equiv 0$	Buyer value if unit is low quality
c	Seller's production cost of a unit
c_b	Buyer's cost of returning a unit
c_s	Seller's cost of processing a return
t	Ex ante probability that unit is high quality
q	Probability of correct quality signal (same for buyers and sellers)

Note that we have set $V_l = 0$, as in all the experiments.

Buyers' expected values and sellers' expected costs

For a given unit, these depend on the warranty offered by the seller and on the return strategy followed by the buyer.

A. *Type 1 ("Full") Warranty*. With this type of warranty, the seller makes a replacement if the buyer makes a claim. If the buyer makes a claim only if he gets a signal indicating low quality, the buyer's expected value (gross of price) is

$$V_1 = tqV_h + t(1\text{-}q)(\text{-}c_b + tV_h) + (1\text{-}t)q(\text{-}c_b + tV_h) \qquad (1)$$

The first term on the RHS is V_h times the probability that the buyer correctly perceives a good unit and does not return. The second term is the probability that the buyer returns what is in fact a good unit—t(1-q)—times the expected value of the replacement *minus* the return cost. The last term corresponds to the buyer returning a low quality unit. A fourth term, corresponding to keeping a bad unit after receiving a 1 signal—$(1\text{-}t)(1\text{-}q)V_l$—is omitted, since we have taken $V_l = 0$.

The seller's expected cost associated with this return strategy is:

$$K_1 = c + [t(1\text{-}q) + (1\text{-}t)q](c + c_s) \qquad (2)$$

B. *Type 2 ("Limited") Warranty*. With this type of warranty, the seller makes a replacement only if he receives a signal that the returned unit is of low quality. If the buyer returns a unit when he receives a signal indicating low quality, his expected value is:

$$V_2 = tqV_h + [t(1\text{-}q) + (1\text{-}t)q]\left[\text{-}c_b + \frac{q(1\text{-}q) + t(1\text{-}q)^2 + (1\text{-}t)q^2}{t(1\text{-}q) + (1\text{-}t)q}\, tV_h\right] \qquad (3)$$

The second term has two components. The first corresponds to the probability that the buyer returns the unit. The second gives the expected value given that a return has been made. Equation (3) may be rewritten as follows:

$$V_2 = tqV_h + [t(1\text{-}q) + (1\text{-}t)q](\text{-}c_b + tV_h), \qquad (4)$$

i.e., $$V_2 = V_1.$$

Buyers are indifferent between Type 1 and Type 2 warranties at the same price.

A seller's expected cost for a Type 2 warranty with the buyer returning perceived low quality units is:

$$K_2 = c + [t(1\text{-}q) + (1\text{-}t)q] \left[c_r + \frac{t(1\text{-}q)^2 + (1\text{-}t)q^2}{t(1\text{-}q) + (1\text{-}t)q}\, c \right] \tag{5}$$

Note that $K_2 < K_1$ for $0<t<1$ and $0<q<1$. Sellers' costs are lower under a Type 2 warranty than under Type 1.

For the parameter values of our experiments, given a Type 1 or 2 warranty, the strategy of returning only units whose signal is 0 dominates all other return strategies. Of course, this does not mean that in laboratory markets all buyers will use this strategy. For completeness, we present in Table 3 buyers' expected values (gross of price) and sellers' expected cost for all possible buyer strategies.[8]

C. *Type 3 ("As Is") Warranty*. This is, effectively, no warranty at all. A buyer's expected value is simply

$$V_3 = tV_h, \tag{6}$$

while a seller's expected (and actual) cost is

$$K_3 = c. \tag{7}$$

Price-Warranty Equilibrium

To define a price-warranty equilibrium, we need a bit more notation. Denote by $K_j(s)$ the expected cost of selling a unit with a type j warranty to a buyer using strategy s. Denote by $V_j(s)$ the expected value to a buyer using strategy s of having a unit with a type j warranty. Then a ***price-warranty equilibrium*** is a set of prices, one for each warranty, $(p_1^*,\ldots,p_J^*)$; a set of strategies, one for each warranty, $(s_1^*,\ldots,s_J^*)$; and a warranty, j^*, such that

$$\begin{array}{rcll} K_j(s_j^*) & = & p_j^* & \text{for all } j \quad (8a)\\ V_j(s_j^*) & \geq & V_j(s) & \text{for all } s, j \quad (8b)\\ V_{j^*}(s_{j^*}^*) - p_{j^*}^* & \geq & V_j(s_j^*) - p_j^* & \text{for all } j \quad (8c) \end{array}$$

Intuitively, a price-warranty equilibrium is a set of prices, one for each warranty, such that sellers earn zero expected profits at those prices, given optimal buyer return strategies, and a particular warranty such that, at the equilibrium prices, no other warranty gives the buyers higher expected surplus (*i.e.*, value net of price).

In a market with competitive, risk neutral sellers and a fixed number of identical risk neutral buyers with "zero-one" demands, at these equilibrium prices all buyers will buy units with warranty j^*, sellers earn zero expected profits,

[8] We exclude the "silly" strategy of returning only if the signal is 1. We saw no evidence of such a strategy being used by buyers in our experiments.

and there is no price for any warranty a seller might offer which would give that seller positive profits.

From inspection of Table 3 one can see that only Type 2 or 3 can be equilibrium warranties, since Type 1 is dominated by Type 2. Indeed, since buyer values under a Type 1 warranty are dominated by values under a Type 2 warranty, for ***any*** return strategy, this result is independent of buyer risk attitudes.[9] Whether Type 2 or Type 3 is the equilibrium depends on whether or not V_2-K_2 exceeds V_3-K_3—*i.e.*, it depends on the parameter values.

[9] We are assuming here that $c_b > 0$. Otherwise it is possible to have Type 1 be an equilibrium warranty. In our experiments $c_b > 0$ for all markets.

Table 3

EXPECTED VALUES (ABOVE DASHED LINES) AND COSTS WITH NO DRM

Buyer return strategy \ Warranty type	1	2	3
Return original unit regardless of signal	tV_h-c_b	$tqV_h-c_b+[t(1-q)+(1-t)q]tV_h$	NA
	$2c+c_s$	$c+c_s+[t(1-q)+(1-t)q]c$	NA
Return original unit only if signal equals 0	$tqV_h+[t(1-q)+(1-t)q](tV_h-c_b)$	$tqV_h+[t(1-q)+(1-t)q](tV_h-c_b)$	NA
	$c+[t(1-q)+(1-t)q](c+c_s)$	$c+[t(1-q)+(1-t)q]c_s$ $+[t(1-q)^2+(1-t)q^2]c$	NA
Never return original unit	tV_h	tV_h	tV_h
	c	c	c

4.2 Market with a DRM

The addition of a DRM adds one more "stage" to the decision-making problem of buyers. Besides having to decide whether or not to return a unit, given a signal, the buyer must decide whether or not to appeal if the seller refuses to make a replacement under the warranty. This expands the number of alternative buyer strategies considerably in the case of Type 2 warranties. There are no disputes with Type 1 or Type 3 warranties, so that the strategies of buyers and the values and costs for those two types of warranties are unaffected by the DRM.

Recall that a DRM is parameterized by the following elements:

- c_m — cost of the mechanism. This can be thought of as the total resource cost of processing each dispute which is brought before the mechanism.
- α — proportion of cost borne by the buyer. Thus αc_m is the effective cost to the buyer of appealing a warranty dispute.
- β — proportion of cost borne by the seller. Thus, βc_m is the effective cost to the seller if a buyer appeals a warranty dispute.
- $\hat{q}$ — probability that the mechanism requires the seller to replace the original unit, conditional on the unit being defective. $\hat{q}$ is also the probability that the mechanism does not require the seller to replace a unit, given it is not defective.

Hence, c_m, α, and β represent the cost parameters of the model, and $\hat{q}$ represents the accuracy of the mechanism. One might think that we are implicitly assuming that the mechanism is neutral (or *unbiased*) since the probability that it erroneously judges a defective unit to be good equals the probability that it erroneously judges a good unit to be defective. That is, the probability of incorrectly ruling in favor of the seller when true quality is bad equals the probability of ruling in favor of the buyer when the unit is good. However, as we shall see, the *overall* probability of erroneously ruling in favor of the buyer or the seller will in general depend upon the proportion of appealed units which are good or defective, which in turn will depend on q, t, and the buyer strategies.

Buyer strategies with Type 2 warranties

We consider the following six strategies the buyer could follow under a Type 2 warranty. They are:

(1) Never return original unit (and, hence, never appeal);
(2) Always return original unit, regardless of signal, and always appeal disputes, regardless of signal;
(3) Always return original unit, regardless of signal, but only appeal

a dispute if signal was 0;
(4) Return units only if signal was 0, appeal disputes;
(5) Return all units, appeal no disputes;
(6) Return units only if signal was 0, appeal no disputes.

Given a mechanism $M = (c_m, \alpha, \beta, \hat{q})$ and the other parameters of the market (V_A, c, c_A, c_s, t, q), we can calculate the expected buyer values and expected seller costs under a Type 2 warranty for each of these buyer strategies.[10] These expressions are given in Table 4. Notice that the entries for strategies (1), (5), and (6) correspond to entries in column 2 of Table 3.

[10] As before, we exclude strategies in which buyers return (or appeal) only if they receive a 1 signal.

Table 4

EXPECTED VALUES AND COSTS WITH DRM UNDER TYPE 2 WARRANTIES

Buyer Strategy	Expected Value / Expected Cost
Return all units Appeal all disputes	$\dfrac{tV_h[t(q-1)+(1-t)q]+tq\hat{q}V_h+[tq(1-\hat{q})+(1-t)(1-q)\hat{q}]tV_h -[tq+(1-t)(1-q)]\alpha c_m - c_b}{c+c_r+[t(1-q)+(1-t)q]c+[tq+(1-t)(1-q)]\beta c_m +[tq(1-\hat{q})+(1-t)(1-q)\hat{q}]c}$
Return all units Appeal disputes only if signal was 0	$\dfrac{tV_h[t(1-q)+(1-t)q]+tq^2V_h-q(1-q)\alpha c_m-c_b +q(1-q)[t\hat{q}V_h+(t(1-\hat{q})+(1-t)\hat{q})tV_h]}{c+c_r+[t(1-q)+(1-t)q]c+q(1-q)\beta c_m +q(1-q)[t(1-\hat{q})+(1-t)\hat{q}]c}$
Return all units Never appeal	$\dfrac{tqV_h-c_b+[t(1-q)+(1-t)q]tV_h}{c+c_r+[t(1-q)+(1-t)q]c}$
Return original unit only if signal equals 0 Appeal disputes	$\dfrac{tqV_h+[t(1-q)^2+(1-t)q^2]tV_h +q(1-q)[t\hat{q}V_h+(t(1-\hat{q})+(1-t)\hat{q})tV_h] -q(1-q)\alpha c_m-[t(1-q)+(1-t)q]c_b}{c+[t(1-q)+(1-t)q]c_r+[t(1-q)^2+(1-t)q^2]c +q(1-q)[t(1-\hat{q})+(1-t)\hat{q}]c+q(1-q)\beta c_m}$
Return original unit only if signal equals 0 Never appeal disputes	$\dfrac{tqV_h+[t(1-q)+(1-t)q](tV_h-c_b)}{c+[t(1-q)+(1-t)q]c_r +[t(1-q)^2+(1-t)q^2]c}$
Never return original unit	$\dfrac{tV_h}{c}$

A definition of price-warranty equilibrium analogous to that given for the no-DRM market applies here. The only difference is that the set of buyer strategies to be considered is larger.

5. Hypotheses

5.1 Hypotheses Derived From Model

The equilibrium model of warranties described above provides a number of specific predictions which suggest a natural way to organize and analyze the data from these eight experiments. By definition, an "equilibrium" is a *set of prices*, a *set of return strategies*, and a *single warranty*. We thus have three distinct groups of hypotheses about market behavior in these experiments:

1. Hypotheses about the transaction prices of units sold under each warranty;
2. Hypotheses about buyer decisions to return as a function of the signal the buyer receives (and, in the case of DRMs, buyer decisions about appealing in case of a dispute); and
3. Hypotheses about which warranties will be purchased.

Prices

In laboratory markets, it is well established that there is considerable variation in transaction prices in the early rounds of an experiment, as buyers and sellers gradually learn about the market. Hence, our hypotheses about prices are made with reference to the transaction prices that occur in the later periods of the experiment. That is, we interpret the model as making predictions about the prices towards which the market is converging. In presenting the results in the next section, we make the notion of "convergence" more specific.

The first set of predictions, based on the model of *risk neutral* sellers and buyers is:

H1: Transaction prices for warranty Types 1, 2, and 3 will converge to the values given in Table 5.

Table 5

EQUILIBRIUM PRICE PREDICTIONS UNDER RISK NEUTRALITY

	Warranty type		
Experiment #	1	2	3
1	1.85	1.62	1.20
2	1.85	1.62	1.20
3	3.65	3.20	-
4	3.65	3.20	2.40
5	1.80	1.69	-
6	1.80	1.84	-
7	1.80	1.58	-
8	1.80	1.84	-

Each row of Table 5 gives the predicted risk neutral equilibrium prices for warranty Types 1, 2, and 3 for an experiment.

If sellers are risk averse, the price predictions from the model do not apply. If they are risk averse, sellers will require a premium above the expected cost of selling a unit under warranty Types 1 and 2. It is straightforward to show that, if sellers are risk averse expected utility maximizers, an appropriately modified equilibrium prediction is the following:

H2: Transaction prices for warranty Types 1 and 2 will converge above the values given in Table 5. Transaction prices for Type 3 warranty will converge to the value given in Table 5.

Buyer Return/Appeal Strategies

Given that a buyer has purchased a unit under some warranty, the model predicts a particular return/appeal strategy that an expected value maximizing buyer should use. For each experiment and for each warranty type, Table 6 presents the optimal strategy. The notation we use for strategies[11] with no DRM is

S_1 = never return
S_2 = return all units
S_3 = return units only if a 0 signal was received

[11] Again, we omit the strategy discussed in footnote 8.

The notation we use with a DRM is

S_1 = never return
S_4 = return all units, appeal all disputes
S_5 = return all units, appeal only if a 0 signal was received
S_6 = return all units, never appeal
S_7 = return units if a 0 signal was received, appeal all disputes
S_8 = return units if a 0 signal was received, never appeal

Not every strategy is feasible under every warranty type. For example, with a DRM, under a Type 1 warranty only strategies S_1, S_6, S_8 are possible, since there are never disputes.

Note that *for a particular buyer with a particular unit*, it is not always possible to tell which of the feasible strategies the buyer is using. Suppose, for example, a buyer purchases a unit under a Type 1 warranty (with no DRM), receives a 1 signal, and subsequently does not return the unit. Then all the experimenter knows is that the buyer was using either strategy S_1 or strategy S_3. Thus, for the purposes of analyzing the data, the predictions in Table 6 have to be modified to take account of this observability problem. Hence, the hypothesis we examine is:

H3: Observed buyer strategies will be consistent with the entries in Table 6.

Table 6

OPTIMAL STRATEGIES FOR BUYERS

	Warranty type		
Experiment #	1	2	3
1	S_3	S_3	S_1
2	S_3	S_3	S_1
3	S_3	S_3	-
4	S_3	S_3	S_1
5	S_8	S_7	-
6	S_8	S_7	-
7	S_8	S_8	-
8	S_8	S_7	-

Warranties

The final prediction of the model concerns which warranty will be purchased by the buyers, as a function of the parameters of the model. For all eight experiments the predicted equilibrium warranty is Type 2.

H4: Once market prices have converged, only Type 2 warranties will be purchased by buyers.

A natural reaction to H4 is that warranty choice must depend on the risk attitudes of buyers and sellers. This is true only to a limited extent. Consider first the case with no DRM. For any return strategy, the distribution of (ex post) payoffs to a buyer of a unit with a Type 2 warranty ***weakly dominates*** the distribution of payoffs with a Type 1 warranty. Hence, if a Type 1 warranty were offered at a higher price than a Type 2 warranty, any rational, informed buyer should never purchase the Type 1 warranty. ***This is true regardless of the risk attitude of the buyer.*** The seller expected costs for a unit sold under a Type 1 warranty are also higher for a Type 2 warranty, so the predicted equilibrium price of a Type 1 warranty should be higher. Hence, ***in equilibrium***, no buyer should ever purchase a Type 1 warranty.

In a market with a DRM, the buyer has the option of never appealing disputes. For any return strategy available to a buyer under a Type 1 warranty (S_1, S_6, or S_8), the above argument about the distribution of buyer payoffs holds. Under a Type 2 warranty, a buyer has the additional option of appealing disputes (*i.e.*, choosing S_4, S_5, or S_7). This strengthens the dominance of Type 2 over Type 1 warranties. However, due to seller appeal costs, Type 2 warranties may be more expensive than Type 1 warranties in equilibrium.[12]

These observations are summarized in the following hypotheses:

H5: If a buyer has a choice between purchasing a Type 1 and a Type 2 warranty and the Type 2 warranty is less expensive than the Type 1 warranty, the buyer will choose the Type 2 warranty.

H6: In experiments 1-4, once the market has converged, no Type 1 warranties will be purchased.

No unambiguous comparison can be made between Type 2 and Type 3 warranties. Despite the fact that in equilibrium Type 2 warranties will yield greater expected surplus to buyers, Type 3 warranties have a much different distribution of returns. In particular, in our design, the "downside risk" to a buyer is minimized with Type 3 warranties since they are cheaper in equilibrium. Thus, one might expect some sufficiently risk averse buyers to purchase them.

[12] This is true in experiments 6 and 8 (*see* Table 5).

5.2 Alternative Theories of Behavior with a DRM

Do the DRMs we investigate in our experiments affect market outcomes? According to our theory, there are clear differences in the predictions of the model depending upon (a) the presence or absence of the DRM and (b) the specific form of the DRM.

A simple approach to investigating this question involves the following alternative models for the behavior of buyers and sellers in these markets. They assume limited rationality in one form or another. Because of the difficulty of the decision problems faced by buyers and sellers with a DRM (recall Figure 2), there is reason to believe that buyers and sellers may be simplifying their decision problem. One way this might happen involves myopic behavior. Buyers, for example, may evaluate alternative warranties being offered at different prices without conditioning on the possibility of appealing a dispute with the seller. The same is possible for sellers. The actual application of Bayes' rule required to make the calculations of probabilities of disputes and the distribution of values (and costs) conditional upon a dispute having been appealed is a difficult task. Similar points have been made by others studying warranties in arguing that consumers do not properly evaluate warranties when making a purchase in product markets, and have great difficulty correctly assessing probabilities.

If both the buyer and seller act in this myopic fashion, a DRM will have no "market" effect. Its *only* effect will be on ex post outcomes (*i.e.*, the quality of units buyers consume, seller costs, and mechanism operating costs). In particular, *prices* will be as if there were no DRM. Similarly, buyer return strategies would be unaffected, as would the equilibrium warranty. The equilibrium price of a Type 1 warranty is unaffected by this kind of myopia. For experiments 5-8 the equilibrium price of a Type 2 warranty that would have been predicted had there been no DRM is $1.58.

H7: In experiments 5-8 prices of Type 2 warranties converge to $1.58.

A second interesting set of hypotheses emerge if we assume that *only buyers are myopic*. When sellers are "fully rational," [13] prices are those given in Table 5 for experiments 5-8. A myopic buyer evaluates expected value (incorrectly) to be the same for Type 1 and Type 2 warranties, and consequently should simply choose whichever one is cheaper. Since in equilibrium Type 1 warranties are cheaper in experiments 6 and 8 and Type 2 warranties are cheaper in experiments 5 and 7, we have:

H8: In experiments 5 and 7, after markets have converged, only Type 2 war-

[13] Since the sellers make no decision in stages 2 and 3, it seems more plausible that sellers are fully rational than that buyers are. One could also consider the alternative model in which *buyers* are fully rational and sellers are myopic. We do not do so here because, given the parameters of our experiments, the predictions would be the same as those of the complete myopia model (H7).

ranties will be purchased, by myopic and sophisticated buyers alike. In experiments 6 and 8, after markets have converged, Type 1 warranties will be purchased by myopic buyers and Type 2 warranties will be purchased by sophisticated buyers.

5.3 Learning and Convergence to Equilibrium

Two very strong ***informational*** assumptions are made in the theoretical model of Section 4: (1) buyers and sellers both know all the relevant distributions and can condition using Bayes' rule; (2) sellers know what return strategies are being used by buyers for each warranty they might offer. In fact, buyers and, particularly, sellers have to acquire this information by trial and error and by observing what decisions are made by buyers in the various stages of the experiment. This learning process is further complicated by some additional problems. Buyers and sellers are also learning the general procedural rules. At the same time, prices of different warranties are constantly fluctuating. Moreover, since buyer signals are private information, sellers cannot directly observe buyer strategies.

In order to focus our attention on one particular feature of these learning processes and at the same time tie it in with our basic model, we isolate the problem of sellers learning the expected cost of a warranty. In fact, sellers do not care precisely which strategy each buyer is using, or the distributions of qualities and signals. All that matters to the sellers is the distribution of costs for units sold under each warranty. We assume no learning is required for sellers to figure out that the cost of a Type 3 warranty is equal to c.

With Type 1 warranties, in experiments 1-4, the sellers could incur costs equal to either c or $2c+c_s$, depending on whether a buyer returns the unit or not. Hence, the only relevant statistic needed for a seller to estimate expected cost of a Type 1 warranty is Π^1_r, the probability of a return under a Type 1 warranty. The laboratory markets were carefully organized so that information about how many buyers have returned units under each warranty, and how many units have been sold under each warranty is public at every point in time. We assume that the seller estimates Π^1_r in period t by using these past frequency data:

$$\hat{\Pi}^1_{rt} = \frac{\text{number of type 1 units returned through period t-1}}{\text{number of type 1 units sold through period t-1}}$$

The estimate of expected cost at time t is then the same for each seller, and is given by $\hat{\Pi}^1_{rt}(2c+c_s) + (1-\hat{\Pi}^1_{rt})c = E^1_t$.

We assume that sellers use similar procedures to estimate expected costs of Type 2 warranties in experiments 1-4. Here, there are three possible costs that the seller could incur: c, $c+c_s$, $2c+c_s$. These correspond to ***no return***, ***return but no replacement***, and ***return and replacement***, respectively. The seller must now estimate two parameters: Π^2_{nr} = probability of a return and no replacement; and Π^2_r = probability of a return and replacement. The necessary information for all sellers to use market aggregate return and replacement frequencies was made public throughout the experiments. Using past frequencies, sellers could obtain estimates $\hat{\Pi}^2_{nrr}$ and $\hat{\Pi}^2_{rr}$. The estimate of expected cost was then $E^2_t = \hat{\Pi}^2_{nrr}(c+c_s) + \hat{\Pi}^2_{rr}(2c+c_s) + (1-\hat{\Pi}^2_{rr}-\hat{\Pi}^2_{nrr})c$.

Under a DRM, with a Type 1 warranty the sellers again only had to estimate Π^1_{rr}. For a Type 2 warranty, however, there are four possible costs: c, 2c, $c+\beta c_m$, $2c+\beta c_m$. (Recall that $c_s = 0$ in experiments 5-8.) Thus, three probabilities have to be estimated:

Π^m_{r} = probability of a return and replacement (no dispute)
Π^m_{na} = probability of return and appeal, no replacement
Π^m_{a} = probability of return and appeal and replacement.

Expected costs are computed by

$$E^m_t = \hat{\Pi}^m_{rr}(2c) + \hat{\Pi}^m_{nar}(c+\beta c_m) + \hat{\Pi}^m_{ar}(2c+\beta c_m) + (1-\hat{\Pi}^m_{rr}-\hat{\Pi}^m_{nar}-\hat{\Pi}^m_{ar})c.$$

This learning model yields the following hypothesis:

H9: In period t, the price of a warranty is given by the estimated expected costs, either E^1_t, E^2_t or E^m_t.

6. Data

6.1 Price Convergence

The entire time series of prices for all warrranties in all experiments is displayed in Figures 3-10. The average transaction price for each warranty in each period is given in Table 7 for Experiments 1-4 (no DRM) and in Table 8 for Experiments 5-8 (with DRM).

Figure 3

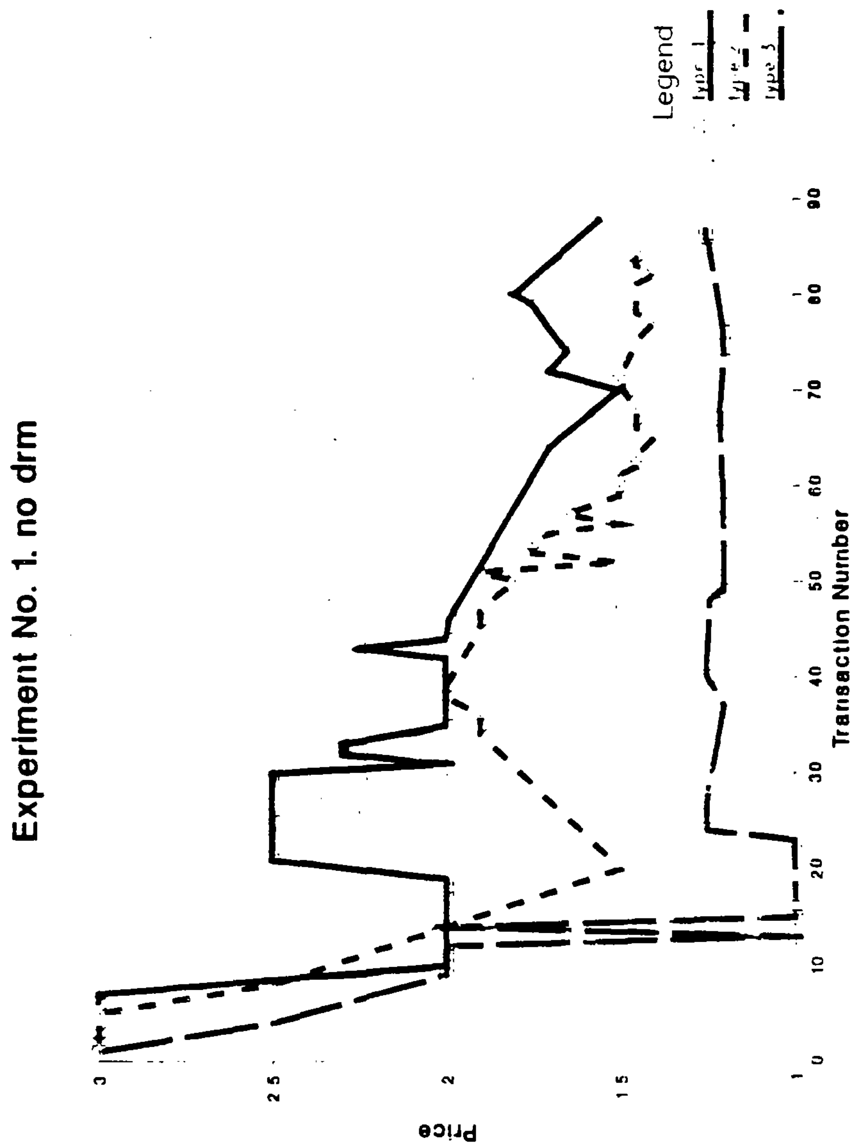

Experiment No. 1. no drm
Legend
Price
3
2.5
2
1.5
1
Transaction Number
0
10
20
30
40
50
60
70
80
90

Figure 4

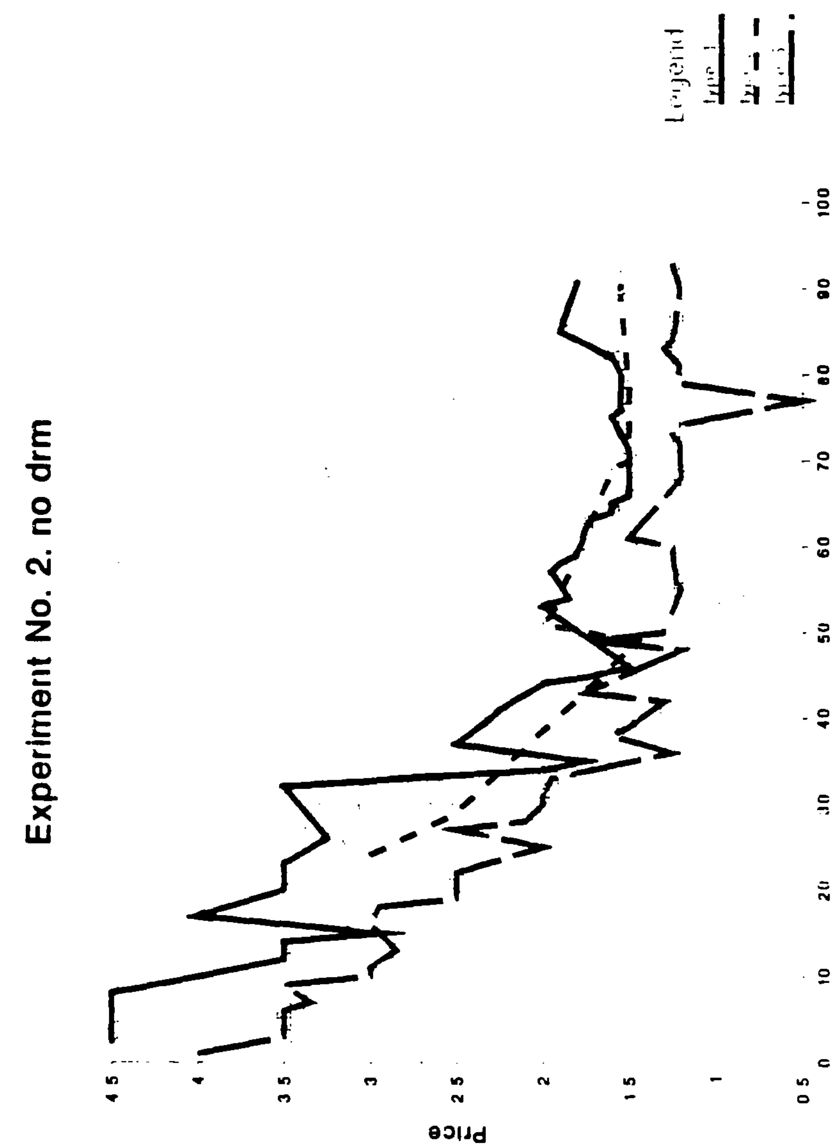

Experiment No. 2. no drm
Price
Transaction Number
0 10 20 30 40 50 60 70 80 90 100
4.5 4 3.5 3 2.5 2 1.5 1 0.5

Figure 5

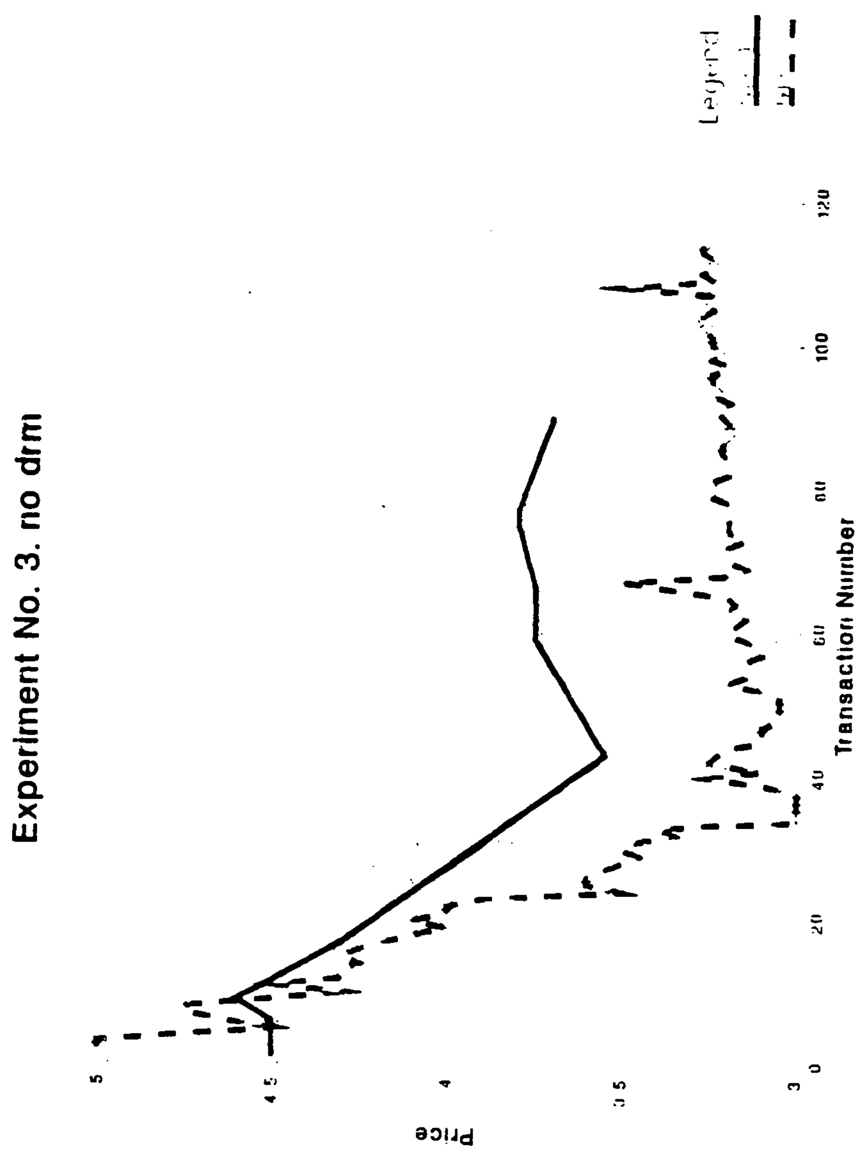

Experiment No. 3. no drm
Price
5
4.5
4
3.5
3
Transaction Number
0
20
40
60
80
100
120

Figure 6

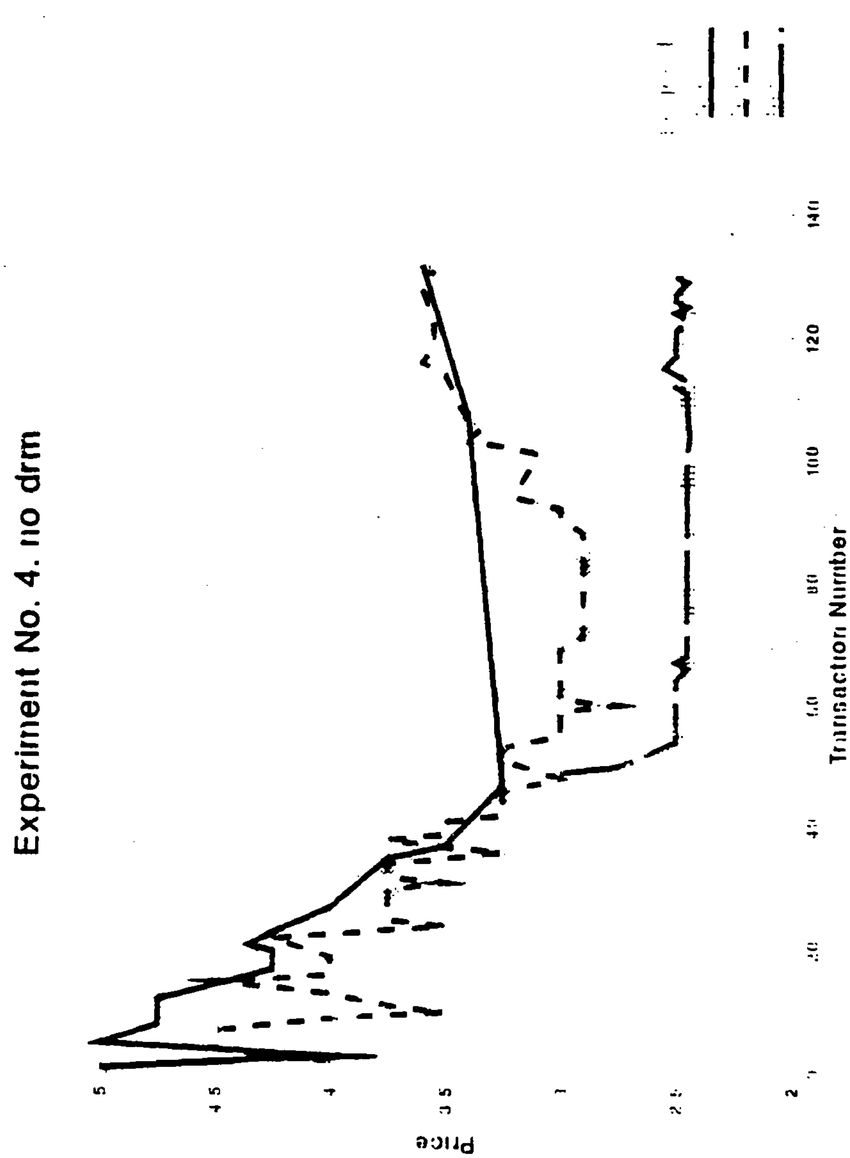

Experiment No. 4. no drm
Price
Transaction Number
5
4.5
4
3.5
2
20
100
120

Figure 7

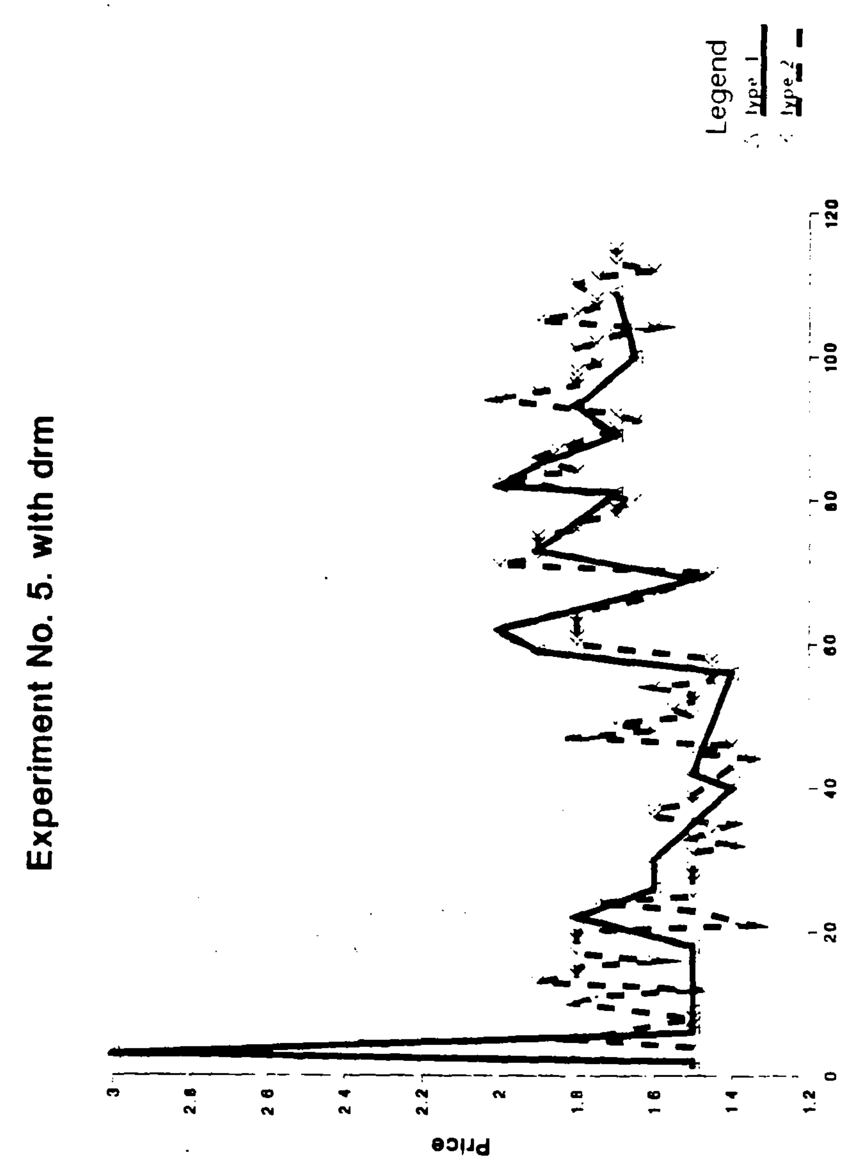

Experiment No. 5. with drm
Legend
type 1
type 2
Price
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
0
20
40
60
80
100
120
Transaction Number

Figure 8

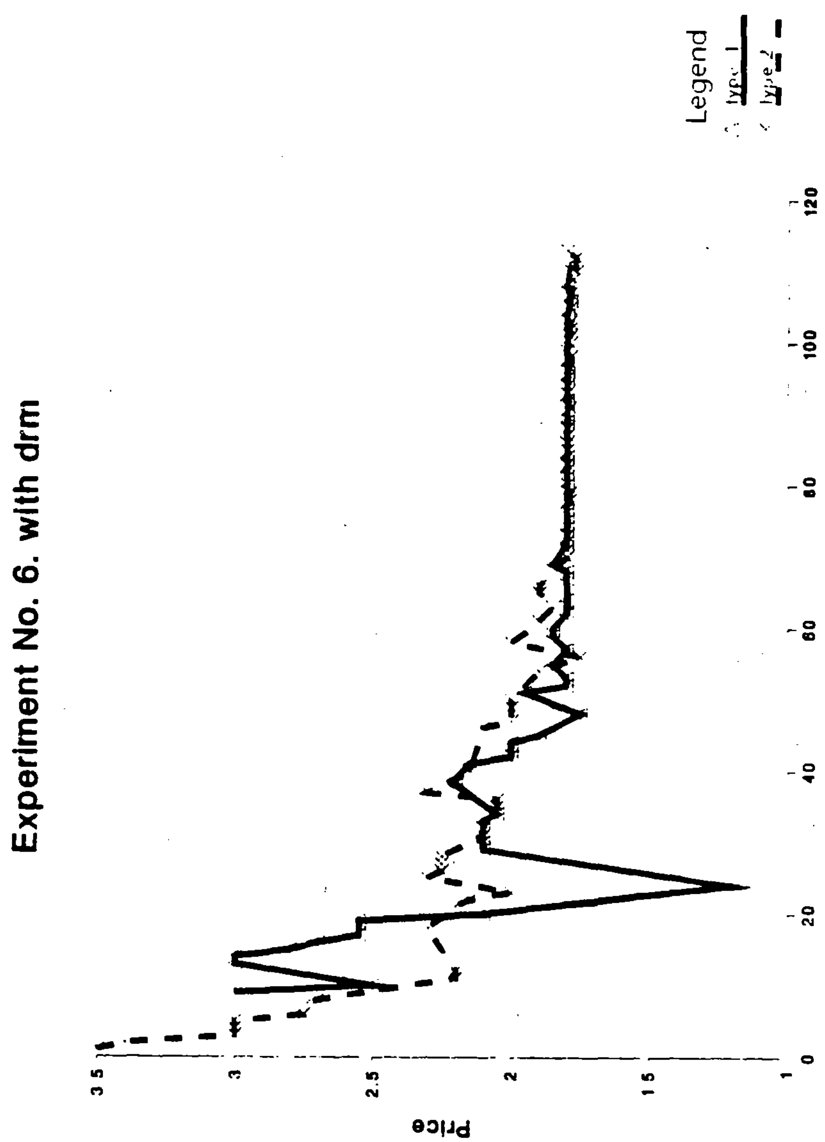

Experiment No. 6. with drm
Legend
type 1
type 2
Price
3.5
3
2.5
2
1.5
1
Transaction Number
0
20
40
60
80
100
120

Figure 9

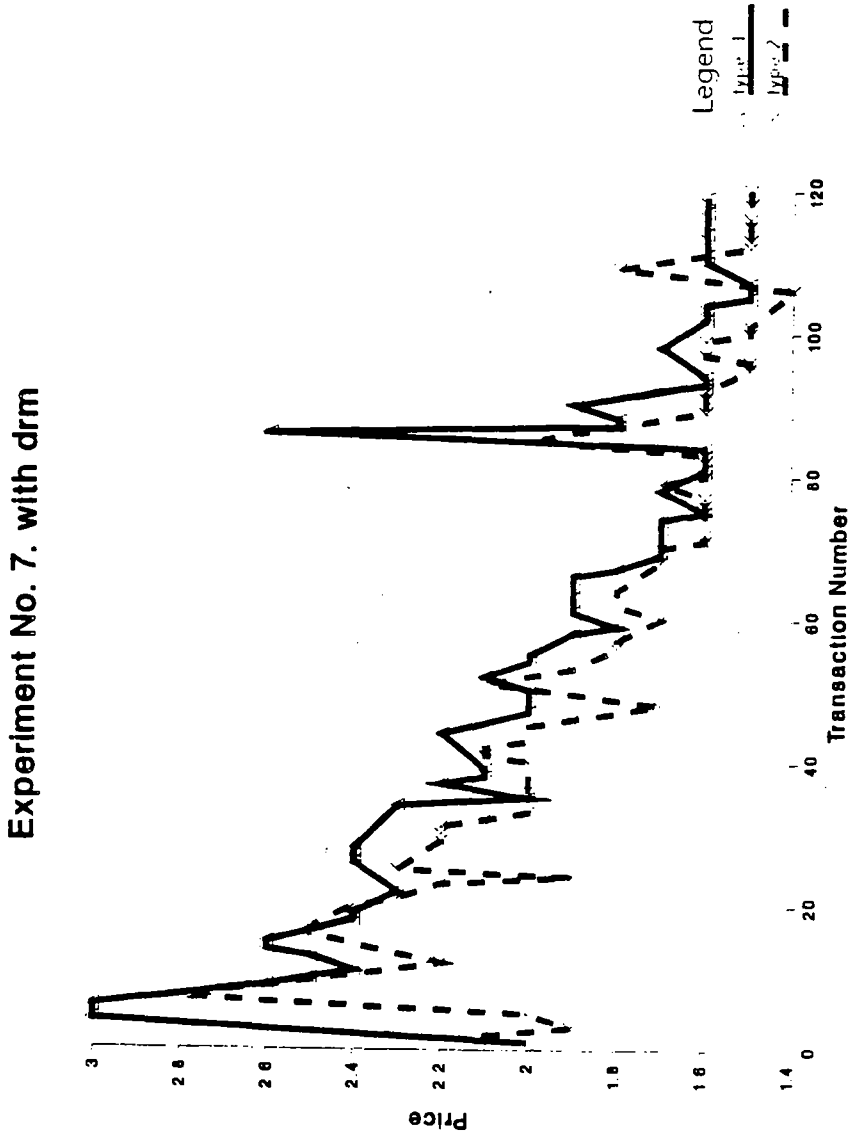

Experiment No. 7. with drm
Price
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
0
20
40
60
80
100
120
Transaction Number
Legend

Figure 10

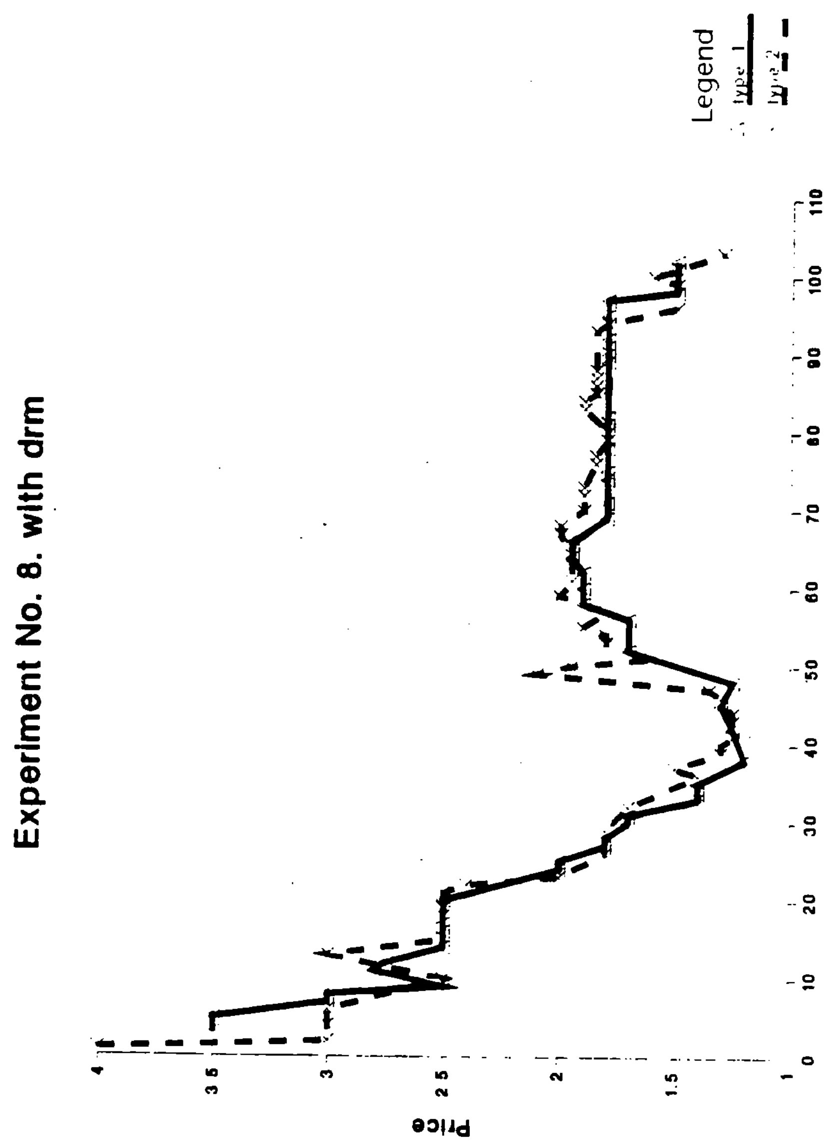

Experiment No. 8. with drm
Legend
type 1
type 2
Price
4
3.5
3
2.5
2
1.5
1
Transaction Number
0
10
20
30
40
50
60
70
80
90
100
110

Table 7

AVERAGE PRICES FOR EXPERIMENTS 1-4.

Experiment #	Warranty Type	Period Number										
		1	2	3	4	5	6	7	8	9	10	11
1	1	2.67	2.00	2.50	2.22	2.06	-	1.70	1.62	1.70	-	-
	2	2.88	1.50	2.00	1.95	1.90	1.69	1.46	1.45	1.44	-	-
	3	2.50	1.40	1.17	1.20	1.25	1.20	1.21	1.20	1.25	-	-
2	1	4.25	3.39	2.50	2.04	1.90	1.62	1.58	1.85	-	-	-
	2	-	3.00	2.50	1.45	1.98	1.55	1.50	1.55	-	-	-
	3	3.42	2.72	1.97	1.47	1.35	1.29	1.23*	1.24	-	-	-
3	1	4.53	4.30	-	3.55	-	3.75	3.75	3.82	3.70	-	-
	2	4.69	4.14	3.41	3.10	3.10	3.14	3.25	3.20	3.22	3.24	3.25
4	1	4.70	4.10	3.88	3.38	-	-	-	-	3.40	-	3.60
	2	4.10	4.06	3.68	3.35	3.07	2.96	2.90	3.00	3.29	3.55	3.57
	3	-	-	-	-	2.65	2.48	2.45	2.45	2.45	2.48	2.48

*One sale at $.50 not included.

Table 8

AVERAGE PRICES FOR EXPERIMENTS 5-8.

		Period Number									
Experiment #	Warranty Type	1	2	3	4	5	6	7	8	9	10
5	1	1.75	1.65	1.60	1.45	1.40	1.80	1.80	1.87	1.73	1.70
	2	1.64	1.71	1.50	1.51	1.53	1.69	1.82	1.80	1.79	1.75
6	1	3.00	2.49	2.09	2.02	1.84	1.82	1.80	1.80	1.80	1.80
	2	2.81	2.16	2.17	2.14	1.93	1.89	1.80	1.80	1.79	1.79
7	1	2.58	2.45	2.32	2.12	1.97	1.85	1.85	1.88	1.58	1.60
	2	2.28	2.28	2.14	1.97	1.86	1.70	1.62	1.63	1.52	1.54
8	1	3.01	2.40	1.68	1.25	1.77	1.93	1.80	1.80	1.68	-
	2	3.10	2.49	1.68	1.31	1.85	1.98	1.86	1.87	1.58	-

The price convergence results are mixed. If one considers each warranty in each experiment as a separate market, then in 9 out of 19 markets, or roughly half the time, the average transaction price in the final period of trade was within \$.05 of the price predictions of the risk neutral model (*see* Table 5). More specifically, 4 out of 8 Type 1 warranty markets (Experiments 2,3,4,6) converged in this sense, 3 out of 8 Type 2 warranty markets (Experiments 3,6,7) converged and 2 out of the 3 Type 3 warranty markets (Experiments 1,2) converged. Thus Hypothesis (H1), although not perfectly accurate, does reasonably well in predicting final period prices.[14] Interestingly, the weaker Hypothesis (H2), which is based on the assumption of risk averse sellers, fares worse. Of the nine Type 1 and Type 2 markets that do *not* converge to within \$.05 of the risk neutral predicted price, six have average final period prices *below* the risk neutral equilibrium, which is inconsistent with risk aversion.

Perhaps even more interesting than the observations about final period prices, and certainly more descriptive, is the general pattern of convergence over time which was observed in all but one of the experiments. In 7 of the 8 experiments, prices began well above the risk neutral predictions for both Type 1 and Type 2 warranties. Over the course the experiment, prices were competed down to a level well *below* the risk neutral predictions. Finally, toward the end of the experiment, these depressed prices gradually recovered back to a level close to the risk neutral equlibrium. The one exception to this pattern of convergence was Experiment 5 in which, by chance, average period 1 prices were within \$.05 of the risk neutral prediction. Yet we still observed the second and third "stages" of convergence, with prices temporarily bid down below the risk neutral equilibrium and subsequently recovering.

This consistent price pattern suggested to us that prices were fluctuating in response to information that was gradually being learned by sellers about their expected costs under the alternative warranties. This prompted the theoretical learning model outlined in the previous section, in which sellers are assumed to begin with diffuse priors over the distribution of costs under the warranties, and use past frequency data to form estimates of these costs. The predicted prices under this model for each warranty in each period (except the first period) of each experiment are given in Table 9. To evaluate this learning model and the corresponding hypothesis (H9), we compare its accuracy with the accuracy of the risk neutral model. We include comparisons for the last several periods, rather than only the last period. As a measure of accuracy for each model, we

[14] Inspection of figures 9 and 10 and table 8 reveals that in Experiments 7 and 8 prices fell off rather dramatically at the conclusion of the experiments. In both cases, there was a seller who became upset that his earnings per sale were essentially zero. One of them later told us that he decided to "punish" other sellers by selling at ridiculously low prices, even if it meant losing money (which it did!). This deviant behavior occurred after these markets had apparently converged for several periods at or very near the risk neutral equilibrium prices. If we were to use average prices two periods from the end of these two experiments, the risk neutral predictions would have performed very well.

use the absolute difference between the model's predicted price and the average transaction price in each period beyond period 5. The absolute deviations from the risk neutral predictions are given in Table 10, and the absolute deviations from the learning model are given in Table 11. Out of 66 total comparisons, the two models perform equally well 5 times, the risk neutral model is better in 39 cases and the learning model is better in 25 cases. Using a simple binomial test based on this data clearly rejects the learning model in favor of the "static" risk neutral predictions. We find this surprising, as it is clear from the overall pattern of convergence that a significant amount of learning is taking place. Evidently, our learning model is not the appropriate one.

Table 9

ACTUAL MEAN (PAST) COSTS FOR TYPE 1 AND TYPE 2 WARRANTIES IN LATER PERIODS
(LEARNING MODEL PRICE PREDICTIONS)

		Experiment Number							
Warranty Type	Period #	1	2	3	4	5	6	7	8
1	6	1.78	1.74	3.90	3.36	1.57	1.78	1.61	1.82
	7	1.78	1.76	4.19	3.36	1.65	1.86	1.61	1.84
	8	1.75	1.77	3.96	3.36	1.60	1.82	1.61	1.82
	9	1.67	-	3.36	3.36	1.80	1.61	1.83	-
	10	-	-	3.54	3.47	1.57	1.83	1.61	-
	11	-	-	3.54	3.47	-	-	-	-
2	6	1.82	1.74	3.20	3.26	1.65	1.78	1.67	1.88
	7	1.79	1.59	3.16	3.19	1.69	1.71	1.69	1.86
	8	1.72	1.54	3.14	3.23	1.71	1.74	1.65	1.84
	9	1.72	-	3.11	3.22	1.74	1.70	1.69	1.83
	10	-	-	3.08	3.23	1.70	1.69	1.70	-
	11	-	-	3.24	3.23	-	-	-	-

Table 10

ABSOLUTE DEVIATIONS OF MEAN TRANSACTED PRICES FROM RISK NEUTRAL EQUILIBRIUM PREDICTIONS

		Experiment Number							
Warranty Type	Period #	1	2	3	4	5	6	7	8
1	6	-	.23	.10	-	0	.02	.05	.13
	7	.15	.27	.10	-	0	0	.05	0
	8	.23	0	.17	-	.07	0	.08	0
	9	.15	-	.05	.25	.07	0	.22	.12
	10	-	-	-	-	.10	0	.20	-
	11	-	-	-	.05	-	-	-	-
2	6	.06	.08	.06	.24	0	.05	.12	.14
	7	.17	.13	.05	.30	.13	.04	.04	.02
	8	.18	.08	0	.20	.11	.04	.05	.03
	9	.19	-	.02	.09	.10	.05	.06	.26
	10	-	-	.04	.35	.06	.05	.04	-
	11	-	-	.05	.37	-	-	-	-

Table 11

ABSOLUTE DEVIATIONS OF MEAN TRANSACTED PRICES FROM LEARNING MODEL PREDICTIONS

		Experiment Number							
Warranty Type	Period #	1	2	3	4	5	6	7	8
1	6	-	.12	.15	-	.23	.04	.24	.11
	7	.08	.18	.44	-	.15	.06	.24	.04
	8	.13	.07	.14	-	.27	.02	.27	.02
	9	.03	-	.05	.04	.19	0	.03	.15
	10	-	-	-	-	.13	.03	.01	-
	11	-	-	-	.13	-	-	-	-
2	6	.13	.19	.06	.30	.04	.11	.03	.10
	7	.33	.09	.09	.29	.13	.09	.07	0
	8	.27	.01	.06	.23	.09	.06	.02	.03
	9	.28	-	.09	.07	.05	.09	.17	.25
	10	-	-	.16	.32	.05	.10	.16	-
	11	-	-	.01	.34	-	-	-	-

6.2 Buyer Behavior

The data on the frequency of inconsistent buyer return behavior are given in Table 12. In experiments 1-4, the predictions about buyer behavior are particularly simple. Under either a Type 1 or a Type 2 warranty a buyer should return units with a 0 clue and retain units with a 1 clue. Only 6 out of 314 buyer decisions were inconsistent with the predictions of the model.

Table 12

PROPORTION OF BUYER DECISIONS INCONSISTENT WITH PREDICTIONS IN TABLE 6
(TOTAL NUMBER OF UNITS IN PARENTHESES)

	Experiment Number							
Warranty Type	1	2	3	4	5	6	7	8
1	0.0 (25)	0.0 (36)	.09 (11)	0.0 (15)	.125 (24)	0.0 (52)	.103 (58)	0.0 (47)
2	.025 (40)	0.1 (10)	0.01 (103)	0.026 (74)	.166 (90)	.082 (61)	.177 (62)	.25 (56)

In experiments 5-8, the possibility of appealing disputes complicates the buyers' decision problem considerably (compare Figures 1 and 2), at least with Type 2 warranties. Type 1 warranties, on the other hand, present the buyers with the same decision problem as in Experiments 1-4. Overall, there were many more inconsistencies—a total of 54 out of 450. This larger number may be partly due to the more complicated decision rule. For example, with Type 2 warranties the strategy of always returning followed by appealing only if the buyer signal was 0 yields an expected payoff to buyers which is only \$.05 less in expected value, even though it is a strictly dominated strategy. Of the 54 inconsistencies, 31 were of this variety. Hence most of the inconsistencies reduced payoffs by a fairly small amount. In experiment 7, where buyers were forced to bear the cost of the DRM, this type of inconsistency occurred only three times.

6.3 Warranties

The data on frequency of warranties are presented in Tables 13 and 14. Hypotheses (H4), (H5), and (H6) all suggest that Type 2 warranties should drive out Type 1 warranties. In experiments 1-4, after period 1 the average price of Type 2 warranties is below the average price of Type 1 warranties ***without exception***. Furthermore, only rarely during the course of a period was a Type 1 offered at a lower price than a Type 2. Thus essentially *all* the observations of Type 1 warranties in the no-DRM experiments contradict our hypotheses about warranty coverage. We have no entirely satisfactory explanation for these occurrences.[15] It is interesting to note, however, that in experiments 3, 4, and 5, Type 2 warranties do succeed in almost completely driving out Type 1 warranties. However, in the other five experiments, Type 1 warranties persist throughout the experiment. Hence there is only weak evidence in favor of our predictions about the equilibrium level of warranty coverage.

[15] One possibility is that subjects prefer a ***simple*** lottery (Type 1 warranty) to an "equivalent" ***compound*** lottery (Type 2 warranty).

Table 13

FREQUENCY OF WARRANTIES IN EXPERIMENTS 1-4

Experiment #	Warranty Type	Period Number										
		1	2	3	4	5	6	7	8	9	10	11
1	1	3	4	2	8	4	0	1	3	3	-	-
	2	4	1	4	4	2	8	8	5	4	-	-
	3	3	5	3	1	3	2	1	2	3	-	-
2	1	4	5	4	6	5	6	4	2	-	-	-
	2	0	1	1	1	2	2	1	2	-	-	-
	3	8	6	7	5	5	4	5	7	-	-	-
3	1	3	1	0	1	0	2	1	2	1	0	0
	2	8	11	12	7	10	6	5	10	11	12	11
4	1	5	4	2	2	0	0	0	1	0	1	-
	2	7	8	10	10	7	7	7	8	3	3	-
	3	-	-	-	-	5	5	5	6	9	8	-

Table 14

FREQUENCY OF WARRANTIES IN EXPERIMENTS 5-8

Experiment #	Warranty Type	Period Number									
		1	2	3	4	5	6	7	8	9	10
5	1	6	2	2	2	1	3	2	3	2	1
	2	6	9	10	10	10	9	9	8	10	10
6	1	1	8	4	8	6	6	6	5	4	4
	2	11	4	8	4	5	6	5	6	8	6
7	1	6	6	5	5	7	6	6	6	6	5
	2	6	6	7	7	5	6	6	6	6	7
8	1	7	5	8	4	3	5	5	5	5	-
	2	5	7	4	8	7	5	7	6	7	-

6.4 Effects of DRM

In order to examine whether the DRM had a significant impact on prices and warranty coverage, in Section 5 we presented an alternative "myopic" theory of behavior. In its strongest version, this theory holds that both sellers and buyers ignore the presence of a DRM and hence the DRM will have no market effect in altering prices (H7). This hypothesis is rejected in favor of the risk neutral theoretical predictions.[16] The price effects of our particular DRMs seem to be accurately predicted by the risk-neutral model of Section 4.

The weaker assumption about limited rationality—that some buyers may be myopic, but sellers are sophisticated—is mildly supported by the data. Hypothesis (H8) implies that if, on average, each experiment has the same mix of myopic and sophisticated buyers, then after convergence more Type 1 warranties should be observed in Experiments 6 and 8 than in Experiments 5 and 7. There is a difference between these pairs of experiments, but this difference is too small to be any more than suggestive.

Including a DRM seems to have had an impact on the market. The evidence is that sellers and a large proportion of buyers acted in a sophisticated manner, anticipating the use of the DRM in case of a dispute.

7. Conclusions

The purpose of these experiments was two-fold. First, we wanted to design and study equilibrium behavior in a laboratory market environment with the following features:

- Buyers and sellers are uncertain about product quality.
- Each seller can offer a variety of warranties simultaneously.
- Disputes over warranty performance occur in the natural course of transactions.

Second, given such an environment, we investigated the effects of varying a DRM's cost and the allocation of this cost between buyer and seller.

The rich specification of these markets is susceptible to interpretation by several theories about market behavior (*i.e.*, about prices and warranty coverage). These theories range from an "ultra-rational" model of Bayesian equilibrium to "irrational" theories postulating that market participants will not properly take account of the complicated procedures of dispute resolution mechanisms. Given how complicated our markets were, particularly compared to previous laboratory markets, the "irrationality" theory seemed likely to predict well. We also put forth an intermediate "learning" model, in which buyers

[16] This comparison is even clearer if one adjusts for the deviant seller behavior described in footnote 14.

and sellers gradually become better informed about the likelihood of defects, disputes, and warranty performance. This model attempts to capture some of the dynamics of the convergence process in these markets.

Our findings in light of these theories are as follows. (These results are very suggestive, but, given only eight experiments we cannot claim that they are conclusive.) First, mechanism costs and the allocation of these costs affect market prices through effects on ***both*** buyer and seller behavior. Placing the burden of the DRM cost on the seller raises the price of goods offered with a warranty, and shifting the burden of cost to the buyer lowers the price of goods offered with a warranty. The effect of the DRM's cost on warranty coverage was inconclusive. We suspect that our design did not allow for enough variety of warranty coverage to be able to discern these subtle effects with only a few experiments.

A second conclusion is that the Bayesian equilibrium model predicts prices very well. This is quite surprising, given the model's relative complexity. Moreover, the Bayesian model significantly outperforms the learning model, even before the last period of an experiment. The overall pattern of the ***dynamics*** by which convergence was achieved in our experiments is not explained well by any of the models we examined. Clearly there is a lot of learning going on about probabilities and market variables (such as prices and what warranties mean). Identification of these learning processes is severely confounded by the fact that participants in the market are constantly adjusting their strategies as they learn. What is needed is a testable theory of how participants simultaneously adjust their ***beliefs*** and their ***strategies*** in response to endogenously generated market information.

Our experiments included some restrictive features to avoid making them too complicated. Some of these could be relaxed, so that other important questions about warranties and DRMs could be investigated. For example, we did not allow "fraudulent" behavior on the part of sellers. It has been argued by many that one important role of DRMs is to enforce "honest" seller behavior. It would be very interesting to run a comparable series of experiments permitting "fraudulent" behavior (*e.g.*, reneging on warranties), to see if DRMs play an important role in policing seller behavior and reinforcing reputations. Other interesting extensions of this work might involve any of the following modifications:

- allow sellers to control average quality (*i.e.*, choose t);
- introduce heterogeneity of sellers and of buyers;
- examine different price-making institutions (such as a double oral auction).

We believe that the basic technology is now established for intensive laboratory study of markets where product quality and warranties play a major role. While the design of these markets is much more complicated than that of usual experiments, it is sufficiently straightforward so that learning the "rules of the game" is not too severe a problem for the participants. (This is, of course, a fundamental consideration with any economics experiment, since we are not merely asking whether subjects understand instructions, but testing models of behavior, given the rules of the game.) It may also be possible to modify the basic design used here to study questions about disputes over contract performance in settings other than warranties in product markets. These might include questioned settlement claims, breach-of-contract disputes, and a variety of other important legal problems[17] where concerns naturally arise about the effects of rules and procedures for processing disputes.

References

R. Cooter, S. Marks and R. Mnookin, "Bargaining in the Shadow of the Law," *Journal of Legal Studies*, June 1982.

D. DeJong, R. Forsythe, and W. Uecker, "The Effects of Alternative Liability Rules and the Public Disclosure of Audit Fees on the Price and Quality of Audit Services: A Laboratory Market Study," Working Paper #84-4, College of Business Administration, University of Iowa, 1984.

Federal Trade Commission, *Handling Consumer Complaints: In-House and Third-Party Strategies*, Washington, D.C. (1980).

M. Lynch, R. Miller, C. Plott, and R. Porter, "Product Quality, Informational Efficiency and Regulation in Experimental Markets," manuscript 1984.

L. Nader (ed.), *No Access to Law: Alternatives to the American Judicial System*, New York: Academic Press (1980).

T. Palfrey and T. Romer, "Warranties, Performance, and the Resolution of Buyer-Seller Disputes," *Bell Journal of Economics*, Spring 1983.

I. P'ng, "Strategic Behavior in Suit, Settlement, and Trial," *Bell Journal of Economics*, Autumn 1983.

L.Ray (ed.), *Consumer Dispute Resolution: Exploring the Alternatives*, Washington, D.C.: American Bar Association (1983).

S. Salant, "Litigation of Settlement Demands Questioned by Bayesian Defendants," Caltech Social Sciences Working Paper #516, March 1984.

S. Salant and G. Rest, "Litigation of Questioned Settlement Claims: A Bayesian-Nash Equilibrium Approach," The Rand Corporation, P-6809, September 1982.

S. Shavell, "Suit, Settlement, and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs," *Journal of Legal Studies*, January 1982.

[17] There has been a recent explosion of interest in developing economic and game-theoretic models of these problems. As yet, these models are untested. *See*, for example, Shavell (1984), Salant and Rest (1983), P'ng (1982), and Cooter, Marks, and Mnookin (1982).

APPENDIX A

Instructions and Subject Record Sheets

Experiments 1-4

These instructions were used for Experiments 1 and 2. Instructions for Experiments 3 and 4 were identical, except that discussion of the "As Is" warranty was omitted. In Experiment 4, this type of warranty was described orally at the beginning of Period 5.

Record sheets were identical in all four experiments.

I.D. No. ________

Market Experiment

This is an experiment in market decision-making. If you make good decisions you can earn a considerable amount of money, which will be paid to you in cash immediately following the experiment.

In this experiment, you are going to participate in a market in which some of you are buyers and some of you are sellers. Your identification number is indicated on the top of this page. This tells you whether you are a buyer or seller. In the instructions that follow, the value of any decisions you might make is described, along with the specific rules for trading and record-keeping. *You are not to communicate with other buyers and sellers* except when necessary to complete transactions. Some of the information contained in these instructions is *private* information. Such information is written in red ink. You are not to reveal it to anyone.

Buying and selling will take place over a sequence of "trading periods." All the rules described below apply equally to each period.

Each period has two stages. In the first stage sellers offer units for sale to buyers and buyers make purchases. Each buyer may purchase at *most* two units in each period. Some of the units are "defective" and some of the units are "good." The *true* quality (good or defective) of a unit is announced only at the *end* of the second stage of the period in which it is sold. At the time of sale, neither the buyer nor the seller knows whether the specific unit being sold is good or defective. Any given unit has a *50%* chance of being a good unit and a *50%* chance of being a defective unit. When a seller wishes to offer a unit for sale, he or she will state a *price* and a *warranty policy*. There are three types of warranty policies the seller can offer:

1. An "as is" policy, which is equivalent to no warranty. If a buyer purchases a unit "as is," the buyer may not return it to the seller for a replacement.
2. A "Type 1" warranty entitles the buyer, if he or she wishes, to return

the unit to the seller and automatically receive a free replacement.
3. The third policy is called a "Type 2" and will be described in more detail in a moment.

In the second stage, all buyers who made a purchase in the first stage will be given a "clue" for each unit they purchased. A clue will help a buyer better evaluate whether a unit is defective or good. Buyers then decide whether or not to "exercise the warranty," that is, ask the seller to replace the unit. Notice that a unit purchased "as is" gives the buyer no choice for that unit, since there is no warranty to exercise. A unit with a warranty leaves the buyer free to keep the unit *or* exercise the warranty.

For each unit a buyer chooses to return under a warranty, the buyer and seller each must pay a small "return fee." For you, this fee is ______. What happens next depends upon whether the returned unit has a "Type 1" or a "Type 2" warranty.

(1) "Type 1" — In this case the buyer gives up that unit and automatically receives a replacement unit.
(2) "Type 2" — In this case, the seller is given a separate clue which enables him or her to better evaluate whether the unit is defective or good. If the clue suggests the unit is defective, the seller replaces the unit. If the clue suggests the unit is good, the seller does not replace the unit. Notice that the seller is not allowed to make a real choice here. The seller's decision depends entirely upon the seller's clue.

After the second stage is over, buyers are told, for each unit they hold, whether that unit is defective or good.

Specific Instructions to Buyers:

The amount of money you earn if you purchase a unit depends on whether it is defective or good. *The following is private information*. Each good unit has a redemption value __________ to you, and each defective unit has a redemption value __________ to you. The amount of money the experimenter pays you for each unit you buy is the redemption value of that unit *minus* the price you paid for it *minus* your return fee (if you exercised a warranty). If you make no purchase during a period, you receive nothing for that period. At the end of the experiment, add up the amount you earned for each unit and the experimenter will pay you that amount.

When you make a purchase, the unit will be assigned a number. You must record on your record sheet the *unit number*, *purchase price*, and *warranty*. Remember that you are allowed to purchase up to two units in each period. At the beginning of the second stage of each period, each of you who purchased one or two units in the first stage of that period will be provided with a clue about the quality of each unit. Each unit will have a separate clue. For each unit you bought, you will be given a card which has the unit number on it and either a "1" or a "0". Record the clue on your record sheet. Roughly speaking, a "1" is a positive clue, in the sense that good units usually get this clue. A "0" is a bad clue in the sense that defective units usually get that clue. Specifically, if a unit is good there is a *75%* chance that you will see a "1" on your card and a *25%* chance of a "0". If a unit is defective, there is a *25%* chance that you will see a "1" on your card and a *75%* chance of a "0". Thus a clue gives you some potentially useful information but does not tell you for sure whether the unit is good or defective. It is only a clue. Your clue is private information which you are not to reveal to anyone else.

Recall that if you purchased a unit "as is," you cannot return it, so its clue does not help you. On the other hand, if a unit has a "Type 1" or "Type 2" warranty, you may wish to use its clue to help you decide whether or not to return the unit. If you choose to return a unit, put an "X" in the column labelled "return" on your record sheet.

If you choose to return a unit, your earnings depend on whether the unit has a "Type 1" or a "Type 2" warranty.

(1) "*Type 1*": In this case the buyer gives up the unit he or she returned and automatically receives a replacement unit.

(2) "*Type 2*": If you return the unit under a "Type 2" warranty, the seller will then receive a separate card which has a "1" or "0" on it. Like your clue, the seller card will have a "1" on it with a *75%* chance, and a "0" with *25%* chance if the unit is actually good. If the unit is actually defective, the seller's clue will be a "1" with *25%* chance and a "0" with *75%* chance. However, just because a seller's clue probabilities are the same as the buyer's does not mean that the seller's card will always have the same clue as the buyer's card. ***The unit you returned will be replaced only if the seller's card has a "0" on it.*** The replacement unit has a *50%* chance of being good and *50%* chance of being defective. If you receive a replacement, record the new unit number on your record sheet.

Buyer earnings: For each unit a buyer holds at the end of the period (either originally purchased units or their replacement), he or she will be told whether that unit is good or defective. For each unit, record the outcome on your record sheet and calculate your earnings for that unit in the following way:

1. If you did *not* return the unit, your earnings are:

good redemption value — purchase price

if the unit is good, or

defective redemption value — purchase price

if the unit is defective.

2. If you returned the unit under warranty, and did *not* get a replacement, then your earnings are

good redemption value — purchase price — ________ (return fee)

if the unit is good, or

defective redemption value — purchase price — ________ (return fee)

if the unit is defective.

3. If you return the unit under a warranty, and receive a replacement, your earnings are

good redemption value — purchase price — ________ (return fee)

if the *replacement unit* is good, or

defective redemption value — purchase price — ________ (return fee)

if the *replacement unit* is defective.

After figuring your earnings for each unit, record the amount on your record sheet. Do not be surprised if your earnings for some units are negative. This will sometimes happen and is not necessarily your fault.

Specific Instructions to Sellers:

In the first stage of each period, you are free to sell as many units as you wish. When you wish to make an offer to sell a unit, raise your hand and I will call on you. Then state a ***price*** and a ***warranty*** ("as is", "Type 1" or "Type 2"). You are free to sell different units at different prices ***and*** with different warranties if you wish. Some offers you make may not be accepted or may be countered by another seller with a "better" offer. You are free to make new offers whenever you wish, but each time you raise your hand and are recognized by me you are allowed to make only one offer. If you wish to make more than one offer (e.g., a "Type 2" warranty offer at one price ***and*** an "as is" warranty offer at another price), you must be called on twice. Be sure you know exactly what you are going to offer before you raise your hand. If you hesitate for too long I will have to call on another seller. If you sell a unit, you must immediately record on your record sheet the unit number, the price you charged for the unit, and the warranty. The units you sell are not "free." You must pay a ________ ***production cost*** for each unit you sell.

In the second stage of each period, some of the buyers who purchased units from you under "Type 1" and "Type 2" warranties may seek replacement units. In case of a return, put an "X" in the "return" column of your record sheet. Each time this happens you must pay a ________ ***return fee*** and each time you replace a unit you must pay a second ________ ***production cost.*** Whether or not you replace a returned unit depends on the type of warranty it is under.

(1) "Type 1": You must always replace a returned unit if it was sold with a "Type 1" warranty.
(2) "Type 2": If the returned unit was sold with a "Type 2" warranty, you will be given a card that has a "1" or a "0"on it if the buyer returns for a replacement. This is your "clue". If the unit is actually good, this card will have a "1" on it with a *75%* chance, and a "0" with a *25%* chance. If the unit is actually defective, the card will have a "1" on it with a *25%* chance, and a "0" with a *75%* chance. You ***must*** replace the unit if your clue is a "0" and you do ***not*** replace it if your clue is a "1". If a replacement is made, record the unit number of the replacement on your record sheet.

Seller earnings:

At the end of the period, calculate your earnings for each unit you sold in stage 1. Sellers also receive an additional $1.00 for each period played.

1. If the buyer did not return the unit, your profit for that unit is:

   purchase price — (production cost)

2. If the unit was returned but not replaced, your profit on that unit is

   purchase price — (production cost) — ________ (return fee)

3. If the unit was returned ***and*** replaced, your profit on that unit is

   purchase price — (2 x production cost) — ________ (return fee)

Do not be surprised if your earnings for some periods are negative. This will sometimes happen and is not necessarily your fault. Add up your profits (or losses) on all units you sold in the period. Then add $1.00 and record the sum on your profit sheet. At the end of the experiment, add up the amount you earned in each period, and the experimenter will pay you that amount.

Buyer Record Sheet

Period	Unit # (Original)	Price	Warranty	Clue	Return	Unit # (Replacement)	Redemption Value of That Unit	Earnings
			⋮			⋮		

Seller Record Sheet

Period	Unit # (Original)	Price	Warranty	Return	Clue	Unit # (Replacement)	Unit Profit
		⋮			⋮		

Seller's Profit Sheet

Period #	Profit
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	

Total:

APPENDIX B

Instructions and Subject Record Sheets

Experiments 5-8

I.D. No. ________

Market Experiment

This is an experiment in market decision-making. If you make good decisions you can earn a considerable amount of money, which will be paid to you in cash immediately following the experiment.

In this experiment, you are going to participate in a market in which some of you are buyers and some of you are sellers. Your identification number is indicated on the top of this page. This tells you whether you are a buyer or seller. In the instructions that follow, the value of any decisions you might make is described, along with the specific rules for trading and record-keeping. *You are not to communicate with other buyers and sellers* except when necessary to complete transactions. Some of the information contained in these instructions is *private* information. Such information is written in red ink.You are not to reveal it to anyone.

Buying and selling will take place over a sequence of "trading periods." All the rules described below apply equally to each period.

Each period has three stages. In the first stage sellers offer units for sale to buyers and buyers make purchases. Each buyer may purchase at *most* two units in each period. Some of the units are "defective" and some of the units are "good." The *true* quality (good or defective) of a unit is announced only at the *end* of the third stage of the period in which it is sold. At the time of sale, neither the buyer nor the seller knows whether the specific unit being sold is good or defective. Any given unit has a *50%* chance of being a good unit and a *50%* chance of being a defective unit. When a seller wishes to offer a unit for sale, he or she will state a *price* and a *warranty policy*. There are two types of warranty policies the seller can offer:

1. A "Type 1" warranty entitles the buyer, if he or she wishes, to return the unit to the seller and automatically receive a free replacement.
2. The second policy is called a "Type 2" warranty and will be described in more detail in a moment.

In the second stage, all buyers who made a purchase in the first stage will be given a "clue" for each unit they purchased. A clue will help a buyer better evaluate whether a unit is defective or good. Buyers then decide whether or not to "exercise the warranty," that is, ask the seller to replace the unit.

For each unit a buyer chooses to return under a warranty, the buyer must pay the experimenter a "return fee". What happens next depends on whether the returned unit has a "Type 1" or a "Type 2" warranty.

(1) "Type 1" — In this case the buyer gives up that unit and automatically receives a replacement unit.
(2) "Type 2" — In this case, the seller is given a separate clue which enables him or her to better evaluate whether the unit is defective or good. If the clue suggests the unit is defective, the seller replaces the unit. If the clue suggests the unit is good, the seller does not replace the unit. Notice that the seller is not allowed to make a real choice here. The seller's decision depends entirely upon the seller's clue.

In the *third stage*, buyers who returned units under a "Type 2" warranty and did not receive a replacement may appeal to an Appeals Judge, if they wish. For each appeal, the buyer and seller must pay an appeal fee. For you, this fee is _________. If there is an appeal, the Appeals Judge gets a clue about the quality of the unit. Based on that clue, the Appeals Judge decides whether or not the buyer gets a replacement unit from the seller.The decision of the Appeals Judge is binding on both buyer and seller.

After the third stage is over, buyers are told, for each unit they hold, whether that unit is defective or good.

Specific Instructions to Buyers:

The amount of money you earn if you purchase a unit depends on whether it is defective or good. *The following is private information.* Each good unit has a redemption value _________ to you, and each defective unit has a redemption value _________ to you. The amount of money the experimenter pays you for each unit you buy is the redemption value of that unit *minus* the price you paid for it *minus* your return fee of _________ (if you exercised a warranty). You may not pay a price that exceeds your good unit redemption value. If you make no purchase during a period, you receive nothing for that period. At the end of the experiment, add up the amount you earned for each unit and the experimenter will pay you that amount.

When you make a purchase, the unit will be assigned a number. You must record on your record sheet the *unit number, purchase price,* and *warranty.* Remember that you are allowed to purchase up to two units in each period. At the beginning of the second stage of each period, each of you who purchased one or two units in the first stage of that period will be provided with a clue about the quality of each unit. Each unit will have a separate clue. For each unit you bought, you will be given a card which has the unit number on it and either a "1" or a "0". Record the clue on your record sheet. Roughly speaking, a "1" is a positive clue, in the sense that good units usually get this clue.

A "0" is a bad clue in the sense that defective units usually get that clue. Specifically, if a unit is good there is a *75%* chance that you will see a "1" on your card and a *25%* chance of a "0". If a unit is defective, there is a *25%* chance that you will see a "1" on your card and a *75%* chance of a "0". Thus a clue gives you some potentially useful information but does not tell you for sure whether the unit is good or defective. It is only a clue. Your clue is private information which you are not to reveal to anyone else.

You may wish to use its clue to help you decide whether or not to return the unit. If you choose to return a unit, put an "X" in the column labelled "return" on your record sheet.

If you choose to return a unit, what happens next depends on whether the unit has a "Type 1" or a "Type 2" warranty:

(1) "*Type 1*": In this case the buyer gives up the unit he or she returned and automatically receives a replacement unit.
(2) "*Type 2*": If you return the unit under a Type 2 warranty, the seller will then receive a separate card which has a "1" or "0" on it. Like your clue, the seller card will have a "1" on it with a *75%* chance, and a "0" with *25%* chance if the unit is actually good. If the unit is actually defective, the seller's clue will be a "1" with *25%* chance and a "0" with *75%* chance. However, just because a seller's clue probabilities are the same as the buyer's does not mean that the seller's card will always have the same clue as the buyer's card. ***The unit you returned will be replaced only if the seller's card has a "0" on it.*** The replacement unit has a *50%* chance of being good and *50%* chance of being defective. If you receive a replacement, record the new unit number on your record sheet.

If you returned a unit under a "Type 2" warranty and did not receive a replacement, you may—if you wish—appeal to the Appeals Judge. If you choose to appeal, you must pay an appeal fee of ________. If you appeal, put an X in the column labelled "Appeal" on your record sheet.

If you appeal, the Appeals Judge will receive a separate card which has a "1" or a "0" on it. The Appeals Judge's card will have a "1" on it with a ______% chance and a "0" with a ______% chance, if the unit is actually good. If the unit is actually defective, the Appeals Judge's clue will be a "1" with a ______% chance and a "0" with a ______% chance. ***The Judge will order the unit to be replaced if his card has a "0" on it.*** Otherwise, the buyer must keep the original unit. If you receive a replacement after an appeal, record the new unit number on your record sheet.

Buyer earnings: For each unit a buyer holds at the end of the period (either originally purchased units or their replacement), he or she will be told whether that unit is good or defective. For each unit, record the outcome on your record sheet and calculate your earnings for that unit in the following way:

1. If you did ***not*** return the unit, your earnings are:

good redemption value — purchase price

if the unit is good, or

defective redemption value — purchase price

if the unit is defective.

2. If you return the unit under a warranty, and receive a replacement, your earnings are

good redemption value — purchase price — _________ (return fee)

if the ***replacement unit*** is good, or

defective redemption value — purchase price — _________ (return fee)

if the ***replacement unit*** is defective.

3. If you returned the unit under warranty, and did ***not*** get a replacement and did ***not*** appeal, then your earnings are

good redemption value — purchase price — _________ (return fee)

if the unit is good, or

defective redemption value — purchase price — _________ (return fee)

if the unit is defective.

4. If you returned the unit under warranty, did ***not*** get a replacement, appealed, and *lost* the appeal, then your earnings are

good redemption value — purchase price — _________ (return fee)
— _________ (appeal fee)

if the unit is good, or

defective redemption value — purchase price — _________ (return fee)
— _________ (appeal fee)

if the unit is defective.

5. If you returned the unit under warranty, did ***not*** get a replacement, appealed, and ***won*** the appeal, then your earnings are

good redemption value — purchase price — _________ (return fee)
— _________ (appeal fee)

if the *replacement unit* is good, or

defective redemption value — purchase price — __________ (return fee)
— __________ (appeal fee)

if the ***replacement unit*** is defective.

After figuring your earnings for each unit, record the amount on your record sheet. Do not be surprised if your earnings for some units are negative. This will sometimes happen and is not necessarily your fault.

At the end of the experiment, add up the amount you earned on each unit. The experimenter will pay you this amount, plus $1.00 for each period played.

Specific Instructions to Sellers:

In the first stage of each period, you are free to sell as many units as you wish. When you wish to make an offer to sell a unit, raise your hand and I will call on you. Then state a ***price*** and a ***warranty*** ("Type 1" or "Type 2"). You are free to sell different units at different prices ***and*** with different warranties if you wish. Some offers you make may not be accepted or may be countered by another seller with a "better" offer. You are free to make new offers whenever you wish, but each time you raise your hand and are recognized by me you are allowed to make only one offer. If you wish to make more than one offer (e.g., a "Type 1" warranty offer at one price ***and*** "Type 2" warranty offer at another price), you must be called on twice. If you sell a unit, you must immediately record on your record sheet the unit number, the price you charged for the unit, and the warranty. The units you sell are not "free." You must pay a __________ ***production cost*** for each unit you sell. You may not sell a unit for less than this amount.

In the second stage of each period, some of the buyers who purchased units from you may seek replacement units. In case of a return, put an "X" in the "return" column of your record sheet. Whether or not you replace a returned unit depends on the type of warranty it is under. Each time you replace a unit you must pay a second __________ production cost.

(1) "*Type 1*": You must always replace a returned unit if it was sold with a "Type 1" warranty.
(2) "*Type 2*": If the returned unit was sold with a "Type 2" warranty, you will be given a card that has a "1" or a "0" on it if the buyer returns for a replacement. This is your "clue". If the unit is actually good, this card will have a "1" on it with a *75%* chance, and a "0" with a *25%* chance. If the unit is actually defective, the card will have a "1" on it with a *25%* chance, and a "0" with a *75%* chance. You must replace the unit if your clue is a "0" and you do ***not*** replace it if your clue is a "1". If a replacement is made, record the unit number of the replacement on your record sheet.

In the third stage of each period, some buyers may appeal after being refused a replacement under a "Type 2" warranty. If a buyer appeals on a unit sold by you, you must pay an appeal fee of ________. Put an "X" in the column marked "Appeal" on your record sheet. The Appeals Judge will then receive a separate card which has a "1" or a "0" on it. The Appeals Judge's card will have a "1" on it with a _____% chance and a "0" with a _____% chance, if the unit is actually good. If the unit is actually defective, the Appeals Judge's clue will be a "1" with a _____% chance and a "0" with a _____% chance. If the Appeals Judge's card has a "0" on it, you must replace the unit. Otherwise, you do not replace the unit. If a replacement is made, record the unit number of the replacement on your record sheet.

Seller earnings:

At the end of the period, calculate your earnings for each unit you sold in stage 1.

1. If the buyer did not return the unit, your profit for that unit is:

   purchase price — (production cost)

2. If the unit was returned but *not* replaced, and there was no appeal, your profit on that unit is

   purchase price — (production cost)

3. If the unit was returned *and* replaced without an appeal, your profit on that unit is

   purchase price — (2 x production cost)

4. If the unit was returned, *not* replaced, and there was an appeal, your profit on that unit is

   purchase price — (production cost) — ________ (appeal fee)

   if you *won* the appeal, or

   purchase price — (2 x production cost) — ________ (appeal fee)

   if you *lost* the appeal (i.e., had to make a replacement after the appeal).

Do not be surprised if your earnings for some periods are negative. This will sometimes happen and is not necessarily your fault. Add up your profits (or losses) on all units you sold in the period. At the end of the experiment, add up the amount you earned in each period, and the experimenter will pay you that amount, *plus* $1.00 for each period played.

Buyer Record Sheet

Period	Unit # (Original)	Price	Warranty	Clue	Return	Appeal	Unit # (Replacement)	Redemption Value of That Unit	Earnings
			⋮				⋮		

Seller Record Sheet

Period	Unit # (Original)	Price	Warranty	Return	Clue	Appeal	Unit # (Replacement)	Unit Profit
			⋮			⋮		

Comments

Robert J. Mackay
University of California, Berkeley

The papers in this section clearly reflect the maturing of the field of experimental economics over the past twenty years. The topics addressed involve important issues of public policy; the theories examined draw on recent advances in the economics of uncertainty and information; the experimental procedures are sophisticated and well developed; and, finally, the experimental results are clearly presented and interpreted. Each of these papers provides us with important insights into complex issues involving the determination and effects of variations in product quality. And, in each case, the insights result from the thorough and careful study of simple special cases that allow for sharp theoretical predictions while maintaining the structural core of the more complex problem. In short, these are fine examples of the likely successes that can be obtained from the application of experimental methods in economics.

After this general praise, some specific comments and criticisms seem in order. The Palfrey and Romer paper examines the case in which product quality is random but producers can respond to this uncertainty by adopting different types of warranties. They study the emergence of various warranty types, including dispute resolution mechanisms, and the effect of these warranties on product price. Their paper represents the state of the art in complexity of experimental design. Moreover, the experimental technology they have developed should prove to be quite valuable in future studies. Several useful extensions of their work come immediately to mind. First, the robustness of their results with respect to the probability of a defect, the accuracy of the inspection mechanism, the accuracy of the dispute resolution mechanism, and the cost of the dispute resolution mechanism should be examined. For example, the probability of a defect is very high in their experiments (*i.e.*, $t = .5$). It would be interesting to see how sensitive their results are to this parameter value. Second, their basic framework could easily be modified to examine the case of asymmetric information where sellers have a more precise estimate of their own product quality than do buyers. Finally, it would be interesting to examine the frequency of warranty types when suppliers can choose their *average* product quality. The interested reader will no doubt be able to think of many other variations and extensions of Palfrey and Romer's experimental design.

The Holt and Sherman paper examines the case in which product quality is random and cannot be determined by the seller prior to sale. Instead, product quality can only be determined by the buyer after purchase. The seller, however, can respond to this uncertainty and avoid the losses resulting from the rejection of low quality units by bundling the units and allowing only bundles to be rejected.

Bundling, in fact, is efficient in the special case they examine since the average quality can be provided more reliably for a bundle than for an individual unit. Their experimental results reveal a strong tendency for bundles to emerge rather than single-unit transactions. Two comments seem worth making. First, the major shortcoming of the Holt and Sherman paper is the lack of repetition of the experiment. More runs of the experiments, examining the robustness of the results to variations in model parameters, would be useful and provide more confidence in the authors' conclusions. Second, it would also be interesting to see how their results would vary with monopolistic supply rather than competitive supply. This modification, moreover, would allow for a more direct comparison of their results with the Kenney and Klein analysis of the bundling of diamonds by De Beers.

The Lynch, Miller, Plott and Porter paper examines the case in which producers can choose quality, but buyers cannot verify quality until after purchase. The authors examine the role of warranties and reputation formation, through provision of seller-specific information on the choice of product quality and price. Their results nicely illustrate both the lemons phenomenon, when seller identities are unknown, and the ability of enforceable warranties to resolve the problem. They also show that the seller reputations, based on public or private information, can have important effects on market outcomes even if they fail to completely resolve the problems associated with asymmetric information.

Their results on reputation formation raise important issues. The crucial point to note is that theoretically we do not have good models of reputation formation in the type of situation they examine. Our most well-developed models rely on arguments based on infinite repetition or, at least, positive probabilities of continuing the interaction. Obviously, neither of these conditions apply to the experimental situation they examine. With a finite number of repetitions, our models predict no reputation formation when players reason backwards from the end of the game. Yet the authors' experimental results clearly show a potential effect of reputation on product quality. In this regard, their results remind one of similar experimental results on collusive behavior in oligopolistic settings. This combined work clearly illustrates the importance of reputation to the most basic issues in consumer protection and antitrust. Additional theoretical and experimental work that focuses on the determinants and effects of reputation in repeated games should yield significant payoffs.

Comments

Ross M. Miller
Boston University

Rather than discuss the papers by Holt and Sherman and by Palfrey and Romer one at a time, I will bundle my comments on the two papers around some specific comments on experimental methods in economics. The papers in this session, unlike the others, provide empirical evidence gathered in a laboratory setting and, as such, are subject to the scientific method more usually thought of in connection with the physical and biological sciences. The three aspects of the scientific methods that I will focus on are controls, replication, and relevance.

The use of control experiments is a practice that I feel should be encouraged. There are, however, very good reasons why experimentalists run fewer controls than would be ideal. First, some controls, such as double-blind experimentation (a virtual necessity for medical research), are difficult to implement with the current technology for running experiments in the social sciences. It is probable that only through computerized experimentation, such as with the PLATO system used by Vernon Smith, will true double-blind experiments be possible.

Second, any experiment, examined carefully, will present many factors that may be subject to control. Because experiments are costly, in both payoffs to subjects and the experimenter's time, an experiment run to demonstrate a new result often takes precedence over one that simply confirms a known result, especially when the result of the control appears obvious in light of the behavior observed in the original experiment. In the Holt-Sherman work, the question that comes to mind is whether the bundling of goods is not just a naturally appealing notion to people and occurs even when it results in inefficiency. Certainly a biological argument can be made that the process of natural selection has favored and will continue in the foreseeable future to favor individuals who possess at least a minimal propensity for "bundling". If individuals like to bundle items, the efficiency of prohibitions on bundling depends on the degree to which the bundling impulse is considered economic and, to use a loaded word, rational.

Finally, a problem with controlling economic experiments is that the number of factors that can be subject to control is potentially enormous. For example, different experimenters have used different terms to refer to the high quality and low quality units where two qualities are available. The zero/one dichotomy used by Palfrey-Romer and the Regular/Super dichotomy that I favor are intended to reinforce subjects with the notion that one type of good is worth more than the other. The Red/Black dichotomy used by Holt-Sherman is possibly more neutral except to businessmen and electricians, and even they could not agree on which color is positive. I do not lose any sleep over such considerations because my impression is that they are minor, but certainly there are psychologists who do.

A problem that I do lose sleep over, particularly with regard to experiments as complex as those of Palfrey-Romer, is the replication of experimental results.

Although the rational expectations and related literature has portrayed humans as capable of acting as if they solve complex and subtle strategic problems, considerable trial and error with the proper feedback may be required before a human being can learn sophisticated strategies. Because there are so many paths and outcomes of the learning process, experiments where learning can play a major role appear to be difficult to replicate between subjects pools or even within a subject pool.

The complexity of the Palfrey-Romer experiments is demonstrated by the persistence of Type 1 ("full") warranties in many of the experiments. For Type 2 ("limited") warranties to drive out Type 1 warranties, buyers must be able to logically deduce or learn through experience that the limitations on the Type 2 warranties, which lower the costs to sellers, do not reduce the value to buyers. Subjects appear to differ in their ability to make such an inference, and so the behavior observed by Palfrey-Romer is less uniform than in a simpler auction setting.

Finally, I will look at the relevance of economic experiments. The two papers are excellent examples of how experiments can be relevant to policymakers, both by their direct application to policy problems and indirectly by causing economists to adjust their theoretical models.

It is significant that both papers examine theories that were developed only recently; in fact, Palfrey and Romer examine a theory from their own recent work. Further, as has been the case in the physical sciences, experimental work in economics is generating results that do not fit existing theories and thus may lead the way for new theories. These new theories are necessary as a basis for examining economic policy and as a foundation for econometric studies.

An unenlightened view of these studies would conclude that because the experiments do not capture the fullness of most naturally-occurring markets they are irrelevant. The problem with constructing laboratory approximations to naturally-occurring markets is that the isolation and testing of individual hypotheses is not facilitated. The feature that both studies isolate, in significantly different settings, is the sharing of risks between buyers and sellers that occurs in a competitive market setting. For warranties, my colleagues and I have compiled a list of over twenty suggestions of features that characterize markets for warrantied products and the list keeps growing. The Lynch, Miller, Plott, and Porter work emphasizes the reputation and moral hazard aspects of warranties. As Palfrey and Romer note, experimental work that combines the risk-sharing with moral hazard is a natural topic for further work. Another favorite of mine is differential information; it would be interesting to see what happens when sellers have better signals of quality than buyers.

Certain relevant features of markets are difficult to incorporate into experimental markets. In particular, zero-profit equilibria are hard to model in economic experiments because then only through lump-sum payments can adequate payoffs be made to sellers.

The Impact of Product Recalls on the Wealth of Sellers

Gregg Jarrell and *Sam Peltzman*
Securities and Exchange Commission — University of Chicago

I. Introduction

This paper has a simple goal: to estimate the losses borne by owners of a firm that recalls a defective product from the market. While we stick close to the "facts," we hope they will shed some light on an important issue in the regulation of consumer protection. This is the extent to which information about product quality is sufficient to deter production of faulty products. In many areas, including the two which are the focus of our analysis—drugs and autos—there is extensive regulation of product quality prior to marketing of the product. One normative justification of such pre-market regulation would be that mere disclosure of any defects after a good is marketed does not impose sufficient costs on the marketer to deter optimally the production of defective products. Such sub-optimal deterrence could occur if, for example, consumers were insufficiently sophisticated in assimilating information about defects or the tort liability system insufficiently compensated them for resulting damages.

While we do not address these normative issues directly, we hope that our results will be useful in assessing the magnitude of the potential normative "problem." Accordingly, we will compare our estimates of losses to owners with independent estimates of some elements of the cost associated with the product defects. In particular, we are able to estimate at least the rough magnitude of elements of the direct costs to firms of recalling defective products. These would include the costs of destroying contaminated batches of drugs, the costs of repairing defective cars, etc. An obvious question—and a test of capital-market efficiency—would be whether the capital market internalizes these costs. If it fails to do so, any presumption of suboptimal deterrence would be strengthened. In many cases, these direct costs will be a major component of the "social" costs of the defect. For example, the repair costs for a potentially defective auto part could even exceed the relevant social costs if the probability that the part will fail is very low. But some cases involve potentially large indirect costs for consumers—*e.g.*, health damages from a dangerous drug. In these cases, optimal deterrence would require a penalty greater than the direct costs we are able to estimate.

The main focus of the paper is on the changes in shareholder wealth which accompany recalls of automobiles and drugs (prescription, over-the-counter and medical devices). We chose these particular products, because each yields a

*We thank James Frieden, Michael Ryngaert, Thaddeus Niemira and Monica Noether for their assistance.

good-sized sample of recalls and because we could obtain associated data on some elements of the direct costs of most of these recalls. The samples also differ in an interesting dimension: drug recalls occur much less frequently (per firm) than auto recalls. Important examples of the latter occur every few weeks or months, while the former occur once or twice in a decade.

Our primary finding is that the capital market in fact penalizes producers of both recalled drugs and autos far more than the direct costs. Indeed, the capital market penalty seems so great that it may even exceed a plausible independent estimate of the relevant social costs. We do not press this point, because we have only the most fragmentary data on the relevant indirect costs and because we want to avoid the implicit issue of capital-market efficiency. But to the empirical question "how much deterrence does the capital market provide against the sale of faulty products?", the answer implied by our data must be "considerable."

We also find that competitors of drug and auto firms with recalled products are not helped by their rival's travail. In fact, in both cases they bear substantial losses.

II. How "Should" the Stock Market React to News of a Product Recall?

The stock market does not react to every event which entails a cost to shareholders, only to those which are not entirely expected. So if product recalls occurred with the same regularity as, say, wage payments we would no more expect stock prices to fall when a recall occurs than on payday, even though both events impose real costs on stockholders. The market can be expected to respond to news of recalls only if the news resolves some uncertainty. And, since recalls are not entirely unexpected, the response will ***understate*** the costs of the recall to stockholders. To see this, let any uncertainty be resolved within a "month," and suppose that only one of two things can happen to a firm next month: either a product is recalled at some cost (K) to shareholders or there is no recall. So, the firm's month-end stock price (S_1^i) will be either:

(1) $S_1^{NR} = V$ if no recall occurs, or

(2) $S_1^{R} = V\text{-}K$ if a recall occurs,

where

V = present value of the firm's profits including all expected recall costs ***except*** those occurring next month, and where we assume independence of successive monthly events.

The firm's stock price at the ***beginning*** of the month is the present value of future profits, or

(3) $$S_0 = p(V\text{-}K) + (1\text{-}p)V = V\text{-}pK,$$

where

p = probability that a recall occurs next month.

Thus, if a recall occurs next month, the stock price will change by (2)-(3) above, or

(4) $$S_1^R - S_0 = -(1\text{-}p)K$$

i.e., by the ***unexpected*** component (1-p) of the recall cost. Only if the recall is entirely unexpected (p = 0) will (4) = K. In months where recalls do ***not*** occur, stockholders get a capital gain of (1)-(3), or

(5) $$S_1^{NR} - S_1^0 = pK$$

So, to get at the market's estimate of K, we would need to subtract (5) from (4)—*i.e.*, to compute the *difference* between the return in months with and without recalls.

In practice, (4) and (4)-(5) will be about equal if p is small. This is the case with drugs where our data indicate that most uncertainty is resolved within a month and where no company in our sample has been involved in more than two distinct recalls in a period of about 100 months. For this sample, then, we use conventional "event study" methodology, more fully described below, in which we, in effect, estimate just (4). But auto recalls are far less of a surprise than drug recalls, so we attempt to estimate (4)-(5) for that sample.

III. Drug Recalls

A. *Selection of Drug Recall Sample*

When a drug product is found to be defective, the manufacturer is required to remove that product from the market. This recall can be initiated either by the manufacturer or the Food & Drug Administration (FDA), and it can involve anything from a few bottles of contaminated or mislabeled product to the permanent removal of a product from the marketplace. The FDA classifies recalls by health hazard: Class I recalls involve product defects that may have seriously adverse health consequences including death; Class II recalls involve temporary or medically reversible health hazards while Class III cases are unlikely to entail adverse health consequences.

Our sample focuses on recalls that involve a serious health hazard and/or a relatively large amount of product. We selected the sample from weekly reports of FDA Recalls and Court Actions in the *Food, Drug and Cosmetic Reporter*, an industry newsletter commonly called the "Pink Sheets." Recalls were included in our sample if the Pink Sheets report gives an estimate either of the direct costs of the recall or, more commonly, of the number of units recalled. In addition, we include those recalls where direct cost estimates are provided in the *Wall Street Journal*. Our sample period runs from 1974 through 1982.

Our sample consists of most of the largest and more hazardous recalls in this period. For example, Class I recalls account for over half of our sample, while they account for less than 2% of the over 3,000 FDA citations reported between 1973 and 1978.[1] Many of our cases received considerable publicity. Over half of our 32 cases were covered by the Wall Street Journal. Five of these cases were serious enough so that the recalled products were withdrawn indefinitely from the market. Table 1, which is elaborated below, summarizes this sample. It shows the names of the manufacturers of the recalled drugs in our sample, the event dates and estimates of the stock market response and direct costs of the recalls incurred by these manufacturers as a result of the recalls. We exclude cases without stock returns data for the manufacturer. This requires that the manufacturer be a publicly-traded firm having stock returns data for a 100 trading day period centered on the day that the recall becomes public information.

B. *Choosing Event Dates*

For each recall, we sought to identify the earliest date at which news of a recall might have become public. This could precede the date on which a recall actually began or was ordered by the FDA. For example, the first hint that a recall may eventually occur might be press reports (we use the *Wall Street Journal*) implicating a drug (*e.g.*, Tylenol) in a health problem (poisoning). In such a case, the date of the press report is our "event date." In general, we use the date of the earliest press story on the troubled product, when there is a WSJ story. For most cases, this is the same date that the recall begins. For recalls not covered by WSJ stories, the event date is the earliest date on which the FDA notified a manufacturer to recall a product. This date is taken from the Pink Sheet story that reports the citation. If these initial FDA communications appear to be strictly private correspondences, then we use the publication date (usually a week or so later) of the Pink Sheet reporting the recall as the event date for these non-WSJ cases.

Sometimes news about essentially the same product defect is spread out over time. For example, two defective batches of a product are found several weeks apart (cases 2.1 and 2.2) or a product defect is found a month before the firm decides that a recall is necessary (26.1 and 26.2). We treated these related events as separate events (and split direct costs evenly among them) if more than three weeks elapsed between the events. These are identified by case numbers with

[1] Lawrence H. Block, "An Evaluation of Drug Product Citations in the FDA Weekly Reports Between 1970 and 1978," *Contemporary Pharmacy Practice*, Vol. 3, No. 3, (Summer, 1980), pp. 171-79.

decimals in the table. (We treat related events less than three weeks apart, like the rest, as a single event beginning on the earliest date of adverse news.)

C. *Direct Costs of Recalls*

For most cases, we estimate the "direct cost" of a drug recall by assuming that all of the violative units become worthless upon recall. Specifically, where the Pink Sheet citation reports the number of units of the violative batch that are in distribution channels, we multiply this figure by the wholesale price of the product as reported in the appropriate yearly issue of the ***American Druggist's Blue Book*** and the ***Drug Topics Red Book*** to estimate "direct costs."

For some of the more publicized recalls, information on the direct cost was available from news stories, because the companies took an extraordinary charge to their income. (Case numbers 6, 7, 14, 26 have their direct losses taken from the WSJ.) For instance, the WSJ reported on 10/29/82 that it would cost Johnson & Johnson about $50 million to recall and destroy 22 million units of Extra-Strength Tylenol capsules. In addition, it was reported that new tamper-proof packaging, additional television advertising, and related efforts to rebuild consumer confidence would cost another $50 million. Therefore, our estimated direct cost to Johnson & Johnson of the Tylenol recall is $100 million.

We make no allowance for tax benefits due to recall costs. Where we use reported extraordinary charges, we use the pre-tax figure, and we ignore any tax savings from inventory losses. Accordingly, our direct cost estimates may be over-generous.

For each recall Table 1 gives the estimated direct cost in dollars (last column) as well as in percent of the market value (just before recall) of the respective manufacturer's common stock. In both dollar and percentage terms, the recall of Procter & Gamble's Rely Tampon (14) entails the largest estimated direct cost ($150 million, 2.5 percent of market value) in our sample, while the Class I recall of Abbott Labs' Plasmatein (1) is the least costly ($5,000, .0005 percent).

D. *Capital Market Returns*

The full costs to manufacturers of recalled drugs is measured by net-of-market (or excess) stock returns in the period surrounding public announcement of the recall. These excess returns are obtained from the Scholes excess return file at the University of Chicago's Center for Research in Security Prices (CRSP).[2] We cumulate excess returns for each manufacturer over several "event windows" of different intervals to allow for pre-event leakage or post-event revision. The most narrow event window is six days, from t = -2 to t = 3, where

[2] For cases 3, 4, 8.1, 8.2, stock returns are unavailable from this source. So we constructed excess return series for these cases by subtracting the return to the New York Stock Exchange Index from returns to these firms' stocks.

Table 1

SAMPLE OF DRUG RECALLS WITH FIRM NAMES, EVENT DATES, STOCK RETURNS, ESTIMATED DOLLAR LOSS, AND WSJ DUMMY

Case #	Firm Name	Event Date	No WSJ[1] Dummy	Direct Cost as Percent of Market Value	Recall Firm's CER % -4 to +5	Estimated Direct Dollar Cost (000's)
1.	Abbott	4/28/76	1	0.00%	-1.91%	$ 5
2.1	Am. Hospital	11/26/74	1	0.03	-4.26	172
2.2	Am. Hospital	1/13/75	0	0.05	-7.10	430
3.	Block	6/11/79	0	2.33	-7.76	2330
4.	Bolar	12/24/80	0	2.34	-10.63	1600
5.	Lilly (V-Cillin)	12/5/77	1	0.05	-1.42	1150
6.	Lilly (Oraflex)	8/3/82	0*	0.66	-10.46	30400
7.	Johnson (Tylenol)	10/1/82	0*	1.35	-15.88	100000
8.1	Mallinckrodt	11/14/74	1	0.00	3.55	10
8.2	Mallinckrodt	12/10/74	0	0.57	-8.43	1300
9.	Merck	8/8/80	1	0.01	-2.76	353
10.1	Milton Roy	7/15/76	1	0.05	3.64	12
10.2	Milton Roy	8/25/76	0	0.05	-1.98	12
11.	Morton Norwich	11/23/79	0	0.06	-2.75	250

12.	Johnson (Ortho)	10/13/75	0	0.13	2.65	6500
13.	Parke Davis	8/13/76	0	0.04	-0.63	1000
14.	Procter-G. (Rely)	9/18/80	0*	2.46	-5.29	150000
15.	Richardson	10/1/78	0	1.68	-8.18	11500
16.	Richardson	9/26/80	1	1.78	-2.35	9530
17.	Robins	4/28/76	1	0.77	-0.33	2100
18.	Searle Labs	11/17/76	1	0.09	-6.56	600
19.	Searle Labs	6/9/81	1	0.07	-1.00	1000
20.	SmithKline	4/26/79	0	0.13	-8.93	3700
21.	Squibb	1/16/75	0	0.24	-9.40	3500
22.	Squibb	11/28/77	1	0.05	-12.11	560
23.1	Sterling	1/14/76	1	0.55	-18.20	6000
23.2	Sterling	2/18/76	0	0.55	-18.84	6000
23.3	Sterling	4/14/76	1	0.03	-2.15	270
24.	Sterling	4/5/78	1	0.02	-6.77	200
25.	American Home Prod.	3/4/82	0	0.05	-0.69	3044
26.1	Robins	5/29/74	0*	0.47	-18.12	2550
26.2	Robins	6/28/74	0*	0.47	-11.16	2550

* 0 = Recalled products were withdrawn from the market.

[1] 1 = Recall *not* reported in WSJ.

t = 0 is the formal event date of the recall. The widest event window is from t = -49 to t = 50.[3]

Table 2 presents mean cumulative excess returns (CER) for various event windows. These are negative for every window from week to 5 months around the event date. But the two-week window (CER(-4, 5)) yields a loss roughly within a percentage point of that for any wider window. This means that essentially all of the market response to the event is compressed into the two surrounding weeks. In addition, there are no systematic "mistakes"—*i.e.*, there is no systematic recovery of some of these losses in the 50 days after the event date (or else the CER (-49, 50) would be smaller than CER (-4, 5)). Finally, note that fully nine-tenths of the sample suffers a loss in the two weeks surrounding a recall (*see* Table 3). It is clear that recalls constitute adverse news for stockholders and that most of the uncertainty about them is resolved in the two weeks surrounding public disclosure of the recall.

It is also clear that the capital losses are substantial by any standard. In particular, they are much larger than our generous estimate of direct costs. The mean CER (-4, 5) is -6.13%, which is fully twelve times the mean relative direct cost of 0.53% (and over 50 times the median). We never fully succeed in explaining this enormous gap.

The last line of Table 2 contains another mystery. This shows the CER (-4, 5) for an equally-weighted portfolio of drug stocks. We conjectured that competitors might benefit from the adversity visited on the seller of the recalled product. Instead, the spillover seems negative. All drug stocks suffer a (significant) mean loss of just over 1 percent in the two weeks surrounding a recall. This cannot be explained by any tendency for recalls to be bunched (in which case one recall would beget expectations of others).[4]

The disproportionate size of capital market losses relative to estimated direct costs led us to see if the capital market losses are systematically related to the degree of publicity surrounding the recall or to whether there was a complete product withdrawal. These may be proxies for costs which we cannot estimate. For example, a complete withdrawal may engender losses to specific assets (*e.g.*, research and development, past advertising) which are not written off. Table 3 presents the mean CERs over three sub-samples of recalls. The mean CER

[3] Some of the wider event windows result in overlap of the related events denoted by decimal case numbers in Table 1. In these cases, we (1) arbitrarily split the time between events in half and attributed the excess return for any day to the event closest in time, and (2) set the remaining excess returns to zero. For example cases 26.1 and 26.2 occur 22 trading days apart. Excess returns for the first 11 days after 5/29/74 are attributed to 26.1 and all subsequent excess returns are set = 0 for that case. Excess returns for the 11 days ending 6/28/74 are attributed to 26.2 and all preceding excess returns are set = 0 for that case. In this way, we avoid double counting of the same excess return.

[4] We have 26 unrelated events in the 9 years 1974-82, or about 3 per year. If recalls were being generated by a Poisson process with a mean of (26/9) per year, the standard deviation would be 1.7. This differs insignificantly from the sample standard deviation of 1.45, so the distribution of recalls seems essentially random.

Table 2

MEANS AND DISPERSION MEASURES FOR CER TO DRUG RECALL FIRMS (VARIOUS INTERVALS), RELATIVE DOLLAR LOSS, AND CER TO DRUG PORTFOLIOS

Variable Name[1]	Mean (%)	S.E. Mean[2] (%)	t-stat. Mean	Minimum (%)	Maximum (%)
CER (-49, 50)	-6.742	3.113	-2.17	-35.94	28.40
CER (-29, 30)	-5.479	2.411	-2.27	-30.84	15.85
CER (-14, 15)	-7.147	1.969	-3.63	-22.46	5.06
CER (-9, 10)	-6.563	1.392	-4.71	-24.10	9.05
CER (-4, 5)	-6.132	0.985	-6.23	-18.84	3.64
CER (-2, 3)	-2.832	0.696	-4.07	-15.39	3.26
BDRUG (-4, 5)	-1.170	0.335	-3.49	-8.53	5.86
Relative Loss %	0.534	0.135	3.95	0.00	2.46

[1] a) CER (-X, Y) is the cumulative excess return (from Scholes Excess Returns Tape, University of Chicago CRSP) from X trading days before to Y trading days after the recall event.

b) BDRUG (-4, 5) is the cumulative excess return to an equal-weighted portfolio of all NYSE or ASE drug manufacturers having a SIC of 2834, 2840, or 2841 (about fifty firms). The cumulative excess return to this drug portfolio is computed from t = -4 to t = 5 for each date on which there was a drug recall that is included in our sample. The drug firm subject to the recall is excluded from the drug portfolio when computing BDRUG for each particular recall event.

c) Relative Loss % is the estimated direct loss expressed as a percent of the market value of the equity of the recall firm. The market value is computed 40 trading days before the recall event date.

[2] The standard error σ of the mean CER (-X, Y) is computed using the formula:

$$\sigma(-X, Y) = \left[\sum_{i=1}^{N} \sigma_i^2 \right]^{1/2} / N$$

where σ_i^2 is the variance of the i^{th} recall firm's excess stock return and N = 32 recalls. σ_i^2 is estimated for each firm by using daily excess returns from t = -49 to t = -5 and t = 5 to t = 50. Let S_i^2 be the variance of the above-defined time series of daily excess returns. Then σ_i^2 is computed by multiplying S_i^2 by T, where T is the number of trading days in the particular event window. (T is 10 for CER (-4, 5), 20 for CER (-9, 10), and so on.) These standard errors are virtually identical to the standard errors of the sample-mean CERs.

Table 3

MEAN CER TO DRUG RECALL FIRMS FOR VARIOUS EVENT INTERVALS BY CLASSES WITH AND WITHOUT WSJ STORIES AND PRODUCT WITHDRAWALS

CER Interval	Mean CER to Recall Firm % All	No WSJ No Wdraw	WSJ No Wdraw	WSJ Wdraw	Percent of All CERs Negative	t-ratio*
-2 to 3	-2.83	-1.16	-3.62	- 5.45	84.4	5.36
-4 to 5	-6.13	-3.76	-6.36	-12.18	90.6	7.87
-9 to 10	-6.56	-3.73	-7.58	-11.86	84.4	5.36
-14 to 15	-7.15	-3.11	-8.80	-14.15	71.9	2.76
-29 to 30	-5.48	0.97	-8.06	-16.82	71.9	2.76
-49 to 50	-6.74	-1.19	-8.41	-17.95	62.5	1.46
Number of Observations	32	14	13	5	32	32

* T-ratio is the percent of All negative minus 50% divided by the standard error (S) from a binomial distribution: *i.e.*, $S = (PQ/N)^{1/2}$, where P is the percent negative, Q is the percent positive, and N is the number of cases (32).

(-4, 5) for the 14 recalls that were not covered by the WSJ and that did not involve withdrawal is -3.76%, while the mean CER for the 13 no-withdrawal recalls covered by the WSJ is -6.36%. The mean CER for the five recalls that resulted in product withdrawals is -12.18%. (All five withdrawals were covered by the WSJ.) These mean CERs imply that a WSJ story engenders an additional CER of -2.6%, and that the product withdrawal distinction adds another -5.8%. So both extra publicity and a withdrawal are costly. But the remaining cases still entail an enormous discrepancy between the capital loss and direct costs.

Table 4 reorganizes the data in Table 3 into specific subperiods. It shows that stocks typically decline both in the week before and the week after our event date. If our event date is, as we intend, the *earliest* date of public information, then Table 4 implies that there is some prior leakage of non-public information. In no other subperiod from t = -50 to t = +50 is there as large a change as in either of the two weeks around the recall date, and only in these two weeks does the frequency of negative CERs exceed significantly what would be expected from a random process.

E. *Cross-sectional CER Regressions*

We investigate the relationship between capital market losses and direct costs more formally in Table 5. Here we regress CERs on the relative dollar loss together with dummy variables for publicity (= 1 if there was no WSJ story) and withdrawal and the CER (-4, 5) to the portfolio of other drug firms. This latter is not really an exogenous variable, given the previously documented spillover effect of recalls. But we include it to account crudely for the industry--specific component of the total loss (as well as "other" industry-specific news). Table 5 confirms the tendency for both publicity and product withdrawal to be costly, though some of the standard errors are large enough to caution against pushing these conclusions too hard. The main new result is the negative relationship between the CERs and the relative direct cost variable. This is consistent with our prior expectation, but the coefficient implies that an extra dollar of direct cost adds $2 to $4 to the stockholders' loss. This implies that even our generous estimates of direct losses are systematically low, but correlated positively with the "true" cost of a recall. The first line of Table 6 sheds further light on the relationship between the stock market loss and direct costs. It reveals that the latter are higher for publicized recalls and for withdrawals. The relevant coefficients are statistically weak, but they are large relative to the mean direct cost. This implies that part of the extra costs of publicity and product withdrawals shown in Table 3 are due to the tendency for these recalls to have larger direct costs.

The larger message of Tables 5 and 6 confirms that of the crude data in Table 3. It is that stockholder losses from recalls go beyond costs which can be attributed to the specific product. Note, from Table 6, that the CER to competitors is much more weakly related to the case-specific variables than is the recall firm's CER. This implies that any recall, regardless of "size," engenders a roughly

Table 4

MEAN CER TO DRUG RECALL FIRMS FOR SUCCESSIVE INTERVALS AROUND EVENT DATE BY CLASSES WITH AND WITHOUT WSJ STORIES AND PRODUCT WITHDRAWALS

CER Interval	Mean CER to Recall Firm % All	No WSJ No Wdraw	WSJ No Wdraw	WSJ Wdraw	Percent of All Negative	t-ratio*
-50 to -35	-1.94	-4.69	-0.13	1.07	50.0	0.00
-34 to -20	1.78	2.65	1.72	-0.50	34.4	-1.86
-19 to -5	-1.16	-0.22	-1.86	-1.97	53.1	0.35
- 4 to 0	-2.36	-0.85	-3.23	-4.29	71.9	2.76
1 to 5	-3.78	-2.91	-3.13	-7.89	81.3	4.54
6 to 19	0.05	1.92	-1.51	-1.13	59.4	1.08
20 to 34	1.02	2.17	1.36	-3.08	34.4	-1.86
35 to 50	-0.36	0.73	-1.62	-0.15	43.8	-0.71
Number of Observations	32	14	13	5	32	32

* *See* Table 3 for explanation.

Table 5

STOCK RETURNS TO DRUG RECALL FIRMS REGRESSED ON RELATIVE DOLLAR LOSS, DRUG PORTFOLIO RETURNS, AND DUMMY VARIABLES FOR NO WSJ STORY AND FOR PRODUCT WITHDRAWAL

	Independent Variables					Summary Statistics	
Dependent Variable	Constant	Direct Cost as % of Market Value	No WSJ Dummy	Withdrawal Dummy	CER (-4, 5) Drug Portfolio	Adj. R^2 F-value	Mean of Dep. Var./ No. of Obs.
CER (-14,15)	-0.052	-0.039	0.040	-0.045	0.735	0.248	-0.071
	(-1.96)	(-1.95)	(1.34)	(-1.13)	(1.64)	3.56	32
CER (-9, 10)	-0.036	-0.035	0.022	-0.040	1.091	0.251	-0.066
	(-1.50)	(-1.94)	(0.81)	(-1.11)	(2.68)	3.59	32
CER (-4, 5)	-0.040	-0.020	0.016	-0.058	0.693	0.250	-0.061
	(-2.09)	(-1.38)	(0.76)	(-2.01)	(2.14)	3.58	32

similar industry wide asset loss. Further, even after allowing for the understatement of direct costs implied by the coefficients in Table 5, we do not come close to rationalizing the 6% average loss of a recall. That is, the regressions imply that an unpublicized recall which does not result in a withdrawal and has trivial direct costs still entails a loss of over 3 percent (based on CER(-4, 5)). Of course, our case-specific variables may be leaving out important product-specific costs. For example, they exclude any estimate of expenses for product liability suits. But we have to doubt that these can amount to much for a case involving a small defective batch of an otherwise safe product. We suspect that the major impact of product liability costs is showing up in the large coefficient of direct costs and the *extra* losses due to product *withdrawals*. Every withdrawal in our sample has engendered well publicized product liability suits.

For one of these cases, we have a full profile of product liability costs. Even though samples of one yield notoriously noisy estimates, it seems worth exploiting these data to get a sense of the likely magnitude of this specific cost. The case involves the Dalkon Shield, an intrauterine birth control device which was implicated in deaths of some users. The two events in our sample (26.1 and 26.2) emanating from this product withdrawal generated CER (-4, 5) values of -18 and -11 percent, or a total loss of around $150 million to the manufacturer, A. H. Robins. Robins took a pretax charge in 1974 of $5.1 million for the costs directly related to withdrawing the product and destroying inventory, and these are shown in Table 1. The company also agreed with the SEC to break out all expenses (extra legal fees and uninsured liability payments) related to litigation over this product in its financial statements. It has done this in every annual report from 1976 to date. The total of the pre-tax charges reported for 1976-82 is $29 million. If the stream of these expenses is discounted at 10 percent per year back to 1974, when the recall occurred, we obtain a 1974 present value of $17 million. A simple regression of the log of the annual elements of this expense stream against time implies a mean increase in these expenses of 21 percent (SE = 11 percent) per year. We then assumed that expenses would continue to be incurred for another five years and would equal the predicted values from this regression in each year from 1983 through 1987. These assumptions imply an additional $21 million of liability costs in 1974 present value, bringing the total to $38 million.

This exercise tells us that, in (partial) hindsight, a reasonably complete independent estimate of the full costs of the recall to Robins is on the order of under 1/3 the stock market loss. (Since Robins has had an average tax rate of over 40 percent in recent years, even this is too high.) So, if the product liability component of this cost is anything like the consumer cost of the product defect, the stock market loss appears to exceed the "social loss." While we hesitate to push these fragmentary data this far,[5] they imply that the stock market losses

[5] A fuller treatment would require us to see if announcement of the liability costs affected the returns to Robins' stock. For example, if the initial reaction overestimated these costs, subsequent announcements of the actual costs would engender positive excess returns.

Table 6

Regression of Relative Direct Cost and Drug Portfolio CER on WSJ and Product Withdrawal Dummy Variables

		Independent Variables			Summary Statistics	
Dependent Variable	Constant	Relative Direct Cost (%)	No WSJ Dummy	Withdrawal Dummy	Adj. R^2 F-value	Mean Dep. Var./ No. Obs.
Relative Direct Cost	0.631	-	-0.383	0.450	0.094	0.534
in %	(3.13)	-	(-1.37)	(1.18)	2.61	32
CER (-4, 5) of	-0.025	0.014	0.008	0.006	0.025	-0.013
Drug Portfolio	(-2.54)	(1.80)	(0.69)	(0.35)	1.27	32

exceed substantially those costs to firms which can plausibly be attributed to the recall of a specific drug.

Another way of putting this is that the stock market is imposing a substantial "goodwill" loss on a firm when a recall occurs that cannot be attributed to costs specific to the recalled product. The stock market appears to expect that news of a recall will reduce consumers' demand for other products sold by the firm and thereby impose additional losses on the firm. We get corroborating evidence for this conjecture when we add the market value of the firm to the regressions in Table 5. A single product typically accounts for a smaller fraction of a firm's profits the larger the firm, so the percentage loss due to recall of a single product should be smaller for larger firms if there is no spillover to other products. But the coefficient of the firm's market value is never as much as a tenth of its standard error, and this result implies that losses do spill over to the firm's other products. Indeed, we showed earlier that the losses may spill over to other firms in the same industry.

This exceptionally conservative (expected) response of consumers to news of a recall is something of a mystery, because there seems to be no easily apprehensible rational basis for expecting one product failure to beget others. As nearly as we can tell from our recall data, product failures occur randomly. However, whatever their source, it seems clear that the costs to drug firms of a recall are so large that they must exert a powerful deterrent effect on the production of defective products.

IV. Auto Recalls

Since the late 1960s, the National Highway Traffic Safety Administration of the Department of Transportation (DOT) has been empowered to order manufacturers to recall and repair autos with defects which compromise safety. Here we use a sample of 116 "major recalls" that occurred in 1967-81 to analyze the stock market's response to the news of this form of product defect. Our analysis here will have to be sensitive to a problem we raised in Section II: Recall announcements occur too frequently to be treated as entirely surprising to the stock market, so the market's response to the news of a recall can understate the full costs it imposes on producers of recalled cars. However, the problem is easier to state than solve, so we defer dealing with it until later in the section.

A. *Recall Sample*

Each recall is initiated by an order from DOT specifying which particular group of cars are to be recalled and what is to be done to fix the car. The distribution of the number of cars per recall is highly skewed. Some involve a few hundred cars or even less, and a few involve millions of cars. We wanted to avoid dealing with a lot of obviously trivial cases while retaining enough variety to analyze the effects of recall size. Accordingly, our sample is drawn from all recall announcements reported in the WSJ involving the domestic "Big 3" (GM, Ford and

Chrysler) for 1967-1981[6] which exceeded the following minimum size criteria: 50,000 cars for GM, 20,000 for Ford and 10,000 for Chrysler. These cutoffs are crudely consistent with the relative market shares (and stock market values) of these firms, and they result in roughly equal representation of each firm in our sample. The sample is described more precisely in Table 7. It is clear that, even after excising the small recalls, there is a very broad range of recalls in our sample, and that our sample remains highly skewed to the right; every relevant coefficient of variation comfortably exceeds one. GM has the biggest recalls, Chrysler the smallest, but these ranks are reversed when recalls are measured relative to market value.

B. *Stock Market Response to Recall Announcements*

For each of the 116 recalls in our sample we computed CERs for various periods around the event date—the date of the WSJ story about the recall. We used the same source (the Scholes excess return file from CRSP) and procedure as for drug recalls. The basic results are in Panel 1 of Table 8. We find significantly negative average CERs for every event window, and the average gets larger absolutely as the windows widens. We did not go beyond the two week window, CER(-5, 5), because recalls are so numerous that much wider windows would have created serious overlap problems.[7] That window yields a mean CER of -1.60 percent. About ½ this total is realized in the 3 days surrounding the event, another ⅓ in the subsequent 4 days (CER (2,5)) with the 1/5 or so remainder leaking out prior to the day before the event. Also, there is a significantly above average frequency of negative recalls for every window, though these do not begin to approach the near-unanimity in the corresponding data for drugs.

Panels A, B and C of Table 8 break out results by company. These reveal that every firm suffers a negative average CER and an above-average frequency of negative CERs for every event window. That unanimity tends to support a conclusion that recalls are costly, even though many of the individual statistics in panels A-C are not significant. The rather wide standard errors on some of these makes us cautious about pushing comparisons among firms too hard, but it appears that GM loses about ½ as much per recall as either of its competitors (based on CER (-5, 5)). The extent to which this is plausibly due to its smaller recalls (per dollar of market value) is discussed later.

1. Does the CER Understate the Cost of a Recall?

Our discussion in Section II implies that the CER for recall periods is an estimate of $-(1 - p) \cdot K$, where K = the cost of a recall to a company and p = probability of a recall. So one way to estimate K would be to estimate p directly

[6] We have no stock market data for foreign producers and American Motors has too few recalls to permit reliable comparisons with the others.

[7] As it is, 4 of our 116 cases overlap. We left the overlaps in our sample, but no result would change very much if the overlapping cases are deleted or if we had made, the same adjustments as for drug recall overlaps.

Table 7

MAJOR RECALLS, 1967-81 DESCRIPTIVE STATISTICS

Sample	# of Recalls	Cars per Recall (000's)				Cars per $Million of Market Value	
		Mean	S.D.	MIN.	MAX.	Mean	S.D.
1. All Recalls	116	717.8	1552.8	14	12000	158.8	342.7
A. GM	41	1244.7	2352.0	50	12000	70.5	141.0
B. Ford	44	567.2	859.8	50	2700	128.0	179.0
C. Chrysler	31	234.6	391.1	14	1300	320.0	582.0
2. All 1967-74 Recalls	53	612.7	1320.8	14	6700	72.0	132.0
3. All 1975-81 Recalls	63	806.1	1729.6	19.6	12000	231.9	437.4

Table 8

MEAN CER FOR AUTO STOCKS. VARIOUS INTERVALS AROUND DAY OF RECALL. 1967-81

	Event Window				
Sample (# of Recalls)	(-5, 5)	(-3, 3)	(-1, 1)	(-5, 1)	(2,5)
1. Total (116)					
- Mean	-1.60%	-.96%	-.81%	-1.07%	-.53%
- t	3.40	2.56	3.30	2.85	1.87
- % Negative	61.2*	62.1*	64.7*	62.1*	60.3*
A. General Motors (41)					
- Mean	-.97	-.80	-.88	-.48	-.49
- t	1.64	1.70	2.86	1.02	1.38
- % Negative	56.1	58.5	65.9*	56.1	56.1
B. Ford (44)					
- Mean	-2.03	-1.58	-.63	-1.51	-.52
- t	3.50	3.42	2.08	3.26	1.49
- % Negative	63.6	68.2*	61.4	63.6	54.5
C. Chrysler (31)					
- Mean	-1.83	-.28	-.98	-1.24	-.59
- t	1.37	.26	1.40	1.16	0.73
- % Negative	64.5	58.1	67.7*	67.7*	74.2*

See text for description of sample, and *see* note to Table 2 for method of computing t.
* = t > 2.0. (*See* note to Table 3).

and divide the CER by (1 - p). To see where such a procedure would lead, note that every company in our sample experienced an average of 2 to 3 major recalls per year in the 1967-81 period, or about 1 in every 10 two-week periods. If uncertainty is typically resolved in the two weeks surrounding a recall, a plausible estimate of p would be around .1, and, using this estimate, we could estimate an average loss due to a recall of around 1.8 percent of market value rather than the 1.6 percent in Table 8.

However, that procedure is based on an implicit assumption that needs to be examined with special care in the case of the auto industry in the 1967-81 period. This is that the other unexpected events impacting auto stocks in the two weeks surrounding a recall were not adverse or favorable on average. To elaborate: ***ex ante*** every CER is zero in expected value. If the ***only*** surprise is that a recall did or did not occur, the *ex post* CER is then -(1 - p)K or pK as in equations (4) and (5). If other surprises occur during recall periods, but are not systematically adverse or favorable, then the -1.6 percent mean CER in these periods is an ***unbiased estimate*** of -(1 - p)K and it therefore underestimates K. We can correct the underestimate as suggested above, or we could estimate pK directly from CERs in non-recall periods. Those CERs would be an unbiased estimate of pK if, again, other surprises in these non-recall periods are neither good nor bad on average.

We know, however, that this last supposition is false for the 1967-81 period. These were hardly the best of times for domestic auto producers. The adverse effects of growing foreign competition, pollution regulation, etc. dominated their stock market performance, and the average CER in non-recall periods was negative. This raises two problems: (1) the CER for non-recall periods is obviously not an unbiased estimate of pK, the capital ***gain*** due to the absence of a recall, (2) the CER for recall periods overstates -(1 - p)K *if* other surprises in these periods were also adverse on average. It helps to state both of these precisely by revising (4) and (5) to include "other" surprises as follows:

$$(4)' \qquad (S_1^R - S_0)_t = -(1 - p_t)K_t + X_t \, ,$$

$$(5)' \qquad (S_1^{NR} - S_0)_t = p_t K_t + Y_t \, ,$$

where

X_t = the gain or loss due to non-recall events, in recall period t, and

Y_t = gain or loss from non-recall events in non-recall period t. The t-subscript indicates that actual returns in any particular period—the $(S_1^i - S_0)$ variables—are generated by a process whose elements can vary from period to period. Our calculated CERs are *ex post* realizations of this process for some particular time interval. They include the realizations of X_t or Y_t, and these can be positive or negative on average. We know that the mean value of Y_t for 1967-81 is negative, because the mean value of $(S_1^{NR} - S_0)_t$ is negative for this period (*see* below). We also know from Table 8 that the mean value of $(S_1^R - S_0)_t$ is negative, but this is insufficient to tell us that the mean of X_r is negative. Therefore, either procedure suggested above for estimating the mean of K has potential pitfalls. If we estimate $\bar{p}_r$ directly and divide the mean return in recall periods by $(1 - \bar{p}_r)$, we get an estimate of: $-\bar{K}_r + \bar{X}_r/(1 - \bar{p}_r)$, which is unbiased only if $\bar{X}_r = 0$. If, on the other hand, we subtract the mean of (5)′ from the mean of (4)′ we obtain

$$-\bar{K}_r + (\bar{X}_r - \bar{Y}_r) ,$$

which is unbiased only if $\bar{X}_r = \bar{Y}_r$. If $\bar{X}_r$ = its *ex ante* value of zero, then this estimate understates $-\bar{K}$, since $\bar{Y}_r < 0$. In addition, this latter estimate is likely to be noisier than any simple transformation of (4)′, because of the variance added by the Y_r series.

Since our readers deserve more than a lecture in elementary statistics, Table 9 presents various estimates of (4)′ - (5)′ for the (-5, 5) event window. These are labeled "Adjusted Mean CER", and the adjustment is as follows: For each year we compute the mean CER(-5, 5) for every non-recall period[8] for each of the three firms. Then we subtract this year-and-company specific non-recall mean CER from the CER(-5, 5) for each recall experienced by the company in the same year. This adjusted mean CER is an unbiased estimate of $\bar{K}$ on the assumption that both recall and non-recall periods within a year share a common company-specific impact of non-recall news. For ease of comparison, we repeat the unadjusted CER(-5, 5) from Table 8, and we provide the added detail of a sub-period breakdown.

None of the results in Table 8 are much affected by our adjustment of the CERs. The adjusted-mean CER (-5, 5) in Panel 1 is a bit smaller than the unadjusted mean, but it remains significantly negative. The main innovation in Table 9 is in the sub-period data of Panels 2 and 3, not in how the CERs are calculated. These reveal a sharp difference in the impact of recalls between periods. The average recall costs less than 1 percent of market value before 1975 (for every firm) regardless of how the CER is measured and it costs more than 2 percent after 1975. This difference is mainly attributable to Ford and Chrysler whose average recall-period CERs in this post-1975 period range from around -2½ to -3½ percent. This post-1975 period is not only more costly per recall, but there are more of them compared to the previous period (63 versus 53). This combination is especially costly for Ford, which bore the brunt of the increased recall activity (27 versus 17 before 1975).

The substantial difference between the stock market response to pre and post 1975 recalls turns out to be more apparent than real. As we show subsequently (*see* Table 11), it is due mainly to the decline in the real value of auto stocks: equally costly recalls translate into a higher percentage loss of market value the lower the market value. In addition, recalls increased in size after 1975 (*see* Table 7).

[8] More precisely, we compute 11 • mean daily ER for non-recall periods.

Table 9

MEAN CERS FOR AUTO RECALLS, ADJUSTED FOR NON-RECALL NEWS, 1967-81

Sample and (Number of Recalls)	Adjusted Mean CER (-5, 5) Mean	Adjusted Mean CER (-5, 5) t	Unadjusted Mean CER (-5, 5) Mean	Unadjusted Mean CER (-5, 5) t
1. All 1967-81 (116)	-1.38%	2.83	-1.60%	3.40
A. General Motors (41)	-.60	0.91	-.97	1.64
B. Ford (44)	-2.05	3.51	-2.03	3.50
C. Chrysler (31)	-1.46	1.06	-1.83	1.37
2. All 1967-74 (53)	-0.60	0.77	-0.55	0.73
A. GM (18)	-0.56	0.57	-0.57	0.63
B. Ford (17)	-0.44	0.63	-0.44	0.60
C. Chrysler (18)	-0.81	0.40	-0.64	0.33
3. All 1975-81 (63)	-2.04	3.35	-2.48	4.19
A. GM (23)	-0.64	0.71	-1.28	1.45
B. Ford (27)	-3.07	3.88	-3.02	3.79
C. Chrysler (13)	-2.37	1.29	-3.48	1.99

See text for description of Adjusted Mean CER(-5,5). Unadjusted mean CER(-5, 5) is computed as in Table 8.

2. How Does the Cost of a Recall Vary with the Number of Cars?

Every recall announcement contains information on the number of cars involved, and we should expect news of big recalls to be more costly than news of smaller recalls. But we found it as difficult to verify this for autos as for drugs. As with drugs, we wanted to allow for a "goodwill" effect of recalls, which we assume is proportional to the firm's market value. So we want an estimate of

$$\$ \text{ RECALL LOSS} = A \cdot (\text{MKT. VALUE}) + B(\text{CARS})$$

across recalls. Here A = "goodwill" cost and B = cost per car. Since

$$\$ \text{ LOSS} = \% \text{ LOSS} \times \text{MKT. VALUE},$$

we estimated

$$\% \text{ RECALL LOSS} = A + B(\text{CARS/VALUE})$$

using - CER (-5,5) as an estimate of % LOSS. A sample of the uniformly disappointing results is in Table 10. The regression estimate of B is rarely more than a standard error from zero and it is negative at least as often as it is positive. Nothing much is gained by allowing for inflation (*i.e.*, assuming that B is proportional to the GNP deflator; *see* lines 1(a)-3(a)) or changing the event window to (-1,1) in order to reduce noise. As with drugs, the A term (not explicitly shown in the table) accounts for essentially all of the recall cost.

One reason for this is that costs per car vary across recalls, so that our regression model is too crude. There is no handy way to estimate independently the dollar cost per car or per recall, as we could with drugs. But, from fragmentary press accounts, we know that the firms' estimates of their explicit costs per car repaired range from something like $10 to $1000, so the measurement error entailed by assuming that B is constant is considerable. There is the additional complication that car owners frequently do not respond to recall notices, and the response rate varies considerably across recalls.[9] We tried some crude adjustments for differences in repair costs and response rates for a sample of recalls, but this failed to sharpen the results in Table 10.[10]

---

[9] The law gives the owner the right, but no obligation, to have his car fixed at zero direct cost. Actual response rates vary over a range from about 1/3 to nearly 100 percent.

[10] Each recall order is published in DOT's annual *Safety Related Recall Campaigns for Motor Vehicles and Motor Vehicle Equipment Including Tires*. Sometimes the order requires repair only if inspection reveals a defect, and other times repair or replacement is mandatory. We allowed B to depend on a dummy = +1 if repair was required. The publication also gives response rates for some recalls, and we allowed B to vary with these. But neither variable worked. Nor did a variety of other adjustments which we tried—*e.g.*, modeling the goodwill loss as a constant dollar, rather than constant percentage, amount.

Table 10

REGRESSION ESTIMATES OF THE MARGINAL COST OF A RECALLED AUTO, BY FIRM AND PERIOD

Sample (# of Recalls)	Cost/Car	t
1. All 1967-81 (116)	-$2.62	0.2
(a) Price Deflated (1972 = 1.00)	-$1.21	0.1
(b) Based on CER (-1, 1)	-8.73	1.3
2. All 1967-74 (53)	-4.98	0.1
(a) Price Deflated	-9.09	0.2
(b) Based on CER (-1, 1)	19.26	0.7
3. All 1975-81 (63)	-10.09	0.7
(a) Price Deflated	-6.63	0.8
(b) Based on CER (-1, 1)	5.83	0.9
A. GM (41)	33.95	0.8
B. Ford (44)	-0.94	0.0
C. Chrysler (31)	-8.80	0.4

Note: Each Cost/Car is the estimate of B in a regression of the general form:

$$-CER(-t, t) = A + B\left(\frac{CARS}{VALUE}\right)$$

where CARS = # of cars recalled, and VALUE = market value of the firm. Negative values of B imply that the marginal cost of a car is negative! t = ratio of B to its standard error. For lines 1(a), 2(a), 3(a) $\frac{CARS}{VALUE}$ is multiplied by the GNP deflator (1972 = 1.0) to account for the effects of inflation on recall costs. CER (-5, 5) is the dependent variable in all regressions, except 1(b), 2(b) and 3(b) which use CER (-1, 1).

We can get some insight into the likely importance of measurement error by comparing the data on cars/market value in Table 7 with subsequent results. For example, we have already noted that GM has the lowest stock market loss per recall and the lowest cars/value. Note also that both cars/value and the mean stock market loss is much larger in 1975-81 than in 1967-74. All of this is consistent with a positive relationship between the market loss and recall size which we suspect is being hidden by measurement error in the disaggregated data summarized in Table 10. But we also have to note that Chrysler has a much higher cars/value than Ford, but no higher mean stock market loss.

Another approach to separating crudely the "goodwill" from direct cost elements of the stock market loss is to see whether it is plausible to assume no goodwill loss at all. For drug recalls, we found that such an assumption would appear utterly implausible because the stock market losses are very much larger than the direct costs. As a first step in such a comparison for autos we have converted the percentage losses (using the unadjusted CER (-5,5)) to (constant 1981) dollar losses per recall and per car. The first column of Table 11 shows that each recall cost around $140 million in market value on average, with GM bearing the highest mean dollar cost and Chrysler the lowest. It also shows that the difference between the pre and post 1975 dollar losses (*see* lines 2 and 3) is much smaller than the difference in the corresponding percentage losses in Table 9. So, with the caution implied by the large standard errors, we can attribute much of that percentage difference to the combined effects of inflation and the poor stock market performance of auto stocks in the late 1970s. Were it not for those twin events, the costs of recalls could well have been lost in the noise of stock prices.

Table 11 also expresses these dollar losses on a per car basis. While we report a mean loss/car, we have little confidence that the high dollar amounts are meaningful. These means are dominated by a few extremely small recalls that generate extremely large losses/car. Accordingly, we show two other measures less affected by these extreme values—the median and the mean dollar loss/mean number of cars in a recall (labeled mean/mean).[11] This last datum is equivalent to aggregate losses in a sample divided by aggregate cars, so it comes closest to summarizing the experience of these firms over long periods. What is perhaps most interesting about this figure is its stability over time and between companies: in any large sample of recalls, the loss/car seems to be around $200.

This last result suggests that—with sufficiently large samples to iron out the random fluctuations—recall costs are proportional to cars recalled. But it appears implausible that the $200 figure is entirely attributable to the direct costs of a recall. Since such costs are a deductible expense, it would imply pre-tax costs of nearly $400 per car, and this would in the high end of the range of per

[11] To illustrate the problem entailed by very small recalls, remember that losses and cars are essentially uncorrelated. So, suppose losses in a 1 car recall and a 100 car recall are each $100. The mean loss/car = ½(100/1 + 100/100) = $50.50, but the total loss from both recalls is only *200* = $1.98 per car. This last figure is our mean/mean for this sample.

Table 11

ESTIMATED DOLLAR LOSSES PER RECALL AND PER CAR, 1967-81, CONSTANT 1981 DOLLARS

	Loss/Recall (Million $)		Loss/Car ($)			
Sample (# of Recalls)	Mean	t	Mean	t	Median	Mean/Mean
1. All 1967-81 (116)	141.1	2.2	813.3	1.5	185.7	196.6
A. GM (41)	235.5	1.4	477.5	0.6	46.6	189.2
B. Ford (44)	128.6	3.7	694.2	2.3	198.2	226.7
C. Chrysler (31)	34.0	1.0	1426.4	0.9	95.7	144.9
2. All 1967-74 (53)	110.1	0.9	1092.9	1.0	64.7	179.7
3. All 1975-81 (63)	167.2	2.7	578.1	2.0	189.0	207.4

Note: Loss/Recall is estimated by multiplying CER (-5, 5) for each recall period by the market value of the firm in that period. Mean Loss/Car is the mean of [(Loss/Recall)/Cars involved in the recall]. The last column (Mean/Mean) is obtained by dividing the mean of Loss/recall, as shown in the first column, by the mean of cars/recall from Table 7. Each loss/recall is deflated by the GNP deflator set to a base of 1981 = 1.0.

car costs which have appeared in press reports about specific recalls.[12] We know of only one publicly available piece of data which permits an estimate the per car cost in a large sample of recalls: GM disclosed that it spent $33 million on recalls in 1982 (Detroit *Free Press*, May 22, 1983). This amounts to about $35 per GM car recalled that year. If this is anywhere close to being typical, then the bulk of the stock market loss represents indirect costs: lost sales and goodwill, liability suits, etc.[13,14] In this sense, there is a rough similarity between auto and drug recalls. But there seems to be a closer connection between the size of the indirect and direct costs for auto recalls than for drug recalls.

3. How Are Competitors Affected by Auto Recalls?

The surprising result that competitors lose rather than gain during a drug recall holds for auto recalls as well. And, as with drugs, the surprise is deepened by the large magnitudes involved. The data are summarized in Tables 12 and 13. Table 12 shows mean CERs to equal weighted "portfolios" of the two competitors during recall periods (*e.g.*, a Chrysler-Ford portfolio during GM recalls). The main result is that competitors lose about 1 percent on average during a two-week recall period, or about ⅔ as much as the recall company loses. All of this is attributable to 1975-81 recalls, where the competitors' loss (-2.40%) virtually matches the recall company's loss! This difference between sub-periods is less intelligible than the similar sort of difference we found for recall companys CERs (*see* Table 9). In that case, we saw that the apparently weak negative CERs for 1967-74 were plausibly masking negative real dollar impacts roughly comparable to those in the later period. In Table 12 we find similarly weak, but positive CERs for competitors in 1967-74. These would be consistent with nontrivial real dollar gains to competitors, a result which would excite no surprise.

---

[12] For example, a recent Detroit *Free Press* series on recalls states that a 1983 recall of 240,000 GM cars "is thought to be the most expensive per-car recall ever." GM's estimate of its total direct cost for the recall is $30 million, or $125 per car. Detroit *Free Press* May 24, 1983.

[13] In this connection, we note recent evidence that new car sales of recalled models appear to decline when the recall is announced. *See* S. M. Crafton, G. E. Hoffer and R. J. Reilly "Testing the Impact of Recalls on the Demand for Automobiles" *Economic Inquiry* v. XIX, Oct. 1981, 694-703 and R. J. Reilly and G. E. Hoffer "Will Retarding the Information Flow on Automobile Recalls Affect Consumer Demand?" *Economic Inquiry* v. XXI July, 1983, 444-447. The latter article estimates that sales of a domestic recalled "line" decline about 5 percent in the month of a recall announcement in the 1977-81 period, but there is no indication that the decline lasts more than a month.

A single month sales decline of this magnitude could not account for very much of the typical stock market loss. There are about 60 domestic "lines" with average monthly sales of around 10,000 cars. Reilly and Hoffer exclude lines with fewer than 8,000 cars per month. If the average line in the sample has 20,000 monthly sales, a 5 percent decline represents 1,000 cars or roughly $10 million sales. The lost pre-tax profits on these sales would amount to under $2 million, based on the industry's margin of sales over material and labor costs.

[14] We also found no serial correlation in recalls. For example the correlation of the number of recalls in successive three-month periods is .03 for Chrysler, .08 for Ford and GM and .17 for the aggregate of all three firms. None of these are significant; auto recalls, like drug recalls, appear to occur randomly.

Table 12

MEAN CERS OF COMPETITORS DURING RECALL PERIODS, VARIOUS EVENT WINDOWS, 1967-1981

	Event Window				
Sample (# of Recalls)	-5, 5	-3, 3	-1, 1	-5, 1	2, 5
1. All 1967-81 (116)					
- Mean	-1.01%	-.92%	-.35%	-.49%	-.52%
- t	2.46	2.76	1.55	1.43	1.94
- % Negative	59.5*	56.0	57.8	62.1*	52.6
2. 1967-74 (53)					
- Mean	.64	.17	.18	.57	.08
- t	1.09	0.42	0.62	1.07	0.22
- % Negative	41.5	37.7	45.3	45.3	45.3
3. 1975-81 (63)					
- Mean	-2.40	-1.84	-.79	-1.37	-1.03
- t	4.65	3.81	2.41	3.30	2.70
- % Negative	74.6*	71.4*	68.3*	76.2*	58.7

Note: Each entry shows the Mean CER (-x, y) for an equal-weighted portfolio of competitors during recall periods. So, *e.g.*, for a GM recall; our "portfolio" is Chrysler and Ford, and our CER (-x, y) for that recall is just the average of the Chrysler and Ford CERs. The ratio of the Mean CER to the sample standard error is shown below the mean.

"% Negative" is the percentage of negative CERs to the portfolio of competitors.

* = t > 2.0.

But the post '74 data clearly describe a much different world. Table 13 organizes the data by company. Reading across any of the rows in panels A-C gives the response of a particular competitor to recalls of its rivals. For example, we find from panel A that GM lost on average a) 2.13% during Ford recalls, b) .83% during Chrysler recalls, c) 1.59% during all Ford and Chrysler recalls, etc. Panel D shows the average response of both rivals to a particular company's recalls. So, *e.g.*, panel D.1. says that Ford and Chrysler lost 1.03% on average during all GM recalls in 1967-81. With due respect given to the large standard errors, this detail reveals some interesting patterns: a) GM loses more during its rivals' recalls (1.59%) than it does during its own recalls (0.97%); the reverse is true for both Chrysler and Ford. The GM response to rivals recalls in 1975-81 (-2.92%) is particularly noteworthy: it loses even more than they do. b) The most damaging recalls for competitors are GM and Ford recalls, particularly in 1975-81. (-2.61% and -2.77% respectively). GM's rivals lose more than GM does in this period. c) By contrast, the relatively small Chrysler recalls cost rivals about half as much as GM and Ford recalls (*see* D.3) and only about ⅓ what they cost Chrysler itself in this 1975-81 period. So Chrysler recalls seems to be treated mainly as idiosyncracies without strong implications for industry wealth.

Since excess returns to auto stocks were negative on average over our sample period, and especially so in 1975-81, the reader may wonder if the negative industry effects which we attribute to recalls have a more general source. This seems implausible. From Table 9, Panel 3 we find that the mean company CER in nonrecall periods for 1975-81 was about -½ percent, or under ¼ the mean CER for competitors during recall periods.[15] So other news would have to be consistently ***especially*** bad during recall periods for a general "bad news" explanation to make sense. This stretches credulity, and we can easily reject a hypothesis of no difference between the mean CER in non-recall periods and the mean CER to competitors in recall periods.

The more intriguing question raised by Table 12 is whether there is any company-specific component at all in the loss due to recalls, given that the mean company-specific and competitor CERs are so similar. Table 13 gives us a hint: there is no general tendency for companies to respond identically to their own recalls and those of competitors. A more formal answer is given by regressing a recall company's CER (-5, 5) on that of the portfolio of its competitors during recall periods. The regression coefficient here gives the company's average share in any industry-wide effects of recalls, and the intercept gives the average company-specific component. We computed the regression for each of the three companies, and obtained a mean intercept of -1.12 percent (t = 2.98).[16] So there

---

[15] Note that the -½ percent ***includes*** the negative returns to competitors during recall periods, so it is too large an estimate of the impact of non-recall news.

[16] The average regression coefficient is .69(t = 7.46). This implies that the typical company-share in an industry-wide recall loss of 1 percent is under 1 percent. The proximate reason for (footnote continued)

Table 13

MEAN CER (-5, 5) OF COMPETITORS DURING RECALL PERIODS, BY COMPANY AND SUB-PERIOD, 1967-81.

Competitor Company	Recall Company			All Recalls of Other Cos. In		
	GM	Ford	Chrysler	'67-'81	'67-'74	'75-'81
A. GM						
- Mean		2.13%	-.83%	-1.59%	-.09%	-2.92%
- t		4.02	1.04	3.49	0.13	5.34
- % Negative		75.0*	54.8	66.7*	45.7	85.0*
- N		44	31	75	35	40
B. Ford						
- Mean	-1.11		.36	-.48	.86	-1.81
- t	1.34		0.47	0.83	1.11	2.24
- % Negative	53.7		54.8	54.2	50.0	58.3
- N	41		31	72	36	36
C. Chrysler						
- Mean	-.94	-.94		-.94	1.16	-2.41
- t	0.83	0.95		1.27	1.12	2.44
- % Negative	53.7	54.6		54.1	37.1	66.0*
- N	41	44		85	35	50

D. All Competitors				
1. 1967-1981				(See Table 12 for aggregation across all recalls.)
- Mean	-1.03	-1.54	-.23	
- t	1.24	2.64	383.2	
- % Negative	56.1	63.6	58.1	
2. 1967-1974				
- Mean	1.00	.42	.50	
- t	0.87	0.45	0.50	
- % Negative	38.9	41.2	44.4	
- N	18	17	18	
3. 1975-81				
- Mean	-2.61	-2.77	-1.25	
- t	2.39	4.22	1.47	
- % Negative	69.6*	77.8*	76.9*	
- N	23	27	13	

Note: Each entry under "Recall Company" is based on the CER (-5, 5) of competitors of that company during that company's recall periods. For example data in Panel A. refer to GM's CER (-5, 5) during Ford and/or Chrysler recalls. The -2.13% under Ford means that GM's Mean CER (-5, 5) was -2.13% during the 44 Ford recall periods, etc. The last 3 columns give data for a specific competitor (*e.g.*, GM in Panel A) aggregated across all recalls of other companies (Ford and Chrysler in Panel A.) for specific time periods.

Panel D. shows the aggregated response of competitors to particular company recalls, again for specified time periods. For example, the -1.03% entry under GM in Panel D.1. is the mean CER (-5, 5) of GM's competitors (Chrysler and Ford) during the 41 GM recalls in 1967-81.

* *See* note to Table 12 for definition of "Mean", "t", and "% Negative". N = number of recalls in the cell. For Panel D., t and % negative are based on CER (-5, 5) of equal weighted portfolio of competitors (*see* note to Table 12).

is a significant company-specific component to recall losses over and above a company's share in the industry-wide loss.

V. Summary

There are striking similarities in the effects of drug and auto recalls on the wealth of shareholders. Both are extremely costly, surely more costly than the direct costs of recalling the defective product and, plausibly, more costly than all of the costs attributable to the specific product defect. In both types of recalls a more general loss of "goodwill" seems to be a large component of the total loss. This result is not unique to this study. One of us has found similarly large "goodwill" losses for FTC false-advertising cases.[17] Just what lies behind these goodwill losses remains something of a mystery which we leave for future research. Our attempts, mainly with drugs, to find answers in costs of product liability suits and in time dependence of recalls succeeded only in deepening the mystery.

Another similarity between drugs and auto recalls—and the source of another mystery—lies in the response of competitors. Their owners lose substantially when a rival product is recalled. Any favorable effects on the demand for substitutes from a recall are swamped by a more general negative effect on the industry. This is another piece of evidence that something much more is involved in a recall than failure of a specific product.

It is difficult to compare the magnitudes of the losses in drug and auto recalls, because both the frequency of recalls and the number of firms involved differ. Per recall, the percentage loss is much greater for drug recalls (6% versus 1½%). But auto recalls occur over twice as frequently, and involve only 3 companies versus 19 for drugs (in our samples). So per-company per-year, auto recalls are considerably more costly. The average loss to rivals is roughly the same (1 percent) for auto and drug recalls, but with about 50 rivals in the case of drugs versus 2 for autos, it is clearly the drug recalls which have the more substantial cross-firm effects.

We began by promising to shed light on the degree to which the capital market might sub-optimally deter production of faulty products. We believe we have done so. It is clear that, in the simple sense of the market's not internalizing even the direct costs, suboptimal deterrence is no problem. It is also clear that to make a suboptimal deterrence story credible requires very generous estimates of the indirect social costs. The only source of such large costs we have found is in the cross-company effects. This might suggest that there is a larger scope for industry cooperation in product design and inspection than economists have

(continued from p. 405.)
this is that Chrysler has more volatile returns than the others, so when Chrysler loses 1 percent the others lose less.

[17] Peltzman, "The Effects of FTC Advertising Regulation," *J. of Law and Econ.*, v. 24, December, 1981, pp. 403-447.

heretofore imagined.

Finally, we hope that our results have begun defining an agenda for future research. They suggest that recall costs are like an iceberg whose easily visible part hides most of what is important. The challenge for future research is to discover just what form—*e.g.*, reduced sales, increased quality costs, lost "political capital"—these large, currently amorphous costs take.

Comments

Gerard R. Butters
Federal Trade Commission

Peltzman and Jarrell's paper presents new and provocative findings. We learn that a major recall of a drug imposes, for each \$1 of estimated direct costs to the manufacturer of the drug, an immediate cost of about \$12 on the stockholders and an additional cost of perhaps \$25-\$50 on the stockholders of other drug manufacturers.[1] A recall of domestic automobiles with a direct cost of perhaps \$50-\$125 per car imposes a cost of \$200 per car on the manufacturer and an additional \$150 or so on each of the other two of the big three auto firms. These losses are large in absolute terms: the median estimated direct cost of a drug recall was about \$2 million, resulting in an estimated industry stock market loss of almost \$100 million. In the case of Johnson and Johnson's recall of Tylenol, the firm's own stock was devalued by over \$1 billion.

Although these losses are not known with much precision, they are not statistical artifacts: the measures of the stock market reaction are robust and are based upon accepted statistical methods. Nor are the losses momentary: they are not reversed by a later compensatory increase in stock values, at least not for two months after the initial event, which is about as far as statistical methods allow one to test. We cannot conclude that the effects are permanent, but if they were, the grand totals would be staggering: ignoring timing considerations, 32 drug recalls which each reduce the industry stock market value by over 1% would cost investors who held the stocks over the period of the recalls over 27.5% of their initial investment.

How can we explain why the market responds so much more dramatically to product recalls than would appear to be warranted by the direct costs estimated by Peltzman and Jarrell? One possibility is that they do not estimate the value of the lost sales during the time that the recalled product was taken off the market. For those products that were totally or permanently withdrawn, the lost sales were huge.[2] They also omit the cost of liability for deaths and damage to health resulting from the use of the withdrawn drug. The prospect of large punitive damages for drug-related deaths, such as the \$6 million recently

[1] This estimate is based upon the assumption that the firms whose products are recalled have capital values two to four times larger than the average capital value of all firms in the 50-firm index used by Peltzman and Jarrell.

[2] I estimate that the withdrawal of Oraflex cost Lilly about \$150,000,000 in potential annual sales, the withdrawal of Rely cost Procter and Gamble roughly \$100,000,000 in potential annual sales, and the recall of Tylenol put the entire \$450,000,000 annual sales of Tylenol products at risk. However, even these huge losses do not seem to account for the stock market losses of approximately \$500,000,000, \$300,000,000, and \$1.2 billion respectively, especially considering that the recalls of Oraflex and Rely were presaged by previous bad publicity and attendant lost sales.

awarded in a case involving Oraflex, could make a substantial contribution to the stock market reaction. Perhaps the best way to measure contemporaneous estimates of the expected magnitude of these losses is to study liability insurance rates and coverage before and after major recalls. However, I have not been able to convince myself that either lost sales or liability payments and increased insurance costs are large enough to account for the capital market reaction, especially for those recalls that were not publicized or did not involve life-threatening defects.

Another possibility is that while a recall does not itself reduce the future earnings of the firms by enough to account for the stock market changes, it indirectly provides information to investors about the firm whose product was recalled, and more broadly about the entire industry, that leads them to revise downwards their estimates of future earnings. For example, the recalls may act as a signal of unrecognized poor product quality control in the firm, of unanticipated difficulties in the production of new drug entities, or of unexpectedly strict government regulation. But even dramatic news should not affect the market valuation of a firm except to the effect that it overturns existing beliefs about product quality or regulatory policy. In the case of automobiles, for example, product quality problems "signaled" by a product recall might well have been previously anticipated as a result of bad *Consumer Reports* ratings, reliability problems of related earlier models, an observable lack of good quality control practices at the factory, previous recalls, etc. In the absense of supporting evidence that recalls are a particularly good "leading indicator" of future quality problems, and given the trivial nature of many of the defects for which cars were recalled, I doubt that signaling can account for much of the stock market effect.

If it is difficult to explain the movement of the stock of the firm whose product was recalled, it is nigh impossible to explain how the capital value of the rest of the industry could fall by even larger amounts (in total dollars). While one might plausibly argue that reputation effects could shift demand away from one firm to the rest of the industry, it is hard to believe that recalls could dramatically reduce overall industry demand. One is forced to the take seriously the possibility that although the losses from recalls are indeed substantial, the stock market overreacts to the prospect of these losses.

This conclusion would be unreasonable if there were an easy way for investors who understood an overreaction to profit from it. But it is quite possible for the market to overreact significantly without providing savvy investors with dramatically increased rates of return. Since future earnings of firms are notoriously difficult to predict, there is a wide range of uncertainty as to the true value of any firm. Thus, if the stock value of a firm falls "too far" in response to a recall, there is no reason why the market price need recover its value rapidly. If the stock price tends to recover only gradually over a period of months or years, then any investor seeking to profit from the overreaction can anticipate only a modestly increased expected return, and is subject to the risk that intervening events will cause the market price to fall rather than rise. When one accounts for the costs of identifying the types of news which the market tends to misjudge and the transactions costs of buying and selling, it would not be surprising to find that arbitrage need not prevent the overreactions.

If it is true that the stock market is overreacting to the news of drug and automobile recalls, then we must question the results of event studies more broadly. In particular, I do not think that one can easily attribute economic significance to small percentage moves in stock market values, even if the changes are both statistically significant and large in absolute dollar terms. Thus, the main contribution of Peltzman and Jarrell's study may be to improve our understanding of the stock market rather than the markets for drugs and automobiles.

Peltzman and Jarrell clearly intend the reverse. In particular, they seek to measure how much deterrence the capital market provides against the sale of faulty products. They conclude that "the costs to drug firms of a recall are so large that they must exert a powerful deterrent effect on the production of defective products." Judging from their overall discussion, they do not mean primarily that the capital market reaction in itself motivates firm executives to be concerned about product quality, but rather that the capital market reflects, at least in a roughly accurate way, real costs to the firms whose products are recalled.

I have little doubt that Peltzman and Jarrell's final conclusion is accurate. Even if the market overreacts to news of a recall, only a fraction of the implied costs to drug manufacturers should be sufficient to encourage stringent quality control. But one must be careful not to generalize beyond their carefully limited statement. First, they make no attempt to measure the external costs of product defects and to compare them to the private costs. Many private costs of a recall (*e.g.*, expected liability suits and losses in sales) have their corresponding social costs (health damages or death and losses in therapeutic benefits). Thus to the extent that the stock market evidence raises our estimates of the private costs of recalls, it should also cause us to raise our estimates of the social costs. Hence it is not clear that the stock market evidence should change one's prior beliefs concerning the likelihood that deterrence of product defects is either sub-optimal or supra-optimal. Second, although the point is obvious, to show that *product recalls* are costly is not the same thing as to show that *product defects* are costly. In the extreme, if recalls were based upon arbitrary and unpredictable regulatory whims, product recalls would not be effective in stimulating improved quality control. Third, it is only in the context of current FDA regulation that recalls (and indirectly, poor product quality) are shown to be costly. One cannot conclude that the response to recalls or disclosures of defects would be the same in a world without direct regulation of quality, especially if one believes that the current market response is being driven by fears of tightening FDA or NHTSA regulation.

Consumer Shopping Behavior In the Retail Coffee Market

Michael L. Katz and *Carl Shapiro**
Princeton University

1. Introduction

In most markets, consumers must choose from among a variety of brands and retail outlets. Price and quality information is costly to acquire, and consumers must engage in search to obtain it. Moreover, even when information about the standard retail prices is known, the consumer may be able to obtain further price reductions through special deals or bargaining. Search and deal-seeking are two elements of "shopping." Given the variety of information sources and deals available in some markets, the problem of determining a fully optimal shopping strategy can be an extremely complex one. As a result of this complexity, the theoretical literature on shopping, such as that dealing with optimal search strategies, has largely bogged down in the mathematical difficulties of computing the optimum.

In this paper, we present an empirical examination of the ways in which consumers and firms respond to the costliness of information and to the possibility of special deals that require consumer action (*e.g.*, redeeming coupons). We use an exceptionally detailed data set to examine consumer shopping behavior in the retail market for ground, caffeinated coffee.[1] The detailed nature of our data allows us to capture much of the richness of consumer shopping behavior, and to identify some of the implications that this shopping behavior has for the demand functions faced by sellers. We are able to use the data to develop econometric estimates of the demand functions for coffee, and to test whether these demand functions conform with various theories of consumer shopping behavior.

The retail coffee market offers an excellent chance to analyze the effects of a variety of marketing practices, and to examine consumer switching among brands. Consumers make frequent purchases of coffee, and many types of promotional activities are used by coffee producers. Promotions include coupons issued by both the manufacturer and the store, in-store promotions (such as end-of-aisle displays), price advertising in local print media (newspapers adver-

---

* We would like to thank John Hilke, David Sappington, and Richard Quandt for helpful comments and suggestions. We are grateful to Information Resources, Inc. for making the data available to us.

[1] Throughout the paper, we use the word "coffee" to refer to ground, caffeinated coffee unless otherwise stated. We do not, in general, include sales of soluble (instant) and decaffeinated coffees in our analysis. The sales of ground, caffeinated coffee make up 43 percent (in terms of the number of cups of coffee that can be made from the purchased quantities) of the total sales of all types of coffee in our sample (including instant and decaffeinated).

tisements, newspaper inserts, and flyers), and nonprice advertising on a spot and network basis on television. Trial sizes also are used. New brands are introduced.

There are two broad sets of questions that we address. The first set of questions concerns consumer shopping behavior and its relationship to manufacturer and retailer promotional activity. We are able to report in detail the pattern of consumer shopping behavior and the pattern of manufacturer and retailer promotional activity. Switching among brands is one measure of consumer shopping behavior that is of particular interest. We construct brand switching matrices for all purchases, as well as for purchases made using coupons or while the brand chosen was advertised.

The second set of questions that we address concerns the existence of market power in the coffee market. The nature of consumer shopping behavior has important implications for market performance. Consumers may respond to price changes much differently in markets where information is costly than in markets where information is costless and complete. Moreover, when shopping is important, seller promotional activities provide several new dimensions to firms' competitive strategies. Our examination of consumer shopping behavior provides some evidence concerning the question of whether promotional activities are likely to promote or stifle interbrand competition. We also test more directly for the existence of retailer or manufacturer market power by estimating aggregate demand equations. We are able to improve upon the usual demand specification by including in these equations important nonprice explanatory variables, such as the use of local print advertising. Our results indicate that estimates based on price variables alone may be misleading when consumer shopping and retailer promotional activities are prevalent.

The remainder of the paper is organized as follows. A general theory of consumer shopping behavior is sketched in Section 2. Section 3 is devoted to providing a fairly detailed description of actual consumer shopping behavior in the retail market for ground, caffeinated coffee. The resultant demand functions for coffee are estimated in Section 4. The paper closes with a short conclusion.

2. A Simple Model of Consumer Shopping Behavior

A. *The Elements of Shopping*

Roughly speaking, we can classify the components of shopping behavior into two categories: (1) obtaining information about the prices and qualities of the available products; and (2) obtaining any sorts of deals that require special action on the part of the consumer, such as clipping coupons or bargaining. Having engaged in shopping, a consumer then makes the actual purchase decisions.

The utility that a household derives from the purchase and consumption of coffee will depend on the attributes (*e.g.*, taste or image) of the brand chosen, the price paid for the brand, and the level of shopping effort. The greater the amount of effort that a consumer devotes to gathering information or to obtaining deals, the more likely it is that he or she will find a desirable brand or a low price. The theory of optimal search or shopping by consumers who have incomplete information about both prices and qualities is extremely complex, and there is no compelling reason to believe that consumers intuitively can apply information gathering and processing algorithms that trained economists and statisticians are unable to develop for any but the simplest search problems.[2] However, it is not an unreasonable hypothesis that in choosing the level of shopping in which to engage, rational consumers will attempt to balance the benefits of shopping against the costs. These costs may include travel costs, the purchase of information, or the opportunity cost of the time spent shopping. In light of these costs, consumers typically will not find it in their interests to become completely informed or to utilize all possible deals.

B. *Information Gathering*

A consumer in the coffee market has a variety of sources from which to construct an information portfolio. In making their information acquisition choices, consumers will balance the credibility and costliness of alternative information sources.[3] Some sources, such as visits to several stores, will provide particularly reliable or accurate information, but may be too costly for the consumer to rely on extensively. Rational consumers may turn to sources that are less reliable or less accurate, but that are less costly as well.

Potential information sources in the retail coffee market include the following ones.

1. Advertising

Coffee manufacturers and retailers made use of four types of advertising during our sample period. First, retailers used *features*, wherein particular brands of coffee are promoted in local fliers, newspaper ads, or newspaper inserts. A featured brand is named in the advertisement, and its price is given. Features were used for 11 percent of the brand-chain weeks included in our sample.[4] Table 1 presents the frequency of feature usage broken down for each of the four leading brands of ground, caffeinated coffee.

Retailers also made use of in-store *displays*. Such displays typically consist of stacking cans or bottles at the end of an aisle, where they will attract passing consumers' attention. Overall, displays were used for 12 percent of the brand-chain weeks. Again, Table 1 gives the breakdown by brand.

Coffee manufacturers made use of advertising on television. Both spot and national television purchases were made. In 1978, manufacturers spent roughly

---

[2] The natural selection arguments applied to competitive firms groping for optimal behavior are inappropriate for consumers.

[3] Porter (1976) gives a much more thorough presentation of a theory of consumer information portfolio choice.

[4] If k chains feature a given brand during a particular week, then the brand is said to be featured for k chain-weeks. We take the average over the entire sample period

90 million dollars nationwide on television advertisements, about half of which went for spot advertising.[5] We do not have data on television expenditures broken out solely for the area (Pittsfield, Massachusetts) covered by our sample of consumers.

Finally, coffee manufacturers made use of nonprice, printed advertising. In 1978, manufacturers spent roughly 10 million dollars nationwide on printed advertising.[6] Again, we do not have data on these expenditures broken out solely for Pittsfield, Massachusetts. These expenditures may include the costs of distributing manufacturers' coupons through newspaper inserts and other print media.

2. Experience

Coffee is like many nondurable consumer items in that it is purchased relatively frequently by most households. Frequent purchases imply that consumers are likely to learn a given brand's attributes rapidly as it is possible to sample several brands in a relatively short period of time. Table 2 reports the frequencies of purchase of ground, caffeinated coffee across households. The mean number of cans purchased over the 108 week period is 30.5.[7] If purchases were evenly spaced, this purchase number would correspond to one purchase every 25 days. Most households' useage rates permit them to sample several brands within our two-year time frame.

3. Testimony of Other Consumers

Word-of-mouth is an important source of information in many markets. Coffee is a candidate for this form of information transmission for two reasons. First, the information to be transmitted is nontechnical—one does not have to be an expert to determine whether a particular brand of coffee tastes good. Second, virtually all coffee consumers know other coffee consumers. A study conducted in 1978 for the International Coffee Organization found that in the Northeastern census region of United States 65 percent of the people who are over 10 years of age drink coffee (including instants and decaffeinated ground) on any given day.[8]

4. Coupons

Both coffee producers and retailers issue coupons. Manufacturer coupons tend to be valid for greater periods of time than do store coupons. One of the three chains of supermarkets in our sample offered a fair number of retailer coupons, but stores in the other two chains made very infrequent use of retailer coupons. The average value of a coupon used for a coffee purchase (including instant and decaffeinated) was 60.3 cents. Coupons are like advertising in many ways, but

[5] *The U.S. Retail Coffee Market (A)* (1982) Harvard Business School, page 14.

[6] *The U.S. Retail Coffee Market (A)* (1982) Harvard Business School, page 14.

[7] For all types of coffee (including soluble and decaffeinated) the corresponding figure was 50.2.

[8] *The U.S. Retail Coffee Market (A)* (1982) Harvard Business School, page 5.

with an obvious difference. As we shall discuss, coupons may play a role in the provision of both price and quality information.

C. *Price and Quality Information*

Consumers need both price and quality information. Some sources are well suited for providing price information, but poor providers of quality information. For example, newspaper advertising conveys accurate price information, but has little information about product quality. Thus, although there is some overlap in the sources of the two types of information, it is useful to consider consumer information portfolio choices separately for price search and quality search.

1. Price Information

Consumers can obtain price information either through in-store search (direct observation of shelf prices) or out-of-store search (media price advertising and word-of-mouth). With the possible exception of word-of-mouth, all of these sources of information are highly reliable. They also are costly, but to varying degrees. The costs of in-store search are the *incremental* costs of going to the store to search and the amount of time or level of effort spent in the store looking at prices. This level of effort is an increasing function of the number of brands examined. The shopping time needed to examine a given number of brands also depends (in a complex way) on the level of in-store advertising. The lower the costs of sampling the price of a given brand, the more likely it is that a consumer will in fact sample that brand. An in-store advertisement for a given brand may have the effect of directly lowering the cost of gathering price information about that brand by having a comparatively large sign that states the price of the brand and can easily be read by the consumer while he or she is moving down the aisle.

A display for a given brand also may serve as a signal of a price reduction for that brand even if the display itself does not directly reduce the cost of gathering price information. If displays and price reductions tend to be positively correlated, then a consumer has a higher expected return to obtaining price information about brands that are on display. The market mechanism leading to such a correlation might be the following one. Displays are costly to the seller. If the effect of a display is to attract consumers' attention and get them to obtain price information (as opposed to blindly purchasing that brand), then a display will lead to a greater increase in sales (and be more profitable) when coupled with a low price rather than with a high one. This notion of display activity as a signal of a low price is a variant of Nelson's (1974) model of advertising as a signal of quality. Here, the disciplining mechanism includes the initial purchase and not solely repeat purchases.

In fact, price *reductions* and in-store displays are positively correlated for several of the brand/chain pairs. Table 3 presents the correlations between retail prices and other promotional activities for the four leading brands of ground, caffeinated coffees. Note that a negative correlation indicates that the promotional activity tended to accompany a price cut. One of the chains (Chain 2) tended to coordinate price cuts with promotions (on feature or on display) more than did the other two chains.

Features (price advertising in print media) may reduce consumers' costs of direct information aquistion by allowing them to search during times when they have comparatively low shadow prices on their time, or by allowing consumers to sample different stores without having to incur the transportation costs or the opportunity costs of the time spent going to the store. Moreover, consumers also may use store features as signals of low prices. Store features are correlated with price reductions and hence potentially could serve as a signal of price cuts, again as shown in Table 3. We shall have more to say on the possible signalling role of features when we present our regression analysis in Section 4.

2. Quality Information

The most reliable way for a consumer to gather information about whether he or she likes the quality or taste attributes of a brand is, of course, to purchase the coffee and drink it. Coffee would appear to conform well to Nelson's (1970) notion of an experience good. It is difficult to convey the taste of the coffee prior to purchase, but after a single purchase the consumer should have a good idea of the attributes of the coffee. A consumer will experiment by purchasing different brands if he or she has some expectation that the trial brand will offer a better price-quality mix than those brands about which the consumer already has information, and if the costs of search are sufficiently low. The cost of this form of information gathering is bounded above by the cost (including the time and transportation) of purchasing the smallest available can. Thus, the costs of this type of search are very low for those brands for which trial sizes are offered by the manufacturers.

In choosing which brands to sample, a consumer may engage in what Nelson (1970) calls "guided sampling," whereby the consumer makes his or her sampling decisions based on the recommendations of acquaintances with whom he or she believes his or her tastes are positively correlated. The sampling decision may also be based on indirect inferences obtained from the presence of promotional activities. For example, the fact that a firm issues coupons or engages in advertising may serve as a signal of product quality if consumers view these promotions as investments in repeat purchases.

In addition to serving a signalling role, consumers may believe the claims made by advertisers about the quality of their products for psychological reasons having to do with something other than the advertising-as-a-signal model. Marketing professionals claim that advertising can be used to influence consumers' beliefs about brands that they already have sampled.[9] In particular, it is claimed that television advertising can influence the image of a brand.

[9] *The U.S. Retail Coffee Market (A)* (1982) Harvard Business School.

D. *Deal Activity—Coupon Use*

Consumer activities conducted in order to obtain special price reductions or deals comprise the final component of shopping behavior. Attempts at bargaining are unlikely to meet with much success in the retail coffee market. A coffee consumer could, however, obtain lower prices through utilizing coupons. Clipping and using coupons is an activity that consumes time, and we would expect it to become increasingly time consuming for a household to find additional coupons. The consumer will balance the price reduction granted by an additional coupon against the costs of finding, clipping, and using it. Here, we mean the incremental cost of using coupons, and we do not include the time that would have been spent making the purchase even in the absence of coupons (*e.g.*, driving to the store—unless more trips are required when coupons are used). The consumer will choose his or her levels of shopping and coupon use to minimize the expected "full price" of purchasing coffee, where the full price is equal to the retail price minus the value of the coupon plus the costs of obtaining and using the coupon.

E. *Implications for Demand*

1. Store and Brand Choice

Having obtained information and formed expectations about prices, qualities, and deal availability, consumers will make purchase decisions. Consumers must choose the brand(s) and quantity of coffee to consume and the store from which to purchase it. The store decision is likely to depend on an entire basket of prices for various items purchased at the supermarket. Coffee prices typically are only a small part of consumers' implicit supermarket price indices. Locational differences among stores also may play a significant role in consumers' store choices. Thus, there is a degree of differentiation among stores that coffee manufacturers cannot influence.

The quantity of a given brand of coffee that a consumer purchases depends on the price of that brand relative to other brands of coffee and relative to the prices of other goods and services. The relative prices of coffee brands are likely to influence the brand choice, while the relative price between a brand of coffee and other goods and services is likely to influence the quantity of the chosen brand that is purchased. The result will be the usual negative relationship between a brand's sales and its price, while the prices of competing brands can be expected to be positively related to the quantity demanded. We would expect that the prices charged at store i would have less influence on the sales of a brand sold at store j than would the prices charged for the other brands at store j for two reasons. First, as discussed above, consumers will have preferences over stores, which will reduce the degree of between-store switching. Second, a household that shops at one store is likely to have less information about the prices charged at other stores than about the prices charged at the household's chosen store since within-store price search is much less costly than between-store price search.

2. Stockpiling

In addition to considering current substitution between the consumption of coffee and other goods, consumers may engage in inventory behavior. Consumers may form expectations about the paths of prices over time. Whenever the consumer believes that prices will rise at a rate greater than storage costs plus interest, the consumer will have incentives to stockpile the good. Of course, consumers also will maintain inventories of coffee so that they can economize on the fixed costs of going to the store; no consumer wants to go to the store every time he or she wants a cup of coffee. The extent to which consumers economize on transactions costs also will be limited by interest and storage costs. In the presence of stockpiling behavior, past sales of coffee may have a negative impact on current sales of coffee.

3. Interaction of Promotions and Price Changes

Consumers can be expected to make purchase decisions based on incomplete information concerning both the physical attributes and the prices of the alternative brands of coffee that are available in the market. Although direct observation provides a very credible source of price information, it is costly, and given the frequent changes in retail prices, consumers are unlikely to have complete price information.

Features, displays, and coupons affect consumer awareness of the product and perceptions of its price. We would expect that price cuts are most effective when they are implemented at the same time that the brand is featured in a local print advertisement or subject to an in-store display. The interaction of feature activity and prices should be particularly pronounced when one looks at the effects between stores, where features give rise to an especially strong reduction in a households' shopping costs. In-store search over prices is unlikely to take place at stores from which the household does not make a purchase due to the high cost of such a search strategy. The household can, however, learn about coffee prices at other stores by looking at retailers' local print advertisements. The between-store interaction effects for price and display activity should be much less important because a household will see a display at a given store only if the household shops at that store. We test these hypotheses below by including interaction terms between price and promotions in our demand function regressions.

4. Information Theory of Kinked Demand

As we have stated, when there are many brands available to the consumer, he or she may not scan the prices of all brands when in the store. Moreover, when there are multiple retail outlets for the brands under consideration, the consumer almost certainly will not know some of the prices at stores other than the one at which he or she makes a purchase. For these two reasons, a price cut for a brand-store combination other than the one that is purchased could go unnoticed. This effect often is given as a reason why price competition may not be effective with imperfect price information; price increases are quickly noted by existing customers, while price reductions are not observed by other consumers. The result of this asymmetry will be a kinked demand curve as illustrated in Figure 1. We test for asymmetric responses to price increases and decreases in our demand analysis presented in Section 4.

3. A Look At Consumer Shopping Behavior

A. *Overview*

Our data were obtained from Information Resources, Inc., a marketing research firm. The data cover the purchases of 935 households over a 108 week period (April 1980-April 1982). For each household, a time series of all of that household's purchases in the coffee category is provided. For each purchase, the time of purchase, the store in which it was made, the price paid, and the use of either manufacturer or retailer coupons are given. In total, about 50,000 purchase events are included in our sample. These events occured at six stores in the Pittsfield, Massachusetts area. The six stores are grouped into three chains. The data were collected through scanners that read the universal product codes of all goods purchased by participants in the study who identified themselves while checking-out.[10] In addition to data on quantities and retail prices, we have information on in-store displays, local print advertising, and the availablity of retailer coupons for each of the six supermarkets on a weekly basis.

We do not have direct evidence of the ways in which consumers incorporate retailer feature and display activities into their information portfolios. Indirect evidence, however, suggests that these activities have an important influence on consumer behavior. Consider features first. As noted in Section 2, retailers ran features for brands 11 percent of the time. Moreover, there was no brand that was featured for more than 27 percent of the weeks by any given chain. Yet, 52.1 percent (by weight) of all ground, caffeinated coffee purchases included in our sample were made during a week when the brand bought was featured in the store's local print advertising.

Looking at display activity, a similar picture emerges. The four leading brands were on display only 12 percent of the time. The single strong outlier was Chock Full of Nuts at Chain 2, where it was on display 42 percent of the time. This brand/chain combination accounted for only 10.5 percent of total sales in our sample. Despite the infrequency of display activity, purchases of brands while

[10] Although we have not tested the sample to determine whether it is a representative one, it is worth noting a point made by Telser (1962) in a very similar context. This data originally was collected to be sold to coffee sellers. Given that sellers are willing to pay a large amount of money for this data, they appear to believe that it accurately reflects the demand conditions that they face.

they were on display accounted for 41.9 percent of all purchases of ground, caffeinated coffee.

Purchases made using coupons comprise a large percentage of total purchases, whether measured in terms of the number of purchases made or the quantities of coffee bought. Consumers used coupons for 24.5 percent (by weight) of all the purchases made in our sample over the full two year period. Of purchases made using coupons, approximately two thirds utilized stores' coupons and one third utilized manufacturers' coupons. About 7 percent of purchases made using coupons utilized store and manufacturer coupons simultaneously. The use of coupons varies greatly across households, as Table 4 illustrates. Over 94 percent of the households in our sample used coupons at least once to purchase some form of coffee (including instant and decaffeinated ground). There is a weak, negative correlation (-0.1) between the total quantity of ground, caffeinated coffee purchased by household and the percentage of purchases made by the household using coupons.

The inferences about the individual importance of various promotions drawn from the data on purchase percentages need to be qualified somewhat. Stores often used promotions in combination with each other. Table 5 gives a sample of the correlations between the use of coupons, features, and displays. Over 78 percent of the purchases of ground, caffeinated coffee that were made using coupons took place while the brand purchased was either on display or on feature, although only 57 percent of all ground, caffeinated purchases were made while on feature or on display. Given these correlations, it is difficult to disentangle the effects of alternative promotions through simple summary statistics.

If consumers switch stores in response to between-store differentials in coffee prices, or if consumers routinely vary their purchases among chains for other reasons (such as differentials in the price of milk), then the demand for a given brand of coffee at one store will be cross-elastic with the brands of coffee sold at other chains. Because the pattern of pricing and promotional activity does not vary across stores within a given chain, we examine switching among retail outlets at the chain rather than store level. On average, a household made 78.5 percent (by weight) of its purchases from its preferred chain. Over 47 percent of the households in our sample switched the chain from which they purchased ground, caffeinated coffee two or fewer times during the two year sample period. A few consumers switched chains quite frequently, however, so that the average number of chain switches per household is 7.67.

To get a sense of the extent to which consumers stockpile coffee, we looked at the variance in the time intervals between the purchases made by a given household and at the variance in the size of purchases made by a given household. If a consumer's only motivation for holding inventories is to minimize transactions costs, if his or her coffee consumption rate is roughly constant, and if there are no smoothing problems due to the fact that only pound multiples may be purchased, then the within-household variances of both the times between coffee purchases and the sizes of the purchases made will be low. For our sample households, the mean coefficient of variation (the ratio of the standard deviation to the mean) for times between purchases was equal to .84, and the mean time between purchase occasions was 7.4 weeks. The mean coefficient of variation for purchase sizes was equal to .36, with a mean purchase size of 23.4 ounces.[11] Both of these numbers indicate that consumers did not pursue steady-state purchasing patterns. These results weakly support the hypothesis that households engage in stocking up for "speculative" reasons as well as to minimize transactions costs.

B. *Brand Switching Behavior*

We turn now to a more detailed description of brand purchase behavior. Specifically, we study the pattern of interbrand switching by consumers. In the next subsection, we will relate this behavior to promotional activity. In the present section, we examine the breadth of consumer switching (*i.e.*, how many brands does a given household purchase) and the frequency of consumer switching (*i.e.*, how often does a household change brands).

As a prelude to this examination, we report, in Table 6, the market shares by brand for the ground, caffeinated segment. The four brands upon which we focus are Maxwell House, Folgers, Chock Full of Nuts, and Hills Brothers. The combined market share of these four brands is 74.4 percent. We treat private label and generic ground, caffeinated coffees as a single composite brand in our analysis of consumers' switching in and out of the dominant branded coffees.

We noted above that coffee is an experience good. We therefore can view the breadth of a household's consumption experience across brands as a measure of the information that it possesses about product attributes. Our data cover a limited period. Given the durability of information about brand attributes, a possible explanation of a finding of low switching and sampling rates during the period covered by our data could be that consumers were well-informed (based on past consumption experience) and had strong brand preferences. A finding that consumers switch brands often would indicate that consumers are likely to be informed about product attributes (they will have gathered information through experience) and thus are not likely to be susceptible to the vendor market power that could otherwise arise due to consumers' imperfect information about the set of alternative brands that are available.

Table 7 reports the number of brands purchased by household over the two year sample period. The most striking feature is the extent to which consumers purchase multiple brands. Only 18 percent of sample households stayed with a single brand over the two years, and half of the households tried at least four brands. A similar picture emerges when one looks at the number of vendors (manufacturers) sampled per household. There is a distinction between brands and vendors because several vendors produce multiple brands. For ground, caf-

[11] The mean purchase size is the number of ounces bought per week averaged over the set of weeks in which the household had positive purchase quantities.

feinated coffee this distinction turns out to be an unimportant one.[12]

Information about product quality is likely to be durable. Given the frequency of purchase (*see* Table 2) and the number of different brands sampled, consumers appear to have good information about the quality attributes of several brands. In light of the fact that prices change frequently and there are several stores that pursue different pricing policies, consumers may actually be better informed about quality than price in this market, at least for the major brands. For other, less frequently purchased and more complicated goods, one typically thinks of quality information as being more difficult to obtain than price information.

We now consider the frequency of brand switching. We are interested in the retention, or loyalty, rates for the major brands, *i.e.*, the fraction of their customers who return to the same brand on their next purchase occasion. For each household we have computed a matrix of switches, where a switch is calculated by looking at the brands chosen on two successive purchase occasions. S_{ij} is the number of switches from brand i to brand j, weighted by the size of the purchase measured in ounces. Thus, for example, S_{ii} is the number of times that households made two consecutive purchases of brand i. We define the ***Markov brand loyalty*** rate as

$$S_{ii} \; / \; \sum_j S_{ij} \; .$$

This rate tells us the percentage of customers that a brand retains for another purchase given that they have purchased the brand today. The first row of Table 8 presents the Markov brand loyalty rates for the four major brands of ground, caffeinated coffee averaged across households.

A comparison of the overall market shares (from Table 6) and the Markov brand loyalty rates verifies the presence of significant brand loyalty. For every brand, the Markov brand loyalty rate is much higher than the brand's market share. This finding is a natural one in differentiated products markets with heterogeneous consumers; informed consumers tend to stay with their preferred brands.

A second way to compute brand loyalty is to compute the percentage of its total purchases of ground, caffeinated coffee that a given household makes of a single brand. For each of the four leading brands, Table 9 gives the percentage of households for whom that brand was their favorite. For the households favoring a particular brand, Table 9 also gives the purchases of these households' favorite brands as a percentage of the households' total purchases. This measure, too, shows that brand loyalty is present.

[12] In the decaffeinated and soluble segments this distinction can be substantial. For example, in the ground, decaffeinated segment the two dominant brands, Brim and Sanka, are produced by the same vendor, General Foods.

C. *Brand Switching and Promotional Activity*

We can use our data on consumer switching among brands to get some idea about possible motivations for firms' promotional behavior. Rational consumers will use coupons to minimize the full price (*i.e.*, the transactions price plus shopping costs) that they pay for coffee. But is it rational for firms to incur the expense of issuing coupons? In particular, from the seller's perspective what is the advantage from offering a coupon rather than making a retail price cut? One difference between a price cut and a coupon is that a coupon may serve as a particular form of advertising (the coupon is a means of getting the consumer's attention). Even viewed strictly as a means of reducing price, though, there are two important differences between coupons and ordinary price cuts: (1) coupons may be a means of price discriminating among consumers; and (2) coupons may be a means of preventing consumers from stocking-up coffee purchased at reduced prices.

Coupons can be a profitable means of facilitating price discrimination or screening when there is a negative correlation between a consumer's willingness to pay for the given brand and the consumer's propensity to use coupons. In such cases, one could think of there being two types of consumer. One type of consumer is price-conscious, uses coupons, and purchases from the brand offering the lowest effective price. The second type of consumer is relatively insensitive to price, does not clip coupons, and makes purchase decisions on the basis of loyalty to, or preferences for, a given brand. When a regular price cut is made, type-two consumers loyal to that brand receive the price reduction even though they would have purchased it anyway. By tying the reduced price to the use of coupons, the seller can avoid giving price reductions to such inframarginal consumers who would have purchased anyway.

Coupons may also serve as a means of targeting new users and preventing established users from stocking up. If a brand makes a price reduction that consumers believe is only temporary, then those consumers who have information about the brand and prefer it to others will stock up on the brand at the low price. In effect, the brand will be taking (future) sales from itself. When the price cut is achieved through the issuance of a coupon, the lower price is available only for the initial unit that a consumer purchases (unless the consumer obtains multiple coupons). The use of coupons will limit existing customers' ability to establish inventories at the low price. At the same time, the coupon will lower the new user's cost of sampling the brand by making a single purchase.[13]

Under the hypothesis that coupon users are price-conscious consumers who switch brands frequently, the loyalty rate among coupon users should be lower than the loyalty rate among consumers who did not use coupons to make the initial purchase in the S_{ij} calculation. Table 8 shows that this is in fact the pat-

[13] The staff of the Federal Trade Commission has suggested that geographic price discrimination, where coupons serve as a means of preventing retailer arbitrage across regions, also may be a motivation for manufacturers to issue coupons.

tern that one finds. For example, overall, 62 percent of the time that a household purchases Chock Full o' Nuts, its next purchase also will be Chock Full o' Nuts. Conditional on the initial purchase being made using a coupon, however, the percentage of repeat purchases is only 49 percent.

This pattern also is consistent with the view that the role of coupons is to attract additional customers without allowing existing customers to stock-up in order to take advantage of the temporarily low prices. When an ordinary price cut, rather than a coupon, is used to attract consumers, households may stock-up on the brand. These stock-up purchases will be recorded as S_{ii} transactions since each can is treated as a separate purchase event. Thus, the set of purchase pairs where coupons were not used for the initial purchase should be more likely to contain a high proportion of S_{ii} pairs.

We have done a similar analysis of feature and display promotional activity (*see* Table 10). Again, we find that consumers who came to brand i while it was being promoted are less likely to remain loyal.[14] This finding is consistent with the view that promotional activity encourages sampling. Note that the effects of features and displays on the Markov loyalty rates seem to be less pronounced than in the case of coupons.

The brands *from* which consumers switch to brand i also can be analyzed using our consumer switching data. The percentage of brand i purchases that come from consumers who last purchased brand i is given by

$$S_{ii} / \sum_j S_{ji} .$$

The first line of Table 11 gives the figures for the four leading brands calculated over all purchase pairs of ground, caffeinated coffee. Again, our main interest is in the split samples, where the division is by whether the *second* purchase in the S_{ji} was made using a coupon. The table shows that coupons do indeed attract a large proportion of "new" users to the brand, *i.e.*, nonrepeat users. For example, 64 percent of overall purchases of Chock Full o' Nuts were repeats, while only 29 percent of Chock Full o' Nuts coupon purchases were repeats. Coupons do appear to selectively discount to price conscious consumers and/or prevent stocking up. Much the same pattern is found when purchases are categorized by display and feature activity (*see* Table 12).

4. The Determinants of Demand

A. *Estimated Demand Functions*

Our final results involve aggregate sales for a given brand/chain combination, rather than purchasing behavior on an individual level. Our data set allows us

[14] These results are consistent with those obtained by Guadagni and Little, who fit a logit model to a different set of panel data for purchases of ground, caffeinated coffee.

to include promotional variables in addition to price in order to test some of the implications derived in our earlier discussion of the theory of consumer shopping. Because each chain has its own pattern of pricing and promotional activity (which does not vary across stores within a given chain), we will examine each chain separately from the others in the analysis below. To reduce the number of parameters to be estimated in our demand equations, the prices of all branded ground, caffeinated coffees other than the four leading ones were aggregated. A price index was constructed using quantities to weight the prices in the index. Similarly, the prices of all private label and generic ground, caffeinated coffees for a given chain were aggregated (again using quantities as the weights) to form a chain-specific private label price index.

In estimating the demand for a given brand at a specific chain, we included the prices and promotional activities for that brand at all three chains as explanatory variables, while for brands other than the one whose demand we were estimating only the prices and promotional activities at the chain under consideration were included. More specifically, we ran regressions of the following form:

$$x^{m}_{it} = \alpha + \sum_{n} \{\beta^{n} p^{n}_{it} + \gamma^{n} F^{n}_{it} + \delta^{n} p^{n}_{it} F^{n}_{it} + \epsilon^{n} p^{n}_{it} F^{n}_{it} + \mu^{n} D^{n}_{it}\}$$

$$+ \sum_{k \neq i} \{\xi_{k} p^{m}_{kt} + \lambda_{k} F^{m}_{kt} + \sigma_{k} p^{m}_{kt} F^{m}_{kt} + \varrho_{k} p^{m}_{kt} F^{m}_{kt} + \tau_{k} D^{m}_{kt}\}$$

where x^{n}_{it} is the quantity of brand i sold by chain n in week t. p^{n}_{it}, F^{n}_{it}, and D^{n}_{it} represent price, feature dummy (=1 if on feature), and display dummy (=1 is on display), with the subscripts and superscripts defined as for the quantity variable.[15] Greek letters denote the coefficients to be estimated. Terms involving feature and display dummies were not included for the two composite "brands."

In the ground, caffeinated segment of the market alone, we have four major brands, each of which is sold at three chains, across which prices and promotional activities vary. This gives us 12 regressions for a given specification. Due to space limitations, we have chosen to report the results from only one of the chains. At two of the chains some of the promotional activities were linearly dependent combinations of others, making it difficult to interpret the results, as well as making some of the estimates unreliable. On this basis we chose to report the regressions results for the remaining chain, Chain 2, only.

[15] Attempts to fit linear-log and log-log (where possible) functional forms proved to be less successful, and the results are not presented here. We also ran regressions using market share as the dependent variable. Those regressions yielded results similar to the ones reported in the text. In some of our share regressions, we also included the average (over the past month) of the market share of the brand whose share we were explaining. These terms were included to capture stockpiling effects and to allow us to distinguish between long and short run effects of price changes. The variables tended to be small in value and insignificant.

These results are displayed in Table 13.[16] We found that the feature variables were significant and had large absolute effects. For example, while Chock Full 'o Nuts had a mean sales level of 421 ounces per week over the full sample period, being on feature tended to increase its weekly sales by 1,435 ounces. For each of the four major brands, being on feature led to increases in sales that were at least this large. The display variables were much less successful as explanators, and often had perverse signs.

The price variables tended to have the expected signs (under the assumption that all of the brands are substitutes with one another), but many of the cross-brand price effects were not statistically significant.[17] The own-price elasticities of demand for the four major brands are of particular interest. Since our specification is a linear one, we must choose a point at which to calculate the own price elasticities. As is conventional, we do so at the means of the relevant variables. The resulting point estimates of the own-price elasticities of demand are shown in Table 14.

B. *Interaction Between Price Advertising and Price Changes Within A Single Store*

The terms that are the product of the brand i's price with its feature activity dummy are included to capture the idea that a price change should be more effective if accompanied by a display or a feature; more consumers will be aware of the price change under such circumstances. The estimated effect of an unadvertised price change of brand i at chain n is given by β^n. The estimated effect of an advertised price change is given by the sum of β^n and δ^n. δ^n captures the extra impact that a price change has when consumers are made better aware of it.

We see from Table 13 that the interaction between a brand's price and feature activity tended to have the predicted effect on the brand's sales (a price increase lost more sales when the brand was on feature). This relationship is illustrated in Table 14, where for each of the four brands the elasticity of demand in response to an advertised price is higher than the elasticity for an unadvertised price. This relationship is consistent with the theory that feature activity conveys price

[16] We speculated that the poor fit for Maxwell House's ground, caffeinated coffee might be the result of our having omitted Maxwell House's flaked coffee, Master Blend, from the regressions. We ran a regression restricting the sample to the period prior to the introduction of Master Blend (roughly half of our sample), and the fit for Maxwell House's regular ground, caffeinated coffee was equally poor.

[17] The reader might be concerned about the omission of the prices of other types of coffee from our regressions. A 1978 study conducted for the International Coffee Organization (*The U.S. Retail Coffee Market (A)* (1982), page 8) found that less than 6 percent of consumers of caffeinated coffee also drink decaffeinated coffee. Similarly, a 1979 study for the same body (*The U.S. Retail Coffee Market (A)* (1982), page 6) found that less than 20 percent of consumers who drink ground coffee also drink soluble, or instant, coffee. These figures include consumption outside the home, and thus probably overstate the extent to which consumers purchase multiple types of coffee in supermarkets.

information. The estimates obtained for the effect that each brand's display-price interaction term had on its sales, on the other hand, do not support the theory that displays provided price information; three of the four estimated coefficients were statistically significantly different from zero and had the opposite of the predicted sign (they were positive).

The most surprising results are that feature activity in itself is so important, even when prices and price-feature interaction terms are included in the regression.[18] One possible explanation is that consumers use features as signals of price cuts. We regressed prices on feature activity to determine the validity of the use of features as signals of price cuts. The week, the square of the week, and display activity also were included to capture any time trends and the effects of other promotions, respectively. Thus, the regressions were of the form:

$$p^{m}_{it} = \alpha + \beta F^{m}_{it} + \gamma D^{m}_{it} + \delta t + \mu t^2 .$$

This regression analysis indicates that on-feature is a weak signal of low prices. While in 9 of 12 of these regressions β did have a negative sign, in 8 of 12 cases the coefficients were insignificant at the .05 level. Display activity is even less valuable as a predictor of price, being significant for only 3 of the 11 brand-chain combinations where displays were used, and for one of those 3 combinations the display coefficient is positive. If consumers do place too much emphasis on feature activity, then we might expect chains to use it more often than they do. An interesting next step in the analysis of feature activity would be to develop estimates of retailers' costs of this type of advertising. Once developed, the estimated costs and benefits to the retailer and manufacturer could be compared to see if they were consistent with the hypothesis that retailers and manufacturers (the latter through giving retailers incentives) made profit-maximizing use of these promotions.

C. *Do Prices and Promotions at Other Stores Matter?*

Changes in the price of a brand at one chain may induce consumers to make substitutions of purchases of that brand at one chain for purchases of the *same* brand at another chain. If such within-brand, between-store switching is important, then attempts to gauge manufacturer market power by looking solely at the effect of price on sales at a single chain will overstate the elasticity of demand faced by the manufacturer.

Do the prices and features at one chain affect the quantity of coffee demanded at another? As discussed in Section 2.E, we would expect that unadvertised prices at other chains would have a very weak effect on sales at a given chain. Advertised prices, those on feature that week, might have a stronger, positive effect, but we would expect this effect to be dampened by households' prefer-

[18] The importance of promotional activity through effects other than interaction with prices also was found by Guadagni and Little.

ences over chains. Feature activity at competing chains might have a negative effect on sales. In terms of our regression, for $n \neq m$, these hypotheses are: $\beta^n = 0$; $\beta^n + \delta^n > 0$; and $\gamma^n < 0$. Our results concerning the effects of prices and promotions at chains 1 and 3 on sales at chain 2 were mixed. For example, an unadvertised rise in the price of Folgers at chain 1 had a positive and significant effect on the sales of Folgers at chain 2. Due to multicollinearity, one cannot determine the corresponding effect for an advertised price increase. For Hills Brothers, $\beta^1 + \delta^1$ is the only one of these terms that was significantly different from zero at the 90 percent confidence level. For display activity, we would expect the cross-chain effects to be minimal, and the regression results support this hypothesis.[19]

Given the cross-chain effects of price on demand in some cases, the elasticities of demand faced by the manufacturer and retailer will differ. To calculate the own-retail-price elasticity of demand faced by a coffee producer, we must look at the total effect on sales across all three chains when the price of a given brand is raised by one percent at all of the chains simultaneously. Again making this calculation at the mean sales level, we find that Chock Full o' Nuts has an unadvertised price elasticity of -2.0, Maxwell House 2.0, Folgers -9.4, and Hills Brothers -20.2.

D. *Kinked Demand Theory*

When price information is costly to gather, a seller may face a kinked demand curve, at least with respect to unadvertised price changes (*see* Section 2.E). We tested this hypothesis by running the following regression, which includes separate coefficients for the effects of price increases and price decreases, in addition to a feature-price interaction term and the first difference of feature activity:

$$x_{it}^m = x_{it-1}^m + \sum_j [\gamma_j(p_{jt}^m - p_{jt-1}^m)U_{jt}^m + \delta_j(p_{jt}^m - p_{jt-1}^m)(1 - U_{jt}^m)$$

$$+ \xi_j(p_{jt}^m - p_{jt-1}^m)F_{jt}^m + \omega_j(F_{jt}^m - F_{jt-1}^m)]$$

where

$$U_{jt}^m = \begin{cases} 1 \text{ if } p_{jt}^m - p_{jt-1}^m \geq 0 \\ 0 \text{ otherwise} . \end{cases}$$

[19] The finding of generally weak cross-store effects appears to contradict the conventional wisdom in food retailing that consumers will switch stores in response to swings in the price of coffee. One explanation of this seeming discrepancy comes from the fact that the consumers' sensitivity to the price of coffee in choosing a store is greatly magnified from the retailer's perspective. When a rise in the price of coffee induces a consumer to shop at another store, the retailer will tend to lose the consumer's entire set of supermarket purchases.

If the kinked demand curve theory is correct, then both γ_j and δ_j will be negative, but γ_j will be larger in absolute value. Constructing an F-test, and again restricting attention to chain 2, for all of the brands except Folgers one cannot reject the null hypothesis that $\gamma_j = \delta_j$ even with a confidence level of only 20 percent. For Folgers, one can reject the null hypothesis of equality even at the 99 percent level. The Folgers' kink, however, is the opposite of the one predicted by the theory; δ_i is larger in absolute value than is γ_i. The cross-brand price effects tend to be insignificant and the kinks in the demand for brand i with respect to the price of brand j, $j \neq i$, tend to be insignificant as well. In this market, at least, the imperfection of consumer information does not appear to give rise to kinked demands.

5. Conclusion

We have examined both consumer and firm behavior in the retail coffee market in some detail. The patterns of consumers' shopping behavior and firms' promotional activities that we observed have important implications for the performance of the retail coffee market.

First, consider consumer shopping behavior. Given the frequency with which households purchase coffee, we would expect experience to be an important source of quality information. The relatively large number of brands (3.7 on average) sampled per household over the two year period covered by our data indicates that consumers should be well informed about the (quality) attributes of the various ground, caffeinated coffees. At the same time, the typical consumer's information about prices, especially at stores other than the one at which the consumer shops, may be highly imperfect.

There are two indicators that consumers' price information may be limited. First, in our regression analysis, we found that prices at rival supermarket chains tend to be insignificant in explaining sales at a given chain. This pattern is consistent with consumers having poor information about the prices at chains other than the ones from which they make purchases. Of course, it also is consistent with consumers having strong preferences among stores. Second, consumers react strongly to local print advertisements (features) that typically provide price information. As we have seen, the strength of consumers' reaction to features even after price-feature interaction effects have been accounted for is surprising; our regressions of prices on time, display activity, and feature activity show that features are a poor predictor of price—although features are weakly correlated with price cuts. Consumers may incorrectly be taking features to be strong signals of price cuts, indicating that consumers are not well-informed about prices.

In summary, we see a pattern that may be typical of consumer non-durables that are purchased relatively frequently: consumers' price information is poor relative to their quality information. This pattern occurs when experience is an important source of information and product qualities change much less frequently than do prices, so that information gathered through experience tends to depreciate less rapidly for qualities than for prices. The opposite pattern tends to hold for less frequently purchased, "big-ticket" items, where experience is less likely to be an important source of information. The appropriate consumer protection policies for different types of goods thus are likely to be rather different.

The use of coupons constitutes another aspect of consumer shopping behavior. Consumers do use coupons, and some households use them quite heavily. Households are willing to switch away from their favorite brands to alternative brands in order to make use of coupons. A household that has purchased a brand using a coupon is less likely to remain with that brand for its next purchase, and is less likely to have bought that brand in the past, than is a household that has not used a coupon to make its purchase.

Turning to the sellers' side of the market, our results indicate the great importance of promotional activity. One dimension of promotional activity that we have investigated from the seller's perspective is the use of store- and manufacturer-issued coupons. Our findings are consistent both with the theory that coupons are a way of reducing prices without permitting loyal customers to stock up, and with the theory that coupons are a means of targeting price reductions at price-conscious consumers.

In our regression analysis, we found that promotional activities are important determinants of the quantities of the brands that are purchased. These results point out a danger in using highly aggregated data to estimate demands. Once data are aggregated across chains, it is difficult to look at promotional activity. In view of the promotional activity which is so prevalent in the retail coffee market, and in view of the correlation between price reductions and promotional activities, prices would appear to have more significance than they in fact do if the promotional variables were omitted from the regressions.

Even when promotional activities are accounted for in the analysis, the own-price effects on demand are significant. There appears to be little market power at either the manufacturer or retailer level. Consumers have enough information about the prices of alternative brands within a single supermarket that they do switch brands in response to price changes. Moreover, the level of within-store price information is sufficient to eliminate any kink in the demand for a particular brand. Further, both the brand switching analysis and the regression results support the hypothesis that feature activity strengthens price competition and reduces market power. Neither product differentiation nor the incompleteness of consumer information is a source of significant market power.

References

Guadagni, Peter M. and Little, John D.C., "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, forthcoming.

Nelson, Phillip, "Information and Consumer Behavior," *Journal of Political Economy*, March-April, 1970, Vol. 78, No.2, 311-329.

____________ "Advertising as Information," *Journal of Political Economy*, July-August, 1974, Vol. 81, No.4, 729-754.

Porter, Michael E., ***Interbrand Choice, Strategy, and Bilateral Market Power***, Cambridge: Massachusetts: Harvard University Press, 1976

Telser, Lester, "The Demand for Branded Goods as Estimated from Consumer Panel Data," *The Review of Economics and Statistics*, August, 1962, Vol. XLIV, No. 3, 300-324.

The U.S. Retail Coffee Market (A) (1982), Harvard Business School Boston: International Case Clearing House 9-582-087.

Figure 1

KINKED DEMAND CURVE

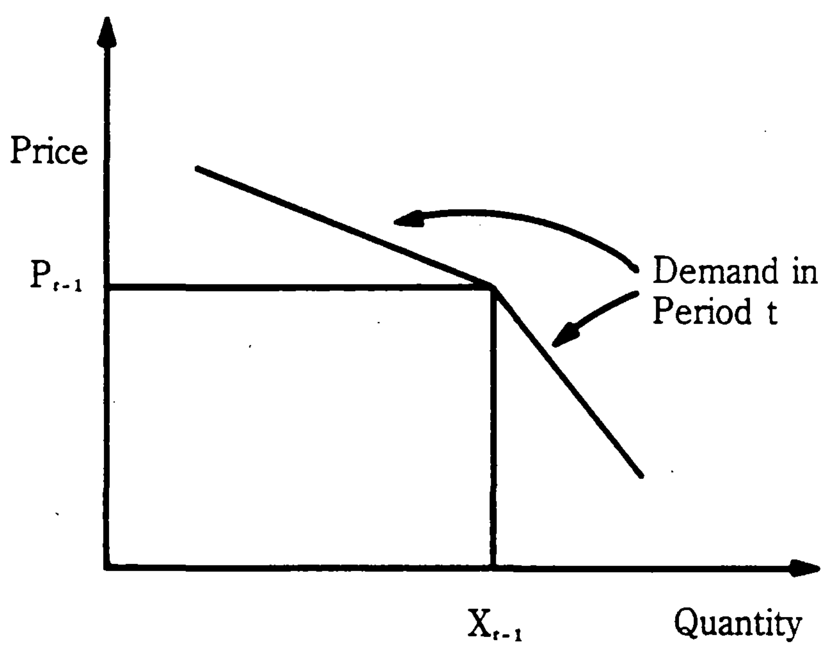


Table 1
RETAILER USE OF PROMOTIONS

Brand	Percentage of chain-weeks featured	Percentage of chain-weeks displayed
Chock Full o' Nuts	15	27
Maxwell House	9	8
Folgers	10	7
Hills Brothers	10	10
Average Across All Four Brands	11	12

Table 2
DISTRIBUTION OF PURCHASE NUMBERS ACROSS HOUSEHOLDS

Number of Purchases	Proportion of Households	Cumulative Proportion
1-5	28 percent	28 percent
6-10	12 percent	40 percent
11-20	14 percent	54 percent
21-40	16 percent	70 percent
41-60	13 percent	84 percent
61-80	7 percent	91 percent
81-100	4 percent	95 percent
> 100	5 percent	100 percent

Mean number of purchases per household: 30.5.

The top 5 percent of the households account for 21.4 percent of the purchases.

Here, a purchase is defined as buying a single can. Thus, a consumer who buys several cans of coffee during a single trip to a store would be reported as making multiple purchases.

Table 3
SAMPLE CORRELATIONS BETWEEN PRICE AND PROMOTIONS

Brand/Chain	Correlation Between Price On Display	Correlation Between Price On Feature	Correlation Between Price and Store Coupon
Maxwell House/1	-.01	.02	-.04
Maxwell House/2	-.17	-.04	-.07
Maxwell House/3	.08	-.11	-.13
Folgers/1	-.23	-.24	-.08
Folgers/2	-.57	-.61	N/A
Folgers/3	.00	-.22	N/A
Chock Full o' Nuts/1	N/A	-.04	-.04
Chock Full o' Nuts/2	-.22	-.39	N/A
Chock Full o' Nuts/3	-.38	-.28	-.17
Hills Brothers/1	-.34	-.37	-.06
Hills Brothers/2	-.55	-.60	N/A
Hills Brothers/3	.50	.30	N/A

N/A : not applicable.

Table 4
DISTRIBUTION OF COUPON USE ACROSS HOUSEHOLDS
(For all types of coffee)

Number of Purchases Made Using Coupons	Proportion of Households	Cumulative Proportion
0	6 percent	6 percent
1-10	47 percent	53 percent
11-20	24 percent	77 percent
21-30	13 percent	89 percent
31-40	6 percent	95 percent
41-50	2 percent	98 percent
> 50	2 percent	100 percent

Mean number of purchases using coupon per household: 13.9.

Table 5
SAMPLE CORRELATIONS AMONG PROMOTIONS
(Maxwell House ground, caffeinated at chain 2)

	On Display	On Feature	Store Coupon Issued
On Display	1.0	0.45	0.24
On Feature	0.45	1.0	0.72
Store Coupon Issued	0.24	0.72	1.0

Table 6
MARKET SHARES BY BRAND

Brand	Market Share	Cumulative Share
Maxwell House	24.4	24.4
Folgers	23.0	47.4
Chock Full o' Nuts	18.3	65.7
Private Label and Generic	18.0	83.7
Hills Brothers	9.0	92.7
Chase and Sanborn	2.6	95.3
Martinson	2.3	97.6
All Other Brands	2.4	100.0

Table 7
NUMBER OF BRANDS PURCHASED, BY HOUSEHOLD

Number of Brands Purchased	Proportion of Households	Cumulative Proportion
1	18 percent	18 percent
2	17 percent	35 percent
3	14 percent	49 percent
4	13 percent	62 percent
5	16 percent	78 percent
6	12 percent	90 percent
7	7 percent	96 percent
8	2 percent	99 percent
9-11	1 percent	100 percent

Mean number of brands purchased per household: 3.7.

Table 8
MARKOV LOYALTY RATE
(With breakdown by coupon use)

	Loyalty Rate			
	Chock Full o' Nuts	Maxwell House	Hills Folgers	Brothers
All Purchases	62	62	63	42
Consumers Who Came with a Purchase Using a Coupon	49	55	56	14
Consumers Who Came with a Purchase Not Using a Coupon	63	70	65	44
Difference Between Coupon and NonCoupon Rates	(14)	(15)	(9)	(30)

Table 9
QUANTITY MEASURE OF BRAND LOYALTY

Household's Most-Purchased Brand	Most-Purchased Brand as Percentage of Household's Total Purchases	Percent of Households for Whom the Brand is Their Most-Purchased One
Chock Full o' Nuts	61	21
Maxwell House	68	29
Folgers	63	22
Hills Brothers	60	2
Private Label*	68	20

*Private Label is a composite of all of the private label and generic brands that are offered.

Table 10
MARKOV LOYALTY RATE
(With breakdown by presence of features and displays)

	Loyalty Rate			
	Chock Full o' Nuts	Maxwell House	Folgers	Hills Brothers
All Purchases	62	62	63	42
Consumers Who Came While Brand Was on Feature or Display	61	55	57	42
Consumers Who Came While Brand Was Not on Feature or Display	67	68	70	43
Difference Between On and Off Rates	(6)	(13)	(13)	(1)

Table 11
WHERE CUSTOMERS COME FROM
(With breakdown by coupon use)

	Proportion of Customers Who Bought Same Brand on Previous Purchase			
	Chock Full o' Nuts	Maxwell House	Folgers	Hills Brothers
All Purchases	64	60	65	46
Consumers Who Came with a Purchase Using a Coupon	29	52	58	31
Consumers Who Came with a Purchase Not Using a Coupon	66	70	66	46
Difference Between Coupon and Non Coupon Rates	(37)	(18)	(8)	(15)

Table 12
WHERE CUSTOMERS COME FROM
(With breakdown by presence of features and displays)

	Proportion of Consumers Who Bought Same Brand on Previous Purchase			
	Chock Full o' Nuts	Maxwell House	Folgers	Hills Brothers
All Purchases	64	60	65	46
Consumers Who Came While Brand Was on Display or Feature	63	53	59	46
Consumers Who Came While Brand Was Not on Display or Feature	67	67	69	44
Difference Between On and Off Rates	(4)	(14)	(10)	(2)

Table 13
DEMAND FUNCTION REGRESSIONS
Naming Conventions

The variables that are used in the analysis are labelled according to the following conventions: CHOCK refers to Chock Full o' Nuts, MXHSE refers to Maxwell House, FOLGR refers to Folgers, HILLS refers to Hills Bros., PRIVT refers to an aggregation of private label and generic coffees, OTHER refers to an index of branded coffees with very small market shares. OWN refers to the brand given by the column index. P as a prefix refers to price (all prices are in cents per 16 oz. can of ground, caffeinated coffee), D as a prefix refers to the display dummy (which is 1 if and only if that brand is on display), F as a prefix refers to the feature dummy (defined analogously). If the variable name contains no number, then the chain referred to is Chain 2. A single digit suffix refers to the chain number.

RESULTS

	Dependent Variable: Chain 2 Quantity Sales:			
Independent Variable	Chock Full o' Nuts	Maxwell House	Folgers	Hills Brothers
INTERCEPT	334 (582)	-578 (376)	355 (393)	-196 (491)
P-CHOCK	-11.7 (2.8)	-1.2 (1.8)	-0.36 (1.9)	0.63 (1.6)
P-MXHSE	-8.9 (3.8)	3.8 (4.2)	1.5 (2.5)	1.1 (2.2)
P-FOLGR	0.60 (2.1)	-0.67 (1.5)	-10.9 (1.4)	0.96 (1.1)
P-HILLS	1.7 (1.5)	0.52 (1.2)	0.97 (1.1)	-3.2 (1.0)
P-PRIVT	1.5 (0.76)	-0.66 (0.53)	0.49 (0.48)	0.61 (0.40)
P-OTHER	-0.16 (0.53)	-0.12 (0.38)	0.32 (0.34)	-0.16 (0.28)
P-OWN-1	21.7 (4.6)	-3.1 (4.1)	4.6 (2.1)	-0.89 (2.4)
P-OWN-3	-4.9 (2.4)	4.0 (2.9)	3.6 (1.8)	2.0 (2.4)
F-CHOCK	1,435 (683)	197 (461)	-6.5 (439)	-204 (350)
F-MXHSE	-1,570 (1,645)	2,571 (1,124)	-406 (1,085)	48 (855)
F-FOLGR	117 (857)	496 (629)	1,248 (571)	-270 (474)
F-HILLS	596 (624)	118 (441)	-1,235 (464)	987 (331)

(Standard Errors in Parentheses)

Table 13 (cont'd)

Independent Variable	Dependent Variable: Chain 2 Quantity Sales: Chock Full o' Nuts	Maxwell House	Folgers	Hills Brothers
F-OWN-1	COL	1,908	1,478	1,367
		(2,757)	(649)	(958)
F-OWN-3	-1,008	-1,762	1,297	COL
	(1,008)	(2,728)	(862)	
P-CHOCK	-4.3	-0.96	0.60	0.81
x F-CHOCK	(2.9)	(2.0)	(1.9)	(1.5)
P-MXHSE	5.4	-7.2	1.6	0.08
x F-MXHSE	(6.1)	(4.2)	(4.0)	(3.2)
P-FOLGR	-0.84	-2.9	-4.6	1.8
x F-FOLGR	(3.7)	(2.8)	(2.5)	(2.1)
P-HILLS	-3.2	-0.70	6.2	-1.6
x F-HILLS	(2.9)	(2.0)	(2.2)	(1.5)
P-OWN-1	1.5*	-5.5	-9.3	-5.5
x F-OWN-1	(1.4)	(10.5)	(2.9)	(4.0)
P-OWN-3	4.0	7.1	-5.2	-1.2*
x F-OWN-3	(4.6)	(10.1)	(3.6)	(0.90)
D-CHOCK	-1,199	-106	77	53
	(560)	(404)	(376)	(316)
D-MXHSE	2,964	860	762	-428
	(1,838)	(977)	(949)	(747)
D-FOLGR	-1,128	-339	-3,086	431
	(1,011)	(733)	(728)	(555)
D-HILLS	-364	12	679	-1,589
	(673)	(462)	(435)	(345)
D-OWN-1	N/A	-1,338	COL	-1,236
		(2,767)		(857)
D-OWN-3	-504	1,138	COL	2,255
	(822)	(2,438)		(8,608)
P-CHOCK	5.3	0.33	-0.10	-0.17
x D-CHOCK	(2.4)	(1.7)	(1.6)	(1.3)
P-MXHSE	-11.4	-3.6	-2.7	1.4
x D-MXHSE	(7.2)	(3.8)	(3.7)	(2.9)
P-FOLGR	4.7	2.2	16.4	-2.6
x D-FOLGR	(4.8)	(3.5)	(3.5)	(2.7)
P-HILLS	2.1	-0.04	-3.3	7.4
x D-HILLS	(3.0)	(2.1)	(2.0)	(1.5)

(Standard Errors in Parentheses)
N/A: Not Applicable, No Display Activity for Chock Full o' Nuts at Chain 1.
COL or *: Variable is Perfectly Collinear with Other Variable(s).

Table 13 (cont'd)

Independent Variable	Dependent Variable: Chain 2 Quantity Sales: Chock Full o' Nuts	Maxwell House	Folgers	Hills Brothers
P-OWN-1 x D-OWN-1	N/A	3.3 (10.5)	1.5* (1.0)	4.5 (3.8)
P-OWN-3 x D-OWN-3	2.2 (3.7)	-4.9 (9.6)	3.1 (1.0)	-7.2 (31.2)
R^2	0.80	0.58	0.87	0.86
Number of Observations	108	108	108	108
Mean of Dependent Variable	421	148	258	186

(Standard Errors in Parentheses)
N/A: Not Applicable. No Display Activity for Chock Full o' Nuts at Chain 1.
COL or *: Variable is Perfectly Collinear with Other Variable(s).

Table 14
OWN-PRICE ELASTICITIES OF DEMAND AT CHAIN 2

Brand	Unadvertised Price Elasticity	Advertised Price Elasticity
Chock Full o' Nuts	-6.5	-8.9
Maxwell House	+6.9*	-6.0
Folgers	-10.6	-15.1
Hills Brothers	-4.2	-6.3

*The own-price coefficient was not significantly different from zero at the .60 confidence level.

Comments

David Sappington
University of Pennsylvania and
Bell Communications Research

This paper by Katz and Shapiro provides some interesting insights concerning purchasing behavior of consumers. Their empirical findings extend our knowledge of how coupons, newspaper advertising, and in-store displays provide information to consumers, and thereby affect purchase decisions. The authors provide some evidence that coupons allow successful price discrimination, and that promotions tend to attract first-time buyers.

My brief comments on this paper focus on two areas. First, I suggest a few caveats concerning interpretation of the econometric results in Section 4. Second, I discuss a few consumer protection issues that might inform future research, should the data permit.

1. Econometric Estimation

In Section 4 of their paper, Katz and Shapiro estimate demand curves for coffee. The results of such estimates are intuitively appealing, but are difficult to interpret because no formal model of either consumer or producer behavior is offered in the paper. Hence, the exact nature of the null hypothesis that is being tested in the regression equations is not apparent. Admittedly, it is a difficult task to formulate a theoretical model and link the parameters of the model to regression coefficients. Recent work by Spady (1984), though, provides keen insight as to how this task might be accomplished. His analysis is nicely suited to the problem at hand here.

There are two particular problems with the estimation techniques adopted by the present authors that warrant brief mention. First, the error terms across estimating equations appear unlikely to be independent. Generalized Least Squares, therefore, seems to be a more appropriate estimating technique than Ordinary Least Squares. Second, all of the right hand variables in the estimating equations are actually endogenous variables in a fully specified model. Hence, the possibility of simultaneous equations bias arises. It should also be mentioned that much richer lag structures than the one reported in footnote 15 should be examined. It is quite possible that current purchases depend upon past prices and anticipated future prices, particularly if firms are known to systematically offer periodic price discounts.

2. Suggestions for Additional Investigations

The detailed analysis that Katz and Shapiro offer is probably best regarded as an examination of how various promotional policies influence consumer purchasing patterns. In as much, their work is a crucial predecessor to an examination of how firms will structure their promotions to maximize profits. Such an examination appears essential to answering the important question of whether features, coupons and displays are employed in the best interest of consumers.

It has been suggested in the industrial organization literature that coupons might be used by incumbent firms to deter entry. The idea is that coupons can make credible threats by incumbents to maintain low prices after entry has occurred. It would be interesting to test whether incumbent firms tend to issue more coupons when entry is deemed more likely. A related test of interest would be whether incumbent firms increase promotional expenditures in response to entry, and whether the effect is to drive new entrants from the market. Similarly, it would be helpful to know whether entrants are afforded the same access to display and feature promotions that incumbents enjoy.

An empirical investigation of whether multiproduct firms employ coupons and other promotions differently from firms that market a single product would also be valuable. For example, do multiproduct firms alternate the brands they promote so as to limit effective competition among the different brands? Our knowledge of the strategies and performance of multiproduct firms is sorely inadequate, and might well be enhanced by a careful analysis along these lines.

It would also be interesting to examine empirically the different effects of different types of coupons. Katz and Shapiro have accurately identified functions of coupons that are printed in newspapers. Coupons that come bundled with the product purchased have different functions. With data that permits distinction among the various types of coupons, future empirical analysis should be sensitive to these distinctions.

References

Spady, Richard, "Non-Cooperative Price-Setting by Asymmetric Multiproduct Firms," Bell Communications Research Working Paper, 1984.

The Impact of Government Policy on the U.S. Cigarette Industry*

Robert H. Porter
State University of New York at Stony Brook

1. Introduction

This paper describes an empirical study of the U.S. cigarette industry from 1947 to 1982. The cigarette industry was very unstable in this period, due to several health information "shocks" and a variety of government policy interventions. This study attempts to determine the effects of these shocks and policy measures on the industry. For example, I shall examine the impact of the ban on television and radio advertising that became effective on January 1, 1971. Structural econometric methods, which estimate demand, price and advertising equations, are employed in order to measure how both producers and consumers responded to the developments in the industry for purposes of comparison with previous studies.

The plan of the paper is as follows. Section 2 contains a brief account of the U.S. cigarette industry since 1946, and the various changes it has undergone. Previous empirical work is reviewed and revised in Section 3. These earlier studies focused, for the most part, on the demand for cigarettes. A second purpose of this study is to examine in detail the decision rules of cigarette manufacturers, so that a simultaneous equation model of the industry as a whole can be estimated. Part of the revision of the earlier studies, then, is a correction for simultaneity bias in their estimates. In particular, if the actions of cigarette companies are considered, how are the estimates of changes in consumer demand in response to the health shocks affected? If, for example, cigarette companies cut prices in response to the publicity about the adverse health effects of smoking, then such a price cut should be viewed as a result of that publicity. The net effect on cigarette consumption will be less than that measured by shifts in aggregate demand. Section 4 then turns to estimation of the firms' decision rules, with emphasis on the determination of prices and advertising levels. Section 5 concludes with a list of caveats and qualifications, as well as suggestions for further research.

* I have benefited from the research assistance of Cristina Mazon and the financial support of the Federal Trade Commission. Pauline Ippolito and Therese McGuire provided helpful comments on an earlier version of this paper.

2. The Cigarette Industry Since 1945 [1]

In 1946, the Supreme Court upheld the conviction of the major cigarette companies of conspiracy in restraint of trade and of monopolization under Sections 1 and 2 of the Sherman Act. However, no structural remedies were adopted. Rather, fines of only $250,000 were assessed.

Since the 1946 decision, six firms have dominated the industry: Reynolds, Philip Morris, Brown and Williamson, American Brands, Lorrilard and Liggett. Together they have accounted for all cigarette sales, for practical purposes. Other companies have occasionally captured a negligible share of the market (less than one per cent), and imported cigarettes have not been a factor. Furthermore, cigarette sales have been dominated by a subset of the major firms. Table 1 shows the evolution of various concentraction indices over five year intervals since 1947. CR_n is the percentage market share of the n largest companies, and H is the Herfindahl index, the sum of the squared percentage market shares of the six major companies. The two largest firms have always had market shares in excess of 50%.

The entries in Table 1 would seem to indicate a stable industry. The Herfindahl index declined after 1946 to a low of 2061 in 1958, remained between 2060 and 2270 until 1976, and then increased slightly each year to its 1982 figure. What is unusual about cigarettes, however, is the instability of individual firm market shares. Table 2 lists individual firm market shares for the period covered in Table 1. This table indicates that there have been large swings in market share, particularly between 1957 and 1972. While Table 1 alone might be indicative of a stable oligopoly which is immune from entry, interfirm market share movements have been dramatic, which could be characteristic of a highly competitive industry. Whether the cigarette industry has behaved competitively is a hypothesis the empirical work will attempt to test.

Much of this instability may be attributable to government policy and health information shocks. In 1953, a report by the American Cancer Society and the British Medical Research Council concluded that mortality rates were significantly higher for smokers than for non-smokers. A 1964 Surgeon General's report, summarizing existing research, concluded that cigarette smoking causes lung cancer. These were the two most prominent reports in a succession of studies which pointed out the health problems associated with smoking. As a consequence of these studies, the federal government intervened in a number of ways. The Federal Cigarette Labelling and Advertising Act of 1965 required that the warning "Caution: Cigarette smoking may be hazardous to your health" be displayed on packages. The FCC required that one anti-smoking commercial be aired for every four pro-smoking advertisements, under the Fairness Doctrine. This campaign was in effect from July 1, 1967 through December 31, 1970.

[1] A more complete, if somewhat dated, description of the industry is given by Tennant (1971) and by Schmalensee (1972). Whitten (1979) contains a detailed account of changes in cigarette characteristics and the brand introduction strategies of the major firms.

In 1967 the FTC began its program which annually publishes the tar and nicotine content of each cigarette brand (as well as carbon monoxide levels after 1981). The Public Health Cigarette Smoking Act of 1970 strengthened the health warnings on packages, to read "Warning: The Surgeon General has determined that cigarette smoking is dangerous to your health." In addition, this Act banned all radio and television advertising as of January 1, 1971. Agreements between the companies and the FTC in 1971 and 1972 ensured that tar and nicotine contents and a "clear and conspicuous" health warning appeared in remaining advertising messages. In addition, there have been a number of federal, state and local measures which restricted areas in which people could smoke.

Further, federal, state, county and city governments have imposed per pack excise taxes on cigarettes. Federal taxes were 7¢ per pack from November 1942 through October 1951, and 8¢ per pack until the end of 1982. State taxes ranged from 2¢ to 25¢ per pack in 1982, with an average levy of 13.1¢ per pack.[2] County and City taxes varied from 1¢ to 15¢ per pack in 1982, with an average levy of 0.56¢ per pack for the U.S. as a whole. Figure 1 plots the variable TAX, the sum of federal, state and local excise taxes collected, divided by total U.S. cigarette consumption and converted to 1967 dollars. (The units of TAX are 1967 cents per pack.) Evidently, real taxes increased until 1969 but declined thereafter, as inflation outstripped excise tax increases. According to the Tobacco Institute (1983), state and federal taxes as a percentage of average retail price peaked at 51.4% in fiscal year 1965, and declined to 26.8% in fiscal year 1982. While cigarettes are very heavily taxed, the real burden decreased after 1969. (That trend was reversed in 1983, when federal taxes increased to 16¢ per pack.)

What has been the impact of these policies and shocks on the industry? Before turning to econometric evidence, we examine the trends of the variables displayed in Figures 2 through 9. Figure 2 plots C, per capita cigarette consumption, for the noninstitutional population 16 and older, in packs per year. (A list of variables appears in Table 3.) It experienced a drop in 1954-55, then grew until 1963, after which it has declined, although not smoothly. PTC in Figure 3 is the percentage of tobacco consumed as cigarettes (as opposed to hand-rolled cigarettes, cigars or pipe tobacco). This percentage has increased since 1947, with temporary drops in 1954-55, 1964 and 1969-70. Figures 4 and 5 show F and L, the percentage market shares of filter and low tar (15 mgs. or less) cigarettes, respectively. Filter cigarettes appeared in 1953, and eventually dominated the market. Low tar cigarettes appeared in 1964, but did not capture 5% of the market until 1972. Their share then increased rapidly until 1981, and declined slightly in 1982. (This decline continued in 1983.) TPC, the weight in pounds of tobacco per cigarette consumed, is displayed in Figure 6. This has declined from 1953 to 1982 by 38%, as manufacturers put less tobacco (and more air, since tobacco was less tightly packed) in individual cigarettes. TPC

[2] Total gross state collections divided by U.S. cigarette sales. The exception is New Hampshire, which imposes an ad valorem tax.

is very highly correlated with average tar and nicotine content per cigarette, when these figures are available. A retail price index for cigarettes, P in Figure 7, is converted into 1967 dollars, and P equals 100 in 1967. Roughly speaking, real retail prices increased until 1972 and declined after that. Of course, much of the variation in retail prices is attributable to real tax rate movements.

Figure 8 shows ADV, real advertising expenditures. These grew until the ban, fell for a few years, and have since grown to record highs. The industry advertising/sales ratio averaged 5% from 1955 through 1967, but it has fallen to approximately 3% in 1981. At their peak in 1967, radio and television advertising accounted for 78.5% of total cigarette advertising expenditures (and 69.1% in 1970). Thus the ban prohibited access to the advertising media favored by cigarette companies, and almost certainly reduced the effectiveness of advertising expenditures thereafter.

Finally, T, per capita consumption of tobacco (for the noninstitutional population 16 and older) is graphed in Figure 9. Unlike per capita cigarette consumption, T peaked in 1952 and has declined ever since. To the extent that the demand for cigarettes is a reflection of the demand for tobacco, this is the most relevant consumption series. As discussed in the next section, C may not have declined as much as T because TPC, tobacco per cigarette, fell, and so consumers increased their consumption of a relative to T in order to maintain their desired tobacco consumption levels.

Another notable statistic is the number of cigarette brands being marketed. This series is available since 1967, when the FTC tar and nicotine measurement program began. For example, 134 brands were marketed in 1972, and 208 in 1982. While the number of brands has increased dramatically, the majority have not been successful, in that few have captured a market share of 0.5% or more.[3] To some extent, this increase in brands is illusory, as many new brands simply employ existing trademarks (*e.g.*, Camel "Lights").

In summary, the following changes in the cigarette industry are evident from the raw data series displayed in Figures 1 through 9. Cigarette consumption per capita has fallen 10.7% from 1963 to 1982, although not as much as tobacco consumption per capita, which fell 33.2% in the same time period and 37.5% from 1953 to 1982. Producers have reduced the amount of tobacco per cigarette (and so tar and nicotine contents), and filter and low-tar cigarettes have come into prominence. Real advertising expenditures fell temporarily after the ban. Perhaps more importantly, they have been diverted into less effective media. Real retail prices have fallen since 1972, but so have real excise taxes.

In order to obtain quantitative rather than qualitative estimates of these movements, and in order to infer causality, structural econometric techniques are employed in the following sections.

[3] *See*, for example, the annual industry studies in *Business Week*, which tabulate market shares for the leading brands.

3. The Demand for Cigarettes

This section summarizes previous research on the demand for cigarettes and, with some modification, applies the methodology of that research to the post-1946 sample. The principal modification is the use of simultaneous equations techniques, so that the results presented in this section consist of a subset of a system of equations describing behavior in the cigarette industry. While this section concentrates on consumer demand, the next section examines firm behavior, particularly pricing and advertising rules.

The data employed are annual observations from 1947 to 1982 inclusive. Besides data availability considerations, the sample period was selected to follow the 1946 Supreme Court decision, and to precede the disruptions associated with the 1983 federal excise tax increase. In 1983, a low price generic brand was introduced by Liggett, and an effective price cut was initiated by Reynolds which offered 25 cigarettes for the price of 20 (their Century brand). Although there were price differences between king-size and regular brands, for the 1947-82 sample the cigarette market equalibrium can be modelled as a situation where comparable brands had the same price at any location or point of sale.

A list of variables is provided in Table 3. Apart from the variables discussed in the previous section, a number of exogenous variables are available. Of particular interest are the dummy variables, which reflect the 1953 and 1964 health information shocks (DA and DB respectively), the Fairness Doctrine period of anti-smoking commercials (DF), and the radio and television advertising ban (DC). When simultaneous equations techniques are called for, I employ two-stage least squares (2SLS), which provides consistent single equation parameter estimates even when other equations in the system are misspecified, unlike three-stage least squares or full information maximum likelihood estimators. As will become apparent, specification issues figure prominently in the ensuing discussion.

A number of studies have examined various aspects of the demand for cigarettes. For example, Lewit, Coate and Grossman (1981) studied a panel data set of teenagers from 1966 to 1970. They found that, unlike adults, teenagers had responsive smoking demand and smoking participation rates with respect to price (elasticities of -1.4 and -1.2, respectively). Also, the Fairness Doctrine had a substantial negative impact on their participation rates. They cite evidence that individuals are much less likely to start smoking after age twenty-five, so that the FCC doctrine may have had a long-term impact as well. They did not present any evidence on the effect of the advertising ban, apart from noting that teenage smoking participation rates rose from 15.2% in 1970 to 15.6% in 1974, but fell to 11.7% in 1979. They attributed the 1970-74 increase to falling real retail prices.

In a study of a 1976 cross-section, Lewit and Coate (1982) estimated the adult price elasticity of demand for cigarettes to be -0.42. This elasticity had two components. The elasticity of smoking participation (-0.26) exceeds that of demand by smokers (-0.10). They also found that the decision to begin smoking regularly is price elastic for young adults, especially males. Income elasticities were significant and positive, but very small (0.08). In conjunction with the results of their earlier study with Grossman, they conclude that higher cigarette prices "appear to effect cigarette demand by affecting the decision to smoke or not rather than by causing existing smokers to reduce the amount of cigarettes they smoke" (p. 136).

Among the studies which employ aggregate annual time series data are those by Hamilton (1974), Ippolito, Murphy and Sant (1979), and Schneider, Klein and Murphy (1981). Hamilton's sample covered 1926-70. He estimated that per capita cigarette consumption fell 8.9% in 1953-70 due to the first health scare, an additional 21.1% in 1964-70 due to the second scare, and a further 20% during the Fairness Doctrine period. Since this latter effect exceeded the positive effect of advertising on consumption by a factor of six, he predicted that the advertising ban would result in an increase in consumption, because anti-smoking commercials were also eliminated. He also predicted that the ban would reduce competition in the industry, as a principal channel for rivalrous product promotion expenditures would be closed, and companies would not fully shift their broadcasting expenditures into other media.

A more recent study is that of Ippolito, Murphy and Sant (1979). In a 1926-75 sample, using Cochrane-Orcutt techniques to correct for serial correlation, they obtained the estimates:

$$\underset{}{ln\hat{C}_t} = K_t + \underset{(4.8)}{0.021\, t} - \underset{(5.6)}{0.176\, DA_t} - \underset{(4.0)}{0.035\, DB_t(t\text{-}1963)} - \underset{(4.6)}{0.811\, lnP_t} + \underset{(8.0)}{0.7\, lnY_t}$$

$$\bar{R}^2 = .986$$

where the notation corresponds to that of Table 3,[4] t refers to calendar year, *ln* indicates natural logarithm and t-statistics are displayed in parentheses. The term K_t includes a constant term and insignificant variables (DA_t(t-1952), DB_t, DC_t and a dummy variable which equals one for 1968-75). Here DB_t(t-1963) is a time counter which equals one in 1964, zero before that and four in 1967, for example.

These results indicate that the 1953 health scare lowered per capita consumption 16% permanently (since 1-exp(-0.176) = 0.16). Consumption had been growing 2.1% per annum before 1964; thereafter it fell at a rate of 1.4% per annum. The 1964 scare did not immediately depress consumption. Rather, the trend growth pattern was reversed. Neither the Fairness Doctrine nor the advertising ban had a significant effect. However, this equation does not forecast very well out of sample. An associated problem is that the estimated coefficients are not robust to changes in sample selection. When a similar equation is estimated

[4] They used total population 17 years and older to create C, and GNP rather than NNP to create Y. Neither distinction is important.

for the 1947-82 sample, with an advertising stock variable (A) included,[5] the trend and dummy variables have comparable coefficients, but the price elasticity is much smaller in absolute value and the income elasticity is negative. (*See* equation 1 in Tables 4, 5 and 6 and equation 2 in Table 5.) The estimates are virtually identical whether estimation is by ordinary least squares (OLS), 2SLS, or single equation maximum likelihood (MLE) with first-order autoregressive (AR(1)) or first-order moving average (MA(1)) errors. The problem is that the specification is somewhat arbitrary. The dummy and trend variables explain most of the variation in *ln*C, yet there is not a good economic justification for the inclusion of the trend variables.[6]

Schneider, Klein and Murphy (1981) explicitly account for the underlying decision rules of consumers. In a 1930-78 sample, they obtain the following OLS equation:[7]

$$
\begin{aligned}
\ln\hat{C}_t = {} & K + 0.462 \ln Y_t - \underset{(5.5)}{1.22} \ln P_t + \underset{(10.5)}{0.97} \ln \widehat{PTC}_t + \underset{(0.7)}{0.046} \ln A_t \\
& - \underset{(1.4)}{0.075}\, DF_t - \underset{(0.8)}{0.0021}\, F_t - \underset{(4.1)}{0.24}\, L_t - \underset{(2.0)}{1.39} \ln TPC_t
\end{aligned}
$$

$$\bar{R}^2 = 0.948 \qquad D.W. = 0.98$$

where the *ln*Y coefficient is prespecified, and obtained from a tobacco demand equation for the same sample. The variable $\ln \widehat{PTC}$ is an instrument obtained from the predicted values of a nonlinear regression of *ln*PTC on a function of income Y. It is included to capture the secular trend of consumers to switch from hand-rolled cigarettes to factory cigarettes as incomes increase, thereby increasing cigarette demand. The advertising variable is obtained by the equation $A = A_1 + .264A_2$ where A_1 is pre-ban advertising stock and A_2 is post-ban advertising stock. In both cases the stock is a distributed lag of past real advertising expenditures where a depreciation rate of ⅓ is used.[8] The coefficient of A_2 is estimated and has a t-statistic of 0.11. Schneider et al. argue that the impact of the health scares was cumulative rather than instantaneous, which is consistent with the Ippolito et al. results for the 1964 scare. The filter and low-tar market shares are intended to proxy the cumulative impact of the 1953 and 1964 scares, respectively. In addition, the TPC variable fell as a result of the health scare, and in response consumers increased their consumption of cigarettes in order to

[5] An explanation of how A is calculated is given in the discussion of Schneider, Klein and Murphy (1981). The results are similar when A is excluded.

[6] For further details of this line of criticism, *see* Schneider, Klein and Murphy (1981). *See*, however, the penultimate paragraph of this section.

[7] They employ total population age 14 and older to calculate C. Again, the distinction is not important.

[8] One problem with the advertising data is that cigarette marketing has evolved into new forms (*e.g.*, free samples, coupons and sponsorship of music and sports events) that are now a non-trival part of marketing expenses. Traditional advertising series do not include these expenses.

maintain desired tobacco consumption. However, the estimated absolute value of the elasticity of C with respect to TPC is greater than 1, which seems implausibly large.

The impact of the advertising ban was two-fold. First, real advertising expenditures fell, as indicated in Figure 9. Second, the advertising stock that was subsequently created (A_2) was much less effective than the previous advertising stock (A_1) in stimulating demand, 26.4% according to the estimated coefficient. (Two caveats are in order. First, the estimated standard error for the coefficient of A_2 is relatively large. Second, there is an implicit assumption of constant marginal effectiveness of advertising stock on demand. Presumably, this should diminish with increases in stock.) However, the coefficient of *ln*A in the consumption equation is small and insignificant, so that neither effect was estimated to be very important.

This equation is also characterized by poor out-of-sample forecasts. For example, the coefficient of L is quite large, and predicts a 57% fall in C from 1978 to 1982 as L increases from 27.5% to 58%. A fall of this magnitude did not occur. Equation 2 in Table 4 displays estimated OLS coefficients for the 1947-82 sample for white noise errors. The coefficients are not directly comparable to those of Schneider et al., in that the income elasticity is estimated, and the actual value of PTC is used. For this sample, the estimated coefficients of Y in the tobacco demand equations of Table 7 are negative (although not significantly so), which conflicts with the estimate from their sample. In general, the estimated coefficients are somewhat different, most notably the smaller (and hence more reasonable) coefficient of L, and the much smaller price elasticity.

Although the impact of the health scares and government policy on the variables DF, F, L, TPC and A are recognized, for estimation purposes they were treated as exogenous variables. In addition, price is assumed to be exogenous, and unaffected by these events. Neither of these assumptions is very satisfactory. Equation 2 of Table 6 shows 2SLS estimates of the parameters, in which F, L, TPC, A, PTC and P are treated as endogenous and replaced by instrumental variables.

A comparison of 2SLS with OLS estimates reveals some distinctions. In particular, advertising is estimated to have a much larger effect. However, so does TPC. Its coefficient is much too large in absolute value. It indicates that as TPC fell by 38%, cigarette consumption rose by a factor of 2.07.

If one examines the residuals from equation 2 in Table 6, it is apparent that this specification misses something captured in equation 1. There was a significant permanent fall in demand after 1953. Filter market share changes were not rapid enough to capture this effect. Equations 3 and 4 in Tables 4, 5 and 6 therefore include DA as an explanatory variable. (The latter also includes PC, a price index for cigars, to see whether the price of this substitute good influenced cigarette demand.) The coefficients in these equations seem more plausible, and both the adjusted R^2 and the Durbin-Watson statistic are satisfactory. (Equations 3 and 4 in Table 5 indicate that the OLS estimates are not affected very much when either AR(1) or MA(1) errors are permitted.) The estimated price and income elasticities, while small in comparison with those obtained in the other time series analyses which used pre-1947 data, are roughly consistent with the 1976 cross-sectional estimates of Lewit and Coate (1982). (Recall that the estimated impact of income on cigarette demand is two-fold. PTC captures the secular shift from hand-rolled to factory cigarettes as income increases. The slightly negative additional impact of income on C, although insignificant, may reflect the fact that cigarettes have been an inferior good since 1946, and particularly since the 1953 health scare.) The estimated effect of advertising on demand, while small, is double that obtained by Schneider, Klein and Murphy (1981).

In order to estimate the cumulative effect of policy and health scares on demand, consider equation 3 in Table 6, for example. Suppose that the changes in the filter and low tar market shares and in the amount of tobacco per cigarette are entirely attributable to the health scares. Also consider the permanent drop in demand reflected by DA. Then, from 1953 to 1982, demand for cigarettes fell only 4.7% because of the health scare. However, a similar calculation for equation 2 of Table 7 reveals that the demand for tobacco fell 40% in the same period. To some extent, cigarette demand did not fall as much because consumers appear to have increased their consumption in order to compensate for the reduced amount of tobacco per cigarette. (The elasticity of C with respect to TPC is estimated to be -0.23.) Tobacco per cigarette fell 37% from 1953 to 1982, causing an increase in cigarette demand of 11.4%. The net effect was a relatively larger decrease in tobacco demand.

An alternative explanation of these particular results is that female smoking behavior differs from that of males, and that these differences affect the empirical results. In particular, female smoking participation has been increasing since the turn of the century, and much of the "trend" in demand captured by Ippolito et al. may be associated with the growth of this segment of the market, rather than a switch from hand-rolled to factory cigarettes, especially since 1946. The negative and significant TPC coefficient may reflect female preference for cigarettes with relatively less tobacco and their growing significance in the market. In an effort to control for this effect, I re-estimated equation 3 of Table 6 and included the U.S. labor force participation rate (LFP) as an explanatory variable. Much of the movement in LFP is attributable to females entering the labor force, which may be correlated with smoking rates. However, the estimated coefficient of LFP was negative, reflecting a secular decline in C since 1946. Furthermore, the estimates of the other coefficients were quite similar. Although most were 20 per cent smaller in absolute value, all retained their original sign and approximate significance. (The estimated price elasticity increased to -0.36.) Nevertheless, the specified demand equation does not adequately proxy the distinction between female and male smoking behavior.

In addition, cigarette demand was affected by the Fairness Doctrine period and the advertising ban. The coefficient of DF in equation 3 of Table 6 indicates that demand was 5.9% lower during the Fairness Doctrine. (This calculation ignores any effects on A.) The radio and television ban affected cigarette demand by reducing the effectiveness of advertising expenditures, and by reducing the amount of advertising done by firms. If one attributes the fall in A from 1970 to 1982 entirely to the ban,[9] then demand fell 7.5% as a result. The figures for tobacco demand are similar; tobacco demand fell 7.5% during the Fairness Doctrine period, and 10% due to the advertising ban. However, the advertising ban also ended the Fairness Doctrine period of anti-smoking commercials, so that the net effect of the ban was a modest decrease in per capita cigarette and tobacco demand (1.6% and 2.5%, respectively).

The next section considers the effect of government policies and the health scares on firms' behavior.

4. The Impact on Firms' Decisions

This section examines how prices and advertising expenditures were affected by the health scares and government policy. First, I briefly review previous studies of the cigarette industry, none of which addresses these issues directly.

The best known earlier studies are those of Telser (1962) and Schmalensee (1972). Both of these focused on interbrand competition and the effect of advertising on individual market shares. Neither attempted to estimate the determinants of prices. Whitten (1979) examined brand introductions, how their success is related to advertising, and whether there was an advantage to early entry. Vernon, Rives and Naylor (1969) estimated an industry model, but concentrated on the tobacco leaf market. They did not estimate a cigarette manufacturer supply function.

Barzel (1976) examined cigarette prices and state excise taxes for a panel data set consisting of states and fiscal years, and found that cigarette prices responded more than proportionately to state tax rate variations (the estimated elasticity was 1.065, which was significantly greater than one). He concluded that this reflected unmeasured quality variations across states. Sumner (1981) showed that these results could arise in a noncompetitive industry with an isoelastic demand curve. Unfortunately, neither Barzel nor Sumner attempted to control for the effects of bootlegging between low-tax and high-tax states, nor did they devote much effort to estimating inter-state cost variations, other than a distance from North Carolina variable, so that their interpretations of the results are not necessarily correct. (Bulow and Pfleiderer (1983) argued that a correction for specification errors may alter their conclusions.)

Appelbaum (1982) attempted to estimate the degree of oligopoly power in several industries, including the cigarette industry. He used full information maximum likelihood techniques for 1947-71 annual data to estimate an industry model. This model consisted of a demand equation, a price equation and cost share equations for each of three factors (labor, capital and an intermediate input

[9] As we shall see later, this is not necessarily an innocuous assumption.

(presumably tobacco)). In order to specify marginal cost in the price equation and the cost share equations, he employed a generalized Leontief cost function. Profit maximization implies the price equation. He derived the equation $(1-\theta/E)P = MC(\overline{W},Q)$, so that a markup $(1-\theta/E)$ times price equals marginal cost, which depends on a factor price vector, $\overline{W}$, and output Q. Both θ, the degree of oligopoly power, or conjectural variation, and E, the absolute value of the price elasticity of demand, affect the markup. Here θ equals zero for a competitive industry (and so price equals marginal cost), the fractional Herfindahl index for an industry of Cournot quantity-setting firms, and one for a perfectly collusive industry (the monopoly outcome). In estimation, θ is identified both by the cross-equation parameter restrictions implied by maximizing behavior, and by the specification that θ varies linearly with factor prices.[10]

The estimated values of θ vary from 0.41 in 1947 to 0.39 in 1971. (The estimate of E is 0.62.) The estimates of θ exceed zero, as well as the Herfindahl index, in both cases significantly. Thus the industry acted more collusively than a Cournot industry would, according to these estimates. In addition, the correlation between θ and H is 0.8 in this sample (both decline monotonically), which is consistent with theories which postulate that non-cooperative schemes are easier to enforce the more concentrated is an industry. However, the sample variation in θ is very small, so these results may not be very significant.

Two caveats are worth noting. First, the demand equation is very simple (total U.S. consumption as a loglinear function of prices and GNP), and so probably misspecified. This could have an important bearing on the estimate of θ, which depends directly on the estimated demand elasticity. Second, the data set is not very rich. As we shall see, taxes are an important determinant of retail prices, and they varied significantly over 1947-71 (*see* Figure 1). Also, advertising was ignored altogether.

Nevertheless, most researchers seem to agree with Appelbaum's conclusion that the cigarette industry is not perfectly competitive (*see*, for example, Geroski (1983)). In that case, prices could respond to demand shocks. For example, a change in the price elasticity of demand would alter the optimal markup of price over marginal cost.

The analysis of this section bears some resemblance to that of Appelbaum. It uses as a starting point the fact that competitive firms should produce to equate price and marginal cost. The hypothesis of perfect competition can then be tested using an approach suggested by Rohlfs (1974). If price is regressed on a list of the determinants of marginal cost, then the residuals of this regression should be uncorrelated with exogenous variables which only affect demand, for competitive firms. As far as competitive firms are concerned, price is a sufficient statistic for exogenous demand variations when they choose quantities. One can test whether these residuals, which are essentially price-cost margins, are uncorrelated with exogenous demand variables by using simultaneous equations

[10] There is no obvious reason why this should be the case.

specification test techniques, which are reviewed by Hausman (1983).

Equation 1 in Table 8 reports the results of regressing *ln*P on several potential determinants of marginal cost. These include CONS, total U.S. consumption; F, the filter market share, as king-size cigarettes are typically more expensive than shorter cigarettes, and are more likely to have filters; TPC, as costs should fall as tobacco content does; TAX, since P is a retail price; PT, since tobacco is the principal input; and W, the real wage rate for tobacco production workers. While the fit is reasonably good, the Durbin-Watson statistic is low. Most of the explanatory power of the regression is provided by taxes and wages. The PT series is a noisy one, and costs probably depend on a distributed lag of past tobacco prices, rather than just the contemporaneous price, as much of the tobacco crop is stored for some time.

When the residuals from this regression are regressed on a complete list of exogenous variables of the system, as suggested by Hausman (1983, p. 433) the null hypothesis of no explanatory power is rejected. The R^2 of this regression is 0.338, which is 12.16 when multiplied by the number of observations (36). When compared to a chi-squared distribution with 5 degrees of freedom,[11] the null hypothesis that the specification suggested by a model of perfect competition is correct is rejected at a 5% significance level. Two caveats are important here. One is that other kinds of specification errors could be responsible for this result. The second is the absence of a cost of capital or capital stock variable. It is notable that much of the explanatory power in the regression of the residuals on the exogenous variables comes from demand side variables, notably PC, DC and DF. PC may be serving as a proxy for the true cost of tobacco used in cigarettes.[12] It is conceivable that DC and DF are correlated with movements in cost factors that were not included as explanatory variables (*e.g.*, capital costs), but not very likely.

Equations 2, 3, and 4 in Table 8 show the results of 2SLS regressions of price on cost determinants and a subset of demand factors. While neither the adjusted R^2 nor the Durbin-Watson statistic are very meaningful for 2SLS regressions, both indicate that these equations provide a better specification than that of equation 1. Note that neither the Herfindhal index (H) nor advertising capital (A) have much explanatory power. One worry here is that the price of tobacco (PT), and so also the price of cigars (PC), are not exogenous, in that tobacco used in cigarettes is a major component of total tobacco demand. Specifications with PT and PC treated as endogenous yielded similar results, and so are not reported

[11] There are 12 exogenous variables in the system, and seven explanatory variables in this equation (including the constant term). This statistic has a chi-squared distribution with $12-7=5$ degrees of freedom under the null hypothesis.

[12] When Equation 1 of Table 8 is re-estimated with PC as an additional explanatory variable, and the residuals from this regression are regressed on the vector of exogenous variables, one obtains an R^2 of 0.2748. The value of the test statistic, $36 \times 0.2748 = 9.89$, when compared to a chi-squared distribution with four degrees of freedom, also leads one to reject the null hypothesis of perfect competition at a 5% significance level.

here. (The same is true of estimation with H endogenous.)

It is also possible that prices are more responsive to federal excise tax changes than they are to state and local tax changes. To the extent that the latter are not uniform across states or municipalities, bootlegging may prevent as large a price response than if tax changes were uniform. This is in fact borne out by the data, in that the elasticity of price with respect to the real federal excise tax exceeds that with respect to real state and local taxes. However, this specification did not significantly alter the other estimated coefficients, and so is also not reported here.

What effect did the advertising ban and the health scares have on prices? Consider equation 3 in Table 8. Again suppose that movements in F and TPC are entirely attributable to the health scares. Then the effect of the health scares on price, as reflected by DA, DB, F and TPC, is negligible, a decrease of less than one percent, and not statistically significant. If one also accounts for the 4.7% fall in C, and hence CONS, obtained from demand estimates, this conclusion is not affected. The health scares seem to have had little direct impact on cigarette prices. Of course, if real excise tax rate increases from 1954 to 1970 were prompted by governmental reaction to the results of the health studies, then the large real price increases in this period were indirectly attributable to the health information shocks.

There is also little evidence of any impact of the Fairness Doctrine anti-smoking advertising campaign on prices. The estimated effect, obtained from the coefficients of DF and ADV, is a price increase of 2-3%, and this figure is not statistically significant. The estimated price increase is slightly smaller if one also accounts for the indirect effect of the Fairness Doctrine on prices via a fall in consumption.

There does appear to be some effect of the advertising ban on prices, as indicated by the significant coefficients of DC and *ln*ADV. Real advertising expenditures fell dramatically after the ban, as indicated by the ADV series depicted in Figure 8, and by the advertising regressions in Table 9, which are discussed in more detail below. The estimated effect on prices appears to have been an increase of 3-6%. Since specification issues are a concern, I employed a number of different specifications of the price equations. While the estimated coefficients vary significantly from one equation to another, the conclusion about the impact on prices of the advertising ban is fairly robust, in terms of sign, magnitude and significance.

This conclusion is consistent with a theoretical model in which advertising generates information about products, and thereby facilitates entry of new brands. Under this theory, the effect of the cigarette advertising ban would be to create barriers to entry, and so to solidify or magnify any monopoly power wielded by existing successful brands. Hence prices would increase as noncompetitive firms increased their markups.

One alternative explanation might be that the advertising ban coincided with a change in the characteristics of a typical smoker, as the cumulative impact of the health scares and the Fairness Doctrine was realized. The results of Lewit, Coate and Grossman (1981, 1982) suggest that much of the decrease in cigarette demand was caused by reductions in the smoking participation rate, as opposed to reductions in cigarette consumption by smokers. If those smokers who quit had a more price elastic demand for cigarettes, for whatever reason, then the aggregate demand curve would not only decrease but also become more inelastic as the quitters left the market. Then optimal markups would have increased. I attempted to test whether this occured by reestimating equation 3 in Table 6 with an additional DCx*ln*P term, to allow for a change in aggregate price elasticity coincident with the advertising ban. The coefficient of this term, although positive as predicted, was small and insignificant, so that this explanation is probably unimportant.

Another possible explanation is that the advertising ban precluded the use of the most efficient advertising technology, and so firms switched to a higher cost technology (print, billboards, etc.). In a competitive equilibrium, the price of cigarettes could have increased, if one views the product as a bundle of cigarettes and advertising services, since the marginal costs of supplying the bundle increased. To some extent, the inclusion of ADV or A as explanatory variables captures this effect, but a price of advertising services variable would be more appropriate.

We now turn to the determinants of advertising expenditures and the estimated impact of various shocks on them. Estimated advertising expediture equations are displayed in Table 9. TV is the percentage of households owning a television. Premultiplying this variable by 1-DC creates a variable that falls from 93.5 in 1970 to zero in 1971. Hence, (1-DC) x TV captures both the growing attractiveness of television as an advertising medium until 1970 and, together with DC, the effect of the advertising ban. Other explanatory variables include CONS, to capture the extent that advertising expenditures follow total consumption; F and L, in case cigarette companies advertise more to promote new cigarette characteristics; H, to capture the influence of market concentration; and the various demand dummy variables. Also included are price P or M, the price residuals from equation 1 of Table 8, which correspond with movements in price that cannot be explained by cost changes. Advertising expenditures might be affected by prices or margins, in that they affect the expected marginal return to advertising.

The results in Table 9 should be regarded as preliminary. The fit is good, but some of the coefficients are difficult to interpret. As expected, advertising expenditures are positively influenced by total consumption, and the advertising ban had a large impact. The increases in advertising at the end of the sample might be attributable to the growth of low tar cigarette market shares, as new brands were heavily promoted. However, the large negative coefficients of P and M are implausible, unless one believes that over the sample there were changes in oligopoly power that resulted in price increases (or decreases) together with advertising decreases (increases). Such a story is consistent with the negative coefficient of H, which indicates that advertising decreases as concentration increases, although this effect is not statistically significant.

Any impact the health scares had on advertising appears to be indirect. Neither DA nor DB have significant explanatory power. However, to the extent that the health scares caused a fall in CONS, advertising expenditures fell. Similarly, L has a significant positive sign, so advertising expenditures may have increased at the end of the sample as companies promoted their new brands more than they had existing brands.

The coefficient of DF is insignificant in both equations, so that the Fairness Doctrine also apparently had no direct effect. However, CONS fell in this period, and so advertising expenditures may also have been indirectly reduced as a result.

Not surprisingly, the advertising ban had a large negative impact on advertising expenditures. These expenditures fell by 33% according to equation 1, or 41% according to equation 2. (These estimates are obtained from the coefficients of DC and (1-DC) x TV.) As noted in the previous section, the post-ban expenditures were much less effective in generating demand.

5. Summary

This section summarizes the findings of the previous sections, points out some caveats, and provides suggestions for future research.

At this point, it is legitimate to ask why I did not employ reduced form methods in the previous sections, and just regress each of the endogenous variables on the complete list of exogenous variables. There are several reasons for using a structural approach. First, I do not capture all of the exogenous factors which could conceivably affect the industry (*e.g.*, either cost of capital or capital stock), and so omitted variables problems might bias the results. Second, by estimating structural equations one estimates parameters which can be compared to those of previous studies, as well as a priori micro-economic predictions (*e.g.*, the sign and magnitude of the price elasticity of demand). For example, by estimating a demand equation one has some idea whether the results are reasonable. Also, one can determine whether the health scares affected consumers' or firms' behavior, or both. Finally, the zero-one dummy variables may inadequately proxy the effects of the health scares and government policy on consumers' perceptions and so on aggregate demand. Both Ippolito et al. (1979) and Schneider et al. (1981) found that the 1964 health scare had a cumulative, rather than instantaneous, effect on the demand for cigarettes, a conclusion supported by the estimates reported in Section 3. Nevertheless, I estimated a reduced form system for purposes of comparison. The qualitative results were very similar to those obtained from the structural equations.

In summary, the 1953 health scare resulted in a permanent drop in the level of demand. In addition, after 1953 filter cigarettes emerged and the amount of tobacco per cigarette (and so the average tar and nicotine content) fell. The net effect was a drop of 12.5% in demand. Neither prices nor advertising were affected very much.

The 1964 Surgeon General's report also had little or no effect on prices or advertising, but had a more gradual effect on demand, affecting the growth rate rather than the level. Low-tar cigarettes appeared as a result, but were not prominent until the late 1970s. The growth in the share of filters accelerated slightly, and tobacco per cigarette continued its drop.

The FCC period of anti-smoking commercials resulted in a 6% drop in demand, and a somewhat larger fall in advertising. Prices also increased 2-3%. All of these effects served to reduce total consumption, on the order of 9%.

The ban on advertising resulted in a large drop in advertising levels, and in the effectiveness of these expenditures in stimulating aggregate demand. Further, prices appear to have increased 3-6%, probably because successful brand introductions were more difficult to achieve, and so the monopoly power of existing brands increased. In other words, barriers to entry for new brands became higher. Together, these two effects served to reduce total consumption, on the order of 12%. Taking into account the end of the Fairness Doctrine period, the net effect of the ban on consumption was a modest fall of 3%.

This study does not directly measure the effect of the FTC tar and nicotine measurement and labelling program. However, by creating a readily observable metric for firms to differentiate their products, the reduction in the amount of tobacco per cigarette, and hence tar and nicotine levels, was undoubtedly partially attributable to this program. To some extent, the effect of this program may be reflected in the cumulative effect attributed to the 1964 health scare.

The quantitative conclusions of this study should be viewed with some skepticism. For example, aggregate annual data for a relatively short period (36 years) was employed. As a result, the conclusions which are most in accord with previous studies of aggregate time series or cross-sectional data are probably the most reliable. These are the demand equation estimates. Enough work has been done that the appropriateness of the specification is not at issue. Nevertheless, simultaneous equations techniques seem to be more appropriate than the single equation methods employed in previous studies. In contrast, very little work has been done before on pricing and advertising decisions in the cigarette industry (or many other industries, for that matter). As a result, these equations are best viewed as preliminary attempts to estimate these decision rules. While the functional forms are influenced by previous work by Appelbaum (1982) on pricing and Schmalensee (1972) on advertising, some of the estimated coefficients indicate that the specification is not correct. Obvious examples are the omission of a cost of capital variable and inadequate proxies (PC or PT) for the true cost of the tobacco leaf used in cigarettes.

As a result, it is appropriate to conclude that more work on the behavior of cigarette companies is needed. A time series of firm-specific data might be more useful to analyze their decision rules.

Table 1

SELECTED CONCENTRATION INDICES, 1947-82

Year	CR_1	CR_2	CR_4	H
1947	34.5	64.2	92.5	2604
1952	33.0	60.2	87.8	2318
1957	29.1	57.8	83.0	2139
1962	35.0	60.5	81.3	2268
1967	32.5	54.7	81.8	2086
1972	31.4	51.4	85.6	2079
1977	33.1	59.8	87.9	2295
1982	33.6	66.4	88.6	2543

Table 2

INDIVIDUAL MARKET SHARES, 1947-82*

Year	R	P	B	A	Lo	Li
1947	29.7	7.0	3.2	34.5	4.3	21.3
1952	27.3	9.6	6.0	33.0	6.3	18.0
1957	28.7	9.3	10.7	29.1	7.7	14.5
1962	35.0	9.4	9.3	25.6	11.0	9.8
1967	32.5	12.7	14.3	22.2	10.2	8.1
1972	31.4	20.0	17.3	16.8	8.9	5.6
1977	33.1	26.7	15.8	12.3	8.7	3.4
1982	33.6	32.9	13.4	8.8	8.6	2.9

* R = Reynolds, P = Philip Morris, B = Brown and Williamson, A = American Brands, Lo = Lorrilard, Li = Liggett. Numbers may not sum to 100.0, due to rounding. Source: Schmalensee (1972) and various issues of *Business Week*.

Table 3

LIST OF VARIABLES*

CONS	aggregate annual cigarette consumption
T	per capita tobacco consumption (noninstitutional population 16 and older)
C	per capita cigarette consumption (noninstitutional population 16 and older), in packs per year
P	retail price index of cigarettes, deflated by CPI
TPC	average annual amount of tobacco per cigarette consumed, in lbs.
F	market share of filter-tip cigarettes
L	market share of low-tar cigarettes (15 mgs. of tar or less)
PTC	percentage of tobacco consumed as cigarettes
ADV	advertising expenditures, deflated by CPI
A	stock of advertising capital
Y	per capita net national product (total population), deflated by CPI
H	Herfindahl index for cigarette sales, divided by 10^4 (Source: Schmalensee (1972) and various issues of *Business Week*.)
TAX	average federal, state and local excise tax collection per pack, deflated by CPI (Source: Tobacco Institute (1983).)
PT	price index of leaf tobacco, deflated by CPI (Source: various issues of *Annual Report on Tobacco Statistics*.)
W	average hourly earnings for tobacco industry production workers, deflated by CPI (Source: various issues of *Survey of Current Business*.)
CPT	consumer price index for tobacco products, deflated by CPI
PC	retail price index of cigars, deflated by CPI (Source: various issues of *Handbook of Labor Statistics* and *CPI Detailed Report*.)
M	price-cost margin estimate (Source: residuals from equation 1 of Table 8.)
TV	percentage of households with television
DA	= 1 from 1954 to 1982 = 0 otherwise; reflecting the 1953 American Cancer Society report
DB	= 1 from 1964 to 1982 = 0 otherwise; reflecting the 1964 Surgeon General's report
DC	= 1 from 1971 to 1982 = 0 otherwise; reflecting the TV and radio advertising ban
DF	= 1 from 1968 to 1970 = .5 in 1967 = 0 otherwise; reflecting the period of anti-smoking commercials

*The sample is annual data from 1947 to 1982. All price indices equal 100 in 1967. All data sources are those of Schneider, Klein and Murphy (1981), unless otherwise indicated.

Table 4

DEMAND EQUATIONS (ORDINARY LEAST SQUARES)*

Variable	Equation Number 1	2	3	4
Constant	-53.33 (7.05)	-15.32 (3.30)	-1.197 (2.441)	-2.547 (2.851)
*ln*P	-0.0527 (0.1174)	-0.2021 (0.1942)	-0.2709 (0.1391)	-0.2937 (0.1417)
*ln*TPC	---	-0.6684 (0.2912)	-0.1260 (0.2324)	-0.1975 (0.2456)
*ln*PTC	---	1.438 (0.430)	1.476 (0.307)	1.557 (0.320)
F.	---	-0.0010 (0.0014)	0.0024 (0.0012)	0.0025 (0.0012)
L	---	-0.0040 (0.0010)	-0.0037 (0.0007)	-0.0033 (0.0008)
*ln*A	-0.0429 (0.0240)	0.0792 (0.0203)	0.0867 (0.0145)	0.0779 (0.0174)
*ln*Y	-0.1667 (0.0963)	-0.1473 (0.1142)	-0.1386 (0.0814)	-0.0948 (0.0945)
*ln*PC	---	---	---	0.1423 (0.1541)
DA	-0.1015 (0.0172)	---	-0.1009 (0.0194)	-0.1007 (0.0194)
DBx(t-1963)	-0.0368 (0.0051)	---	---	---
DC	0.0130 (0.0382)	---	---	---
DF	-0.0039 (0.0261)	-0.0481 (0.0212)	-0.0580 (0.0152)	-0.0558 (0.0155)
t	0.0303 (0.0037)	---	---	---
$\bar{R}^2$	0.912	0.826	0.912	0.911
D.W.	1.14	0.88	1.52	1.54

*Estimated standard errors in parentheses. The dependent variable is *ln*C.

Table 5

DEMAND EQUATIONS (AR(1) OR MA(1))*

	Equation Number			
	1. AR(1)	2. MA(1)	3. AR(1)	4. MA(1)
Constant	-51.38 (6.56)	-52.62 (6.92)	-1.065 (2.121)	-1.247 (2.544)
*ln*P	-0.1089 (0.825)	-0.0626 (0.0839)	-0.2842 (0.1234)	-0.2527 (0.1469)
*ln*TPC	---	---	-0.2109 (0.1788)	-0.1859 (0.2212)
*ln*PTC	---	---	1.345 (0.290)	1.392 (0.333)
F	---	---	0.0016 (0.0010)	0.0019 (0.0012)
L	---	---	-0.0036 (0.0007)	-0.0036 (0.0007)
*ln*A	-0.0571 (0.0195)	-0.0609 (0.0197)	0.0798 (0.0156)	0.0821 (0.0167)
*ln*Y	-0.0937 (0.0829)	-0.1341 (0.0850)	-0.1037 (0.0766)	-0.1300 (0.0879)
DA	-0.0865 (0.0168)	-0.0815 (0.0173)	-0.0718 (0.0174)	-0.0804 (0.0202)
DBx(t-1963)	-0.0371 (0.0041)	-0.0371 (0.0042)	---	---
DF	---	---	-0.0426 (0.0144)	-0.0480 (0.0171)
t	0.0294 (0.0034)	0.0300 (0.0036)	---	---
$\bar{R}^2$	0.934	0.941	0.921	0.922
	ρ = 0.488 (0.145)	θ = 0.641 (0.134)	ρ = 0.470 (0.147)	θ = 0.405 (0.163)

*Estimated standard errors in parentheses.

Table 6

DEMAND EQUATIONS (TWO-STAGE LEAST SQUARES)*

	Equation Number			
Variable	1	2	3	4
Constant	-52.99 (7.76)	-11.55 (5.23)	-2.191 (3.929)	-3.253 (4.797)
*ln*P	-0.0502 (0.1205)	-0.2589 (0.2477)	-0.2774 (0.1537)	-0.2902 (0.1589)
*ln*TPC	---	-1.524 (0.535)	-0.2332 (0.4510)	-0.3192 (0.5057)
*ln*PTC	---	1.919 (0.575)	1.559 (0.367)	-1.622 (0.4051)
F	---	-0.0043 (0.0023)	0.0019 (0.0010)	0.0017 (0.0021)
L	---	-0.0058 (0.0015)	-0.0040 (0.0010)	-0.0039 (0.0011)
*ln*A	-0.1016 (0.0173)	0.1201 (0.0312)	0.0936 (0.0203)	0.0909 (0.0217)
*ln*Y	-0.1669 (0.0971)	-0.2285 (0.1424)	-0.1461 (0.0905)	-0.1300 (0.1003)
*ln*PC	---	---	---	0.0688 (0.1737)
DA	-0.1016 (0.0173)	---	-0.1010 (0.0239)	-0.0989 (0.0247)
DBx(t-1963)	-0.0365 (0.0056)	---	---	---
DC	0.0132 (0.0394)	---	---	---
DF	-0.0047 (0.0268)	-0.0614 (0.0262)	-0.0607 (0.0163)	-0.0599 (0.0166)
t	0.0301 (0.0041)	---	---	---
$\bar{R}^2$	0.912	0.767	0.911	0.908
D.W.	1.14	1.37	1.53	1.52

*Estimated standard errors in parentheses.

Table 7

TOBACCO DEMAND EQUATIONS*

	Equation Number	
	1 (OLS)	2 (2SLS)
Constant	3.436 (0.744)	3.965 (0.945)
F	-0.0015 (0.0006)	-0.0014 (0.0007)
L	-0.0048 (0.0009)	-0.0054 (0.0011)
*ln*A	0.1217 (0.0159)	0.1260 (0.0171)
*ln*Y	-0.1436 (0.0888)	-0.1078 (0.0927)
*ln*CPT	-0.3375 (0.1774)	-0.4694 (0.2202)
DA	-0.0590 (0.0223)	-0.0645 (0.0228)
DF	-0.0782 (0.0190)	-0.0778 (0.0203)
$\bar{R}^2$	0.978	0.978
D.W.	1.39	1.45

*Estimated standard errors in parentheses. CPT is a consumer tobacco retail price index.

Table 8

PRICE EQUATIONS (TWO-STAGE LEAST SQUARES)*

Variable	Equation Number			
	1	2	3	4
Constant	3.91	0.495	1.90	-0.331
	(0.90)	(2.834)	(3.53)	(3.124)
*ln*CONS	-0.078	0.087	0.172	0.039
	(0.070)	(0.145)	(0.181)	(0.160)
F	0.0008	-0.0004	-0.0002	-0.0011
	(0.0005)	(0.0015)	(0.0019)	(0.0018)
*ln*TPC	0.0460	-0.169	-0.090	-0.306
	(0.1104)	(0.313)	(0.395)	(0.380)
*ln*TAX	0.445	0.367	0.304	0.384
	(0.020)	(0.053)	(0.077)	(0.059)
*ln*PT	-0.014	-0.039	-0.080	0.057
	(0.079)	(0.092)	(0.113)	(0.145)
*ln*W	0.325	0.345	0.450	0.309
	(0.061)	(0.103)	(0.141)	(0.116)
*ln*ADV/*ln*A	---	---	-0.091	0.030
			(0.059)	(0.034)
H	---	---	-0.486	0.145
			(0.567)	(0.413)
*ln*PC	---	0.305	0.178	0.229
		(0.156)	(0.204)	(0.184)
DA	---	0.014	0.034	0.016
		(0.028)	(0.035)	(0.029)
		0.032	0.007	0.031
DB	---	0.009	0.009	0.005
		(0.015)	(0.018)	(0.016)
DC	---	0.033	-0.048	0.055
		(0.028)	(0.063)	(0.038)
DF	---	0.032	0.007	0.031
		(0.022)	(0.031)	(0.023)
$\bar{R}^2$	0.983	0.981	0.973	0.979
D.W.	1.20	1.61	1.71	1.84

*Estimated standard errors in parentheses. Equation 3 employs *ln*ADV, equation 4 uses *ln*A.

Table 9

ADVERTISING EXPENDITURE EQUATIONS
(TWO-STAGE LEAST SQUARES)*

Variable	Equation Numbers 1	2
Constant	4.680 (7.203)	-3.884 (6.171)
*ln*CONS	1.631 (0.998)	1.533 (1.077)
F	0.0019 (0.0058)	-0.0011 (0.0060)
L	0.0067 (0.0059)	.0176 (0.0036)
H	-3.321 (3.663)	-3.441 (3.953)
*ln*P/*ln*M	-2.104 (1.015)	-4.187 (2.167)
DA	0.2068 (0.2295)	0.1611 (0.2484)
DB	0.1676 (0.1064)	0.0618 (0.1028)
DC	0.2823 (0.4944)	0.0703 (0.5182)
$(1\text{-DC})^4$TV	0.0071 (0.0045)	0.0062 (0.0048)
DF	0.0158 (0.1352)	-0.0856 (0.1317)
$\bar{R}^2$	0.958	0.952
D.W.	1.23	1.28

*Estimated standard errors in parentheses. Equation 1 employs *ln*P; equation 2 uses *ln*M, the residuals from equation 1 in Table 8. The dependent variable is *ln*ADV.

Figure 1

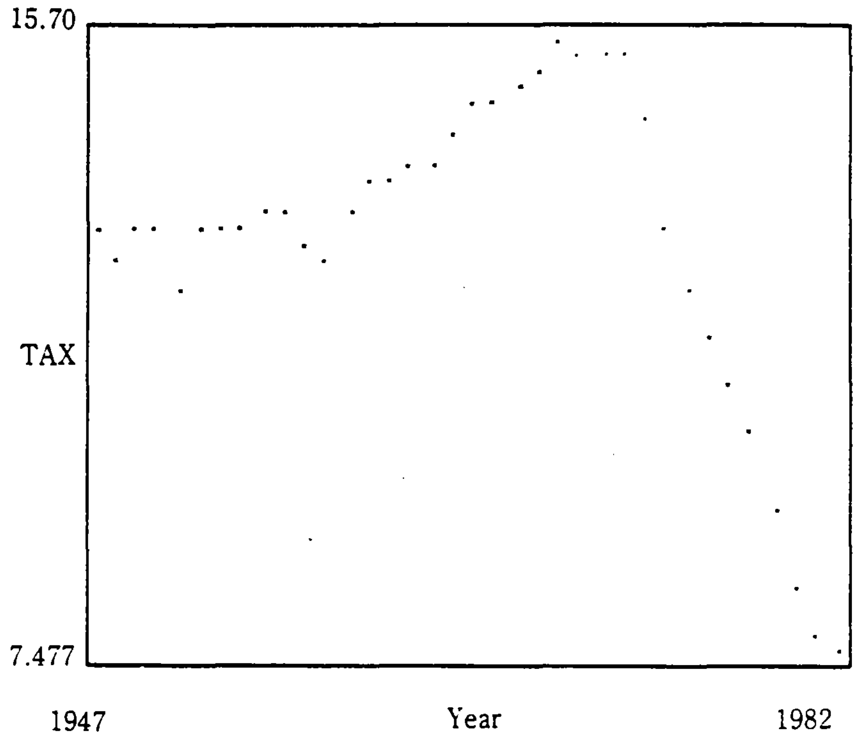


Figure 2

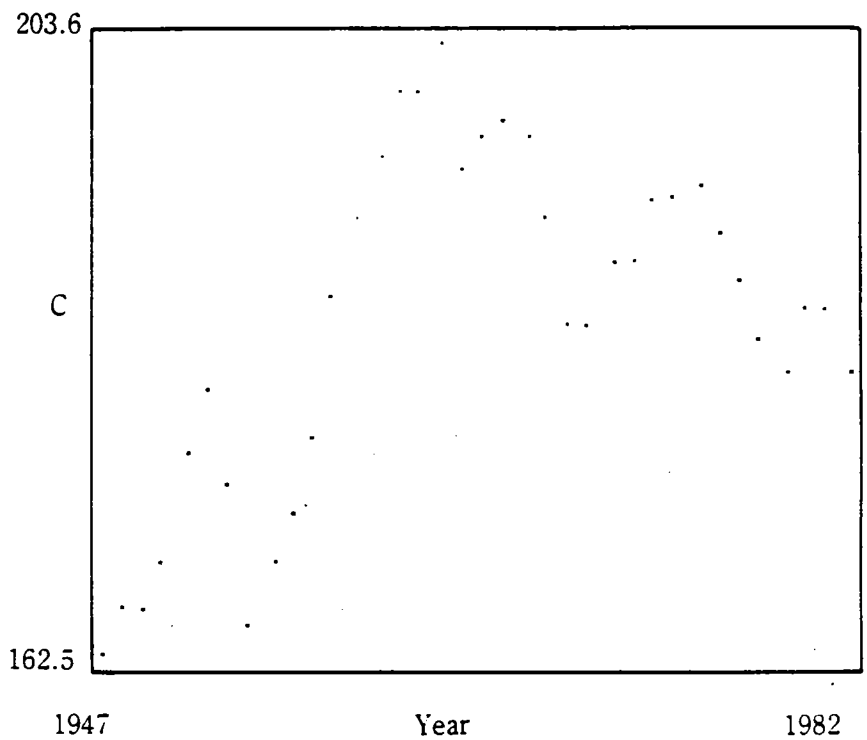


Figure 3

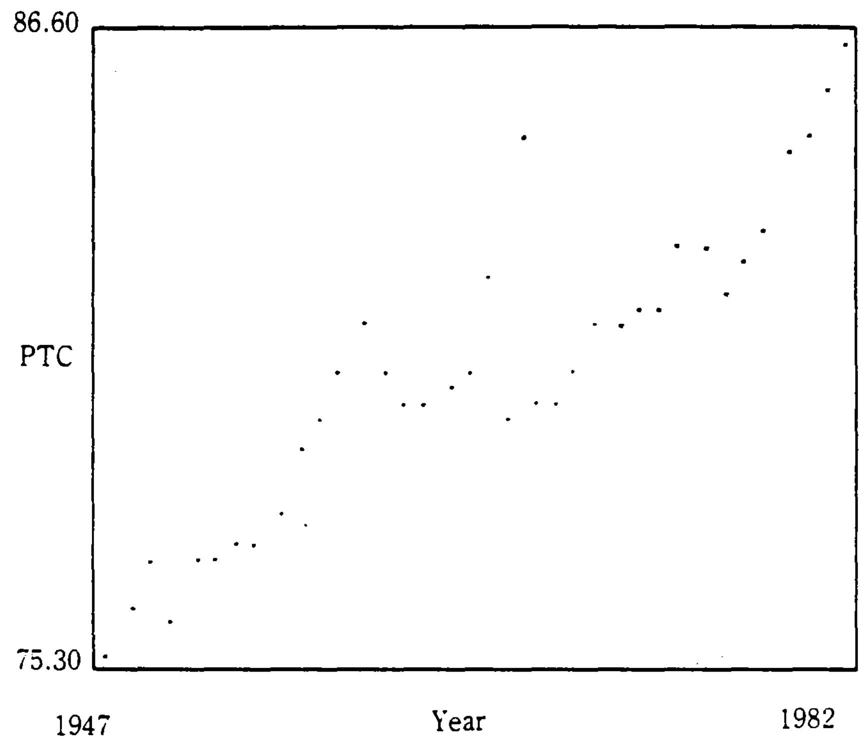


Figure 4

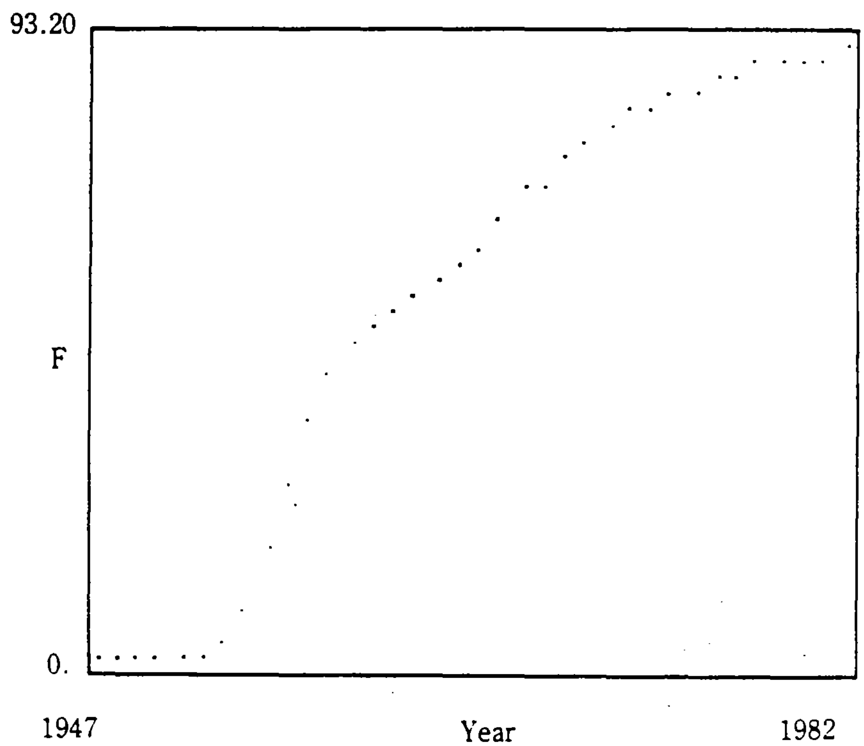


Figure 5

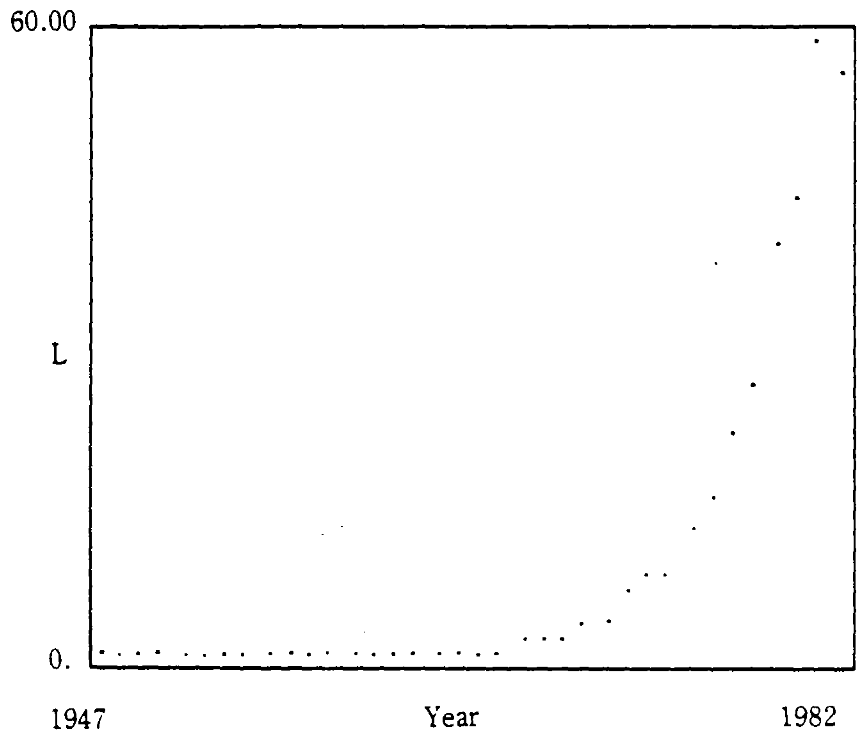


Figure 6

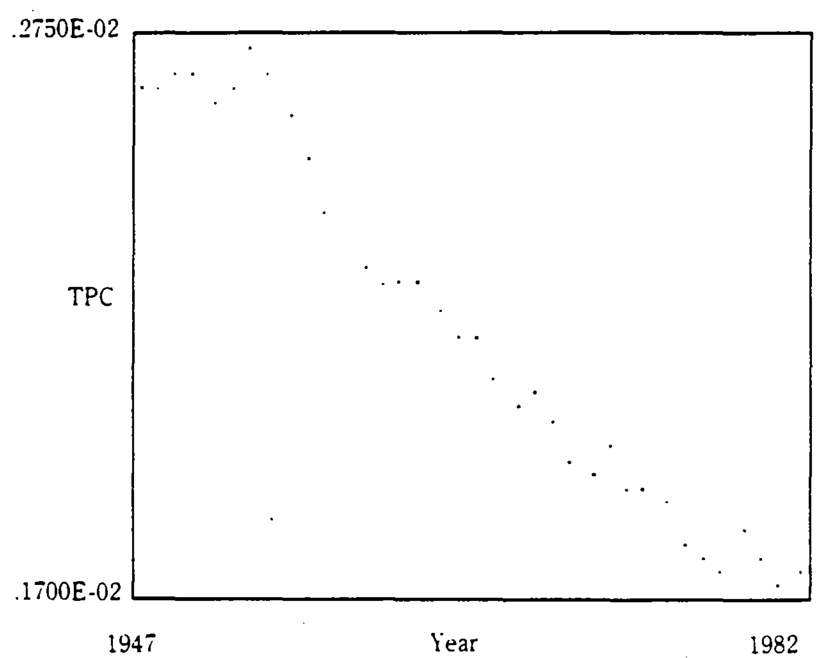


Figure 7

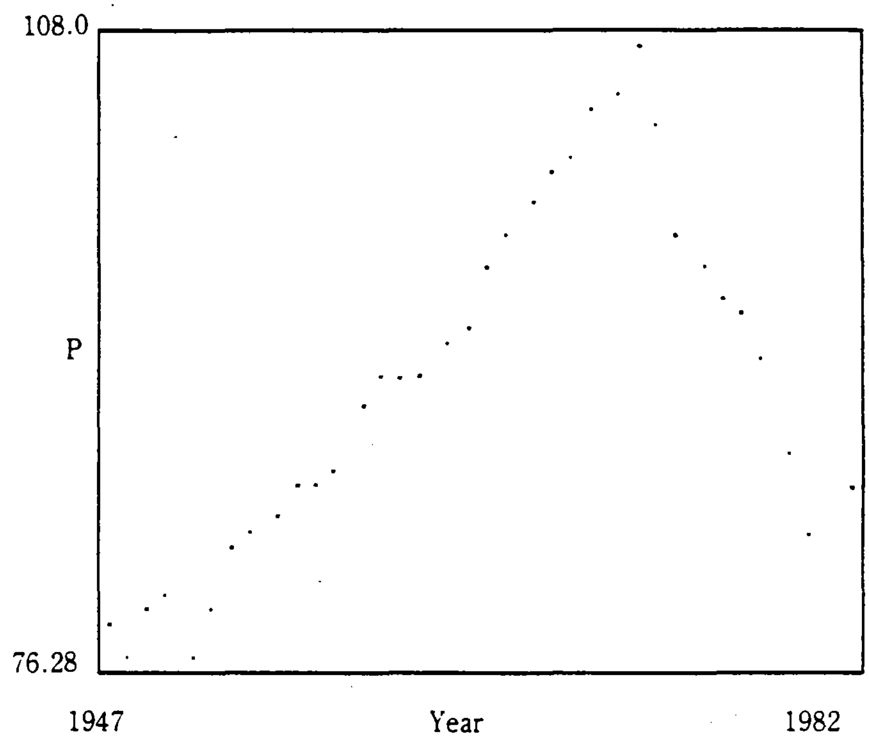


Figure 8

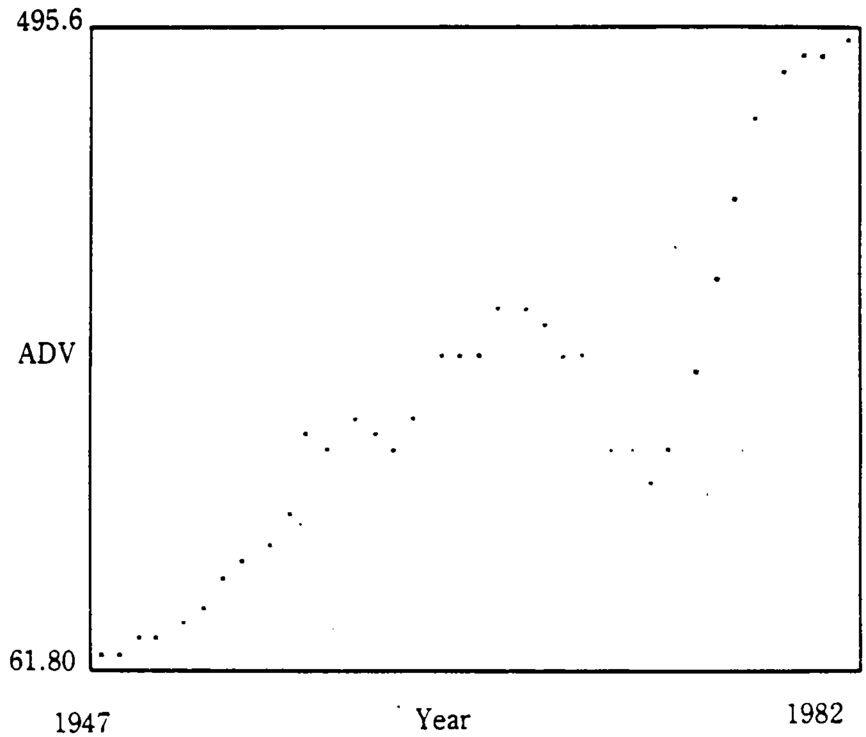
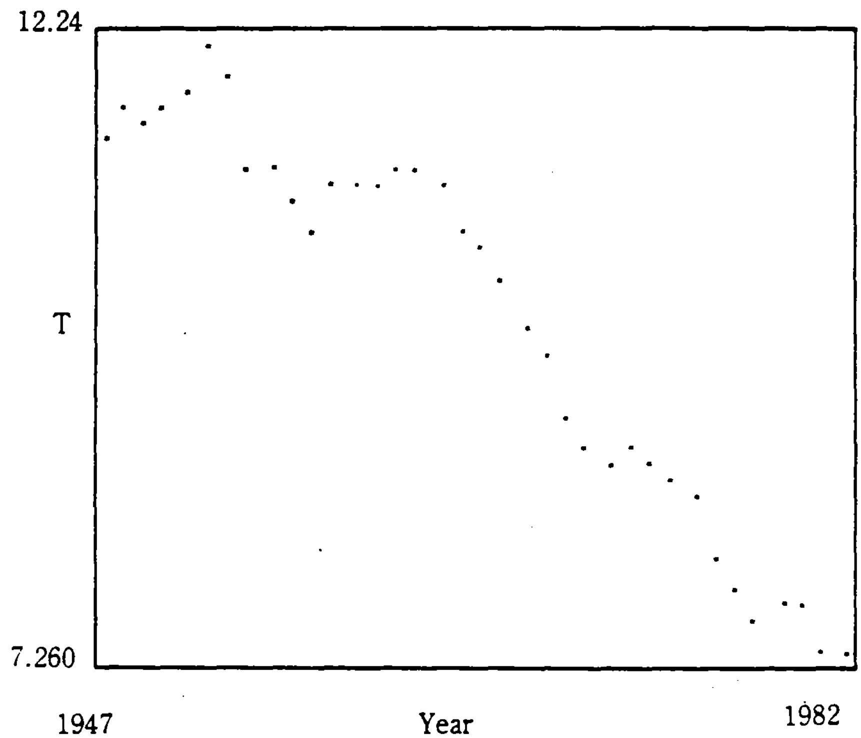


Figure 9



References

Appelbaum, Elie, "The estimation of the degree of oligopoly power," *Journal of Econometrics* 19 (1982), 287-299.

Barzel, Yoram, "An alternative approach to the analysis of taxation," *Journal of Political Economy* 84 (1976),1177-1198.

Bulow, Jeremy I., and Pfleiderer, Paul, "A note on the effect of cost changes on prices," *Journal of Political Economy* 91 (1983), 182-185.

Geroski, Paul A., "Some reflections on the theory and application of concentration indices," *International Journal of Industrial Organization* 1 (1983), 79-94.

Hamilton, James L., "The demand for cigarettes: Advertising, the health scare and the cigarette advertising ban," *Review of Economics and Statistics* 54 (1974), 401-411.

Hausman, Jerry A., "Specification and estimation of simultaneous equation models," Chapter 7 in Z. Griliches and M. Intriligator, eds., *Handbook of Econometrics*, Volume 1, Amsterdam: North-Holland (1983), 391-448.

Ippolito, Richard A., Murphy, Dennis R., and Sant, Donald, *Consumer Responses to Cigarette Health Information*, Federal Trade Commission Staff Report, Washington, D.C., August 1979.

Lewit, Eugene M., and Coate, Douglas, "The potential for using excise taxes to reduce smoking," *Journal of Health Economics* 1 (1982) 121-145.

Lewit, Eugene M.; Coate, Douglas; and Grossman, Michael, "The effects of government regulation on teenage smoking," *Journal of Law and Economics* 24 (1981), 545-569.

Rohlfs, Jeffrey, "Econometric analysis of supply in concentrated markets," *International Economic Review* 15 (1974),69-74.

Schmalensee, Richard, *The Economics of Advertising*, Amsterdam: North-Holland, 1972.

Schneider, Lynne; Klein, Benjamin; and Murphy, Kevin M., "Government regulation of cigarette health information," *Journal of Law and Economics* 24 (1981), 575-612.

Sumner, Daniel A., "Measurement of monopoly behavior: An application to the cigarette industry," *Journal of Political Economy* 89 (1981), 1010-1019.

Telser, Lester G., "Advertising and cigarettes," *Journal of Political Economy* 70 (1962), 471-499.

Tennant, Richard B., "The cigarette industry," Chapter 7 in W. Adams, ed., *The Structure of American Industry*, 4th edition, New York: Macmillan (1971), 216-255.

Tobacco Institute, *The Tax Burden on Tobacco: Historical Compilation*, Volume 18, Washington, D.C., 1983.

Vernon, John M.; Rives, Norfleet W., Jr.; and Naylor, Thomas H.,"An econometric model of the tobacco industry," *Review of Economics and Statistics* 49 (1969), 149-158.

Whitten, Ira Taylor, *Brand Performance in the Cigarette Industry and the Advantage of Early Entry,* Federal Trade Commission Staff Report, Washington, D.C. June 1979.

Comments

David T. Scheffman
Federal Trade Commission

Porter attempts to estimate the structural effects of various government policies designed to alter the flow of information and advertising received by consumers about cigarettes. Using simultaneous equations techniques Porter provides more refined estimates for some of the models of cigarette demand that have appeared in the literature. This is a useful contribution. As is usually the case for comparisons of OLS and 2SLS estimates, significant differences do not arise. In the second part of the paper some tentative beginnings are made toward modelling the effects of the government policies on the "supply" side of the cigarette industry. Some testing of industry conduct is performed, and an attempt is made to determine the effects of the policy changes on advertising. Perhaps the main contribution of the paper is the attempt to quantify the separate effects of price changes and government policies on the amount of cigarette and tobacco consumption.

Although the paper contributes to our understanding of the effects of government policies on the cigarette industry, I am surprised by all the obvious avenues that Porter did not pursue in his research. The policies studied in this paper provide a particularly rich menu of information remedies: third party disclosure, government disclosure, mandatory producer disclosure, "corrective" advertising, and a ban on radio and TV advertising. It would be very interesting to compare the relative effectiveness of these remedies on demand. To some extent this is possible by comparing the coefficients on the policy dummy variables in the Porter demand equations.

Unfortunately, the modelling of the policy variables as simply dummy variables does not feature enough of the structure of the policies to yield very interesting comparisons. In particular, the demand equation specifications assume particular types of responses to the policies when one of the interesting issues to be resolved is what are the characteristics of the responses? For example, Porter follows Ippolito, Murphy and Sant in specifying the form of the effects of the American Cancer Society and Surgeon General's Reports. This does not allow testing of the hypothesis that the effects of affirmative disclosure remedies are perishable. The defects in the modelling of the policy variables are to some extent intertwined with the simplistic modelling of advertising used throughout the paper.

The richest vein of potential research left untapped in this paper lies in the effects of the policies on the supply side of the market. In attempting to assess the effects of the policies on industry advertising strategies, the aggregation of advertising flows or stock expenditures is particularly undesirable because of the differential effect of the Fairness Doctrine and advertising ban on TV and radio advertising. Certainly, the Fairness Doctrine cries out for a cost-of-advertising approach.

The shocks experienced by this industry should make it particularly amenable to an empirical testing of market conduct. However, to begin such testing it is important not to model industry demand so that the price elasticity is assumed to be constant since one of the revealing characteristics about industry conduct is whether pricing is influenced by demand elasticity, *ceteris paribus*. Porter indicates that some testing of this assumption was done, but the assumption should be retested in a demand equation in which the policy variables, advertising, and the effects of filter and low tar cigarettes are modelled in a more satisfactory fashion.

In order to test for market conduct, some theoretical analysis of the effects of the policies under competition and imperfect competition is required. Explicit modelling of the industry's advertising and product introduction strategies is not too difficult and is amenable to empirical testing. Perhaps the most important question to be answered here is whether the increase in new product introductions is a response arising from competition or from imperfect competitors dissipating monopoly rents. This has an important bearing on the basic consumer protection policy issue arising in the supply side of the market: what are the supply side costs of the government policies?

To sum up, although this paper is of interest, I hope that Porter will pursue further research on the effects of government policies on the cigarette industry. The major questions are still to be answered.