



Office of the Chairman

UNITED STATES OF AMERICA
FEDERAL TRADE COMMISSION
WASHINGTON, D.C. 20580

August 25, 2003

James E. Wells
Director
Natural Resources & Environment
U.S. General Accounting Office
441 G. St. N.W.
Washington, DC 20548

Dear Mr. Wells:

This letter submits the preliminary view of the Federal Trade Commission ("Commission") on the General Accounting Office ("GAO") report entitled "Effects of Mergers and Market Concentration in the U.S. Petroleum Industry in the 1990s" ("Report"). Our response consists of this letter and the three enclosures. The Report purports to examine the effects of recent mergers on several aspects of the petroleum industry, including wholesale gasoline prices, concentration, vertical integration, and barriers to entry. We understand that the GAO will publish it shortly. Regrettably, the Commission has had only a limited opportunity to review the report, receiving it in early August just a few weeks before the Report was to be published.¹

The subject of this Report is important and timely, and warrants careful and reliable analysis for Congress to be able to make informed policy determinations. Unfortunately, the Report in its present form is so flawed that reliable judgments cannot be formed regarding the competitive effects of mergers in the petroleum industry. These flaws include:

- Methodological mistakes that make the Report's quantitative analyses wholly unreliable. For example, the Report does not use obvious controls for isolating the effect of a merger. It does not properly compare supposedly affected areas

¹ Commission staff were not provided with a draft of the Report until August 4, 2003 and were not permitted to make or retain copies of the Report, despite the fact that it is roughly 200 pages long and includes complex econometric analyses that took the GAO most of a year to complete. The Commissioners received copies on August 12 but were not allowed to share these copies with the FTC staff. GAO also declined to provide the Commission with the underlying data used for the Report, and did not supply it with a detailed description of the Report's final methodology. It has therefore effectively been impossible for the Commission and its staff to analyze, replicate, or test fully the Report's methodology. A more detailed commentary on the econometric analysis in the report is attached at Enclosure 1.

with unaffected areas. It also does not include non-merger factors that almost certainly will affect price, like seasonality, supply disruptions, and temperature.

- Critical factual assumptions that are both unstated and unjustified. For example, the Report simply assumes that state boundaries delimit meaningful geographic markets – an assumption that in most cases is devoid, to our knowledge, of any empirical basis or support. These assumptions often then are combined in the Report with further methodological flaws that do not meaningfully distinguish correlation from causation.
- Conclusions that lack any quantitative foundation. For example, there appears to be no quantitative basis for the Report's conclusion that unbranded gasoline has become less available. At the same time, the Report makes no effort to assess the (major) regional differences regarding the availability of unbranded gasoline, making the Report's treatment of regional differences inconsistent as well as arbitrary.

The Commission has spent significant resources investigating consummated mergers, both to determine whether past enforcement actions were correct, and to identify anticompetitive mergers the effects of which could be attenuated by future Agency action. As a result, we have accumulated substantial methodological expertise and have applied that expertise to the oil industry as part of our enforcement mission. Based on this expertise and our initial review of the analyses in the report, we find that the event study and the price-concentration regression are fundamentally flawed.²

I. THE REPORT'S ECONOMETRIC ANALYSES ARE FUNDAMENTALLY FLAWED

The heart of the Report consists of two econometric analyses. The first performs what is sometimes called an "event study." The analysis attempts to isolate the impact of eight petroleum mergers that occurred in the late 1990s on the price of wholesale gasoline (adjusting for crude oil costs). The results purport to show that six of the eight mergers in question were

² The Commission staff previously provided GAO staff with preliminary oral and written comments in December 2002 on issues, among others, relating to the data and methodology encompassed in the Report's econometric analyses. (See Enclosure 2 to this letter.) GAO staff appears to have ignored most of the comments provided by Commission staff about the basic methodology that the GAO staff proposed to use at that time. GAO staff also failed to apprise Commission staff of methodological changes made subsequent to December 2002, which changes further undermine the reliability of the conclusions in the Report's econometric analyses. FTC staff also forwarded the attached report by Professor John Geweke on "Empirical Evidence on the Competitive Effects of Mergers in the Gasoline Industry," unpublished draft, July 16, 2003, to GAO staff. (See Enclosure 3) Professor Geweke is one of the most widely respected econometricians in the United States.

associated with statistically significant price increases, ranging from less than one cent per gallon to over five cents per gallon for at least one type of gasoline. (According to the event study, the other two mergers were associated with price decreases.) The second econometric analysis contained in the Report seeks to describe the relationship between wholesale gasoline prices (adjusting for crude oil costs) and wholesale concentration measured at the state level. This price-concentration analysis purports to show a positive and significant relationship between higher wholesale gasoline prices and industry concentration. Depending upon a variety of factors, such as fuel type and region of the country, the analysis estimates that an increase in the Herfindahl-Hirschman Index (HHI) of 100 points may lead to an increase in the price of wholesale gasoline of as much as four cents per gallon.

Based on our initial review of these econometric analyses, the methodologies underlying both analyses are fundamentally flawed. Five primary reasons support this conclusion.

First, the models used do not control for the many factors that could cause prices to increase. Isolating the effect on price from a merger necessarily requires the correct and comprehensive identification of factors that might influence demand (seasonality, temperature, income) as well as those that might influence supply (supply disruptions, changes in gasoline formulation). The Report is conspicuous in its failure to control for *any* of these factors. For example, the period at issue was characterized by several supply shortages, which can cause short-term price spikes entirely unrelated to the mergers under investigation. Similarly, seasonality is a crucial factor in analyzing this market: gasoline prices tend to increase in the summer in response to increased demand. Not controlling for seasonal effects is especially problematic because in some cases the post-merger period contains only a short time period, encompassing just one season, while the pre-merger period includes at least an entire year. By not controlling for such factors, the Report fails to provide meaningful information regarding whether price changes were merger-related or not.

An approach superior to that of the study would be to compare price changes in the affected markets with price changes in carefully selected comparable non-merger markets. If post-merger prices in non-merger markets went up as much as those in the merger market, then there is no rational basis for concluding that the merger caused the price increases. No such carefully defined "natural experiment" was conducted by GAO staff; at least none is included in the Report. Instead, the preferred estimates in the Report simply compare the post-merger prices in the areas affected by the merger to the pre-merger prices in those areas.

Second, the price-concentration methodology used by the GAO is subject to several well known problems that make it unacceptable as an alternative to a well-conducted event study. The most important of these problems is the difficulty in distinguishing between correlation and causation. Simply because two factors move together does not mean that one caused the other.

Third, any reliable price-concentration analysis necessarily requires that concentration be calculated in an economically well-defined market – that is, an area in which a particular merger or other increase in concentration is likely to have an economic effect. The Report's assumptions

of state-wide geographic markets are unjustified. We are not aware of any supporting empirical data that markets generally coincide with state boundaries. Indeed, all of the data with which we are familiar point to the conclusion that wholesale markets in this industry rarely coincide with state limits. Accordingly, while price-concentration analyses may provide some useful information on general industry trends in concentration, they cannot be used to determine if an economically meaningful relationship exists between price and concentration.

Fourth, the results in the Report are, in many cases, not robust. Economists usually consider various approaches to estimating a model to determine whether the results from one approach (or “specification”) are consistent with those using alternative approaches. This procedure is known as checking the “robustness” of the results. If results differ substantially for different methods or approaches, the reliability of the results is questionable. The Report’s results in fact differ substantially across models. For example, in some cases, when estimating the effect of a particular merger on wholesale gasoline prices, the Report finds positive effects with some specifications and negative effects in other specifications.

Finally, documentation of the technical work on the econometric models is incomplete. For example, there is no discussion regarding how divestitures were treated, or which terminal racks were used in which regressions, or how price observations were constructed. The Report’s approach to these issues, had they been better documented, might raise more concerns about the methodology used to reach the reported results. Moreover, given the failure to provide the underlying data, it is impossible to replicate independently the Report’s results or to perform more rigorous robustness tests (including taking into account the missing factors that influence price changes as discussed above). Results that cannot be replicated or thoroughly analyzed for robustness are of little scientific value.

II. THE REPORT’S ASSERTIONS ABOUT STRUCTURAL CHANGES IN THE PETROLEUM INDUSTRY, AND THEIR COMPETITIVE EFFECT ON GASOLINE MARKETS, ARE ALSO FLAWED

In addition to the flaws in the Report’s quantitative analyses, there are several conclusions in the Report that appear to be without quantitative support. Other observations appear to overlook important factual issues, or invite unwarranted conclusions about the effect of particular facts on the extent of competition in the market.

For example, the Report suggests that the mergers in question have raised barriers to entry, while acknowledging that the effect of these mergers on entry barriers could not be quantified. The Report further observes that mergers may have made it more difficult for smaller firms to compete, or for new competitors to enter these markets. These observations, however, even if true, do not mean that competition in the petroleum industry has been harmed or eliminated. For example, to the extent that mergers confer cost-reducing scale advantages (as the Report suggests), consumers will benefit when cost savings are passed on through lower prices. Complaints from small competitors that competition with larger-scale entities is putting them in

jeopardy therefore may well suggest *enhanced* competition, as all firms feel pressured by competition to reduce costs by whatever means possible and thereby reduce retail prices to consumers. Similarly, other structural factors detailed in the Report – such as minimum volume requirements and the alleged preferences of refiners to deal with larger distributors – are as consistent with a theory of *enhanced* competition as they are with a theory that petroleum industry consolidation has adversely affected consumers.

Another finding in the Report without quantitative support is the conclusion that vertical integration between refining and marketing has increased. Characterizing the degree of vertical integration between functional levels is more complicated than the Report suggests. EIA data on volume of gasoline distributed by channel of distribution indicate that, for the nation as a whole, the sale of gasoline through independent distributors – by far the leading channel of gasoline distribution – has increased in recent years. These data indicate that overall vertical integration between marketing and refining has *not* increased in recent years. These data, and our own experience, also reveal that vertical integration between refining and marketing differs significantly across different geographic areas within the United States. Moreover, the competitive implications of vertical integration are complex, with the potential for procompetitive as well as anticompetitive effects.

Finally, the Report finds that unbranded gasoline has become less available. This conclusion appears to have no firm quantitative foundation, but is instead based on interviews with various industry participants. In fact, the availability of unbranded gasoline varies significantly across geographic areas. The Report specifically notes that hypermarkets almost always supply unbranded gasoline and are growing significantly. The success of hypermarkets and other unbranded marketers in some areas of the country raises important questions about the competitive significance of “branded gasoline” in attracting consumers, and suggests that generalizations about possible impediments to the expansion of unbranded marketers are unwarranted.

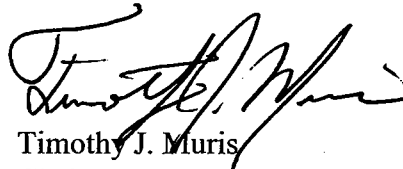
III. CONCLUSION

The Commission and its staff stand ready to provide further assistance to the GAO. The Report deals with a timely and important topic, and its findings have potentially important implications for public policy regarding petroleum mergers. In the Commission's view, however, this report does not meet the high standards of “accountability, integrity, and reliability”³ we

³ “Accountability describes the nature of GAO's work. GAO helps the Congress oversee federal programs and operations to ensure accountability to the American people.” See www.gao.gov. “Integrity describes the high standards that GAO sets for itself in the conduct of its work. GAO takes a professional, objective, fact-based, nonpartisan, nonideological, fair, and balanced approach to all of its activities. Integrity is the foundation of reputation, and GAO's approach to its work assures both.” *Id.* “Reliability describes GAO's goal for how its work is viewed by the Congress and the American public. GAO produces high quality reports,

would expect from GAO's reports and publications. We, too, are governed by similar standards for protecting the public interest.⁴ Accordingly, we remain willing and eager to assist in the production of a more accurate report.

By direction of the Commission.



Timothy J. Muris
Chairman

Enclosures

- (1) Discussion of Deficiencies in Chapter 5 of the GAO Report "Effects of Mergers and Market Concentration in the U.S. Petroleum Industry in the 1990's"
- (2) FTC Staff Comments on the GAO methodology for "Econometric Analysis of Effects of Market Concentration and Mergers on U.S. Wholesale Gasoline Prices in the 1990's," December 20, 2002
- (3) Professor John Geweke, "Empirical Evidence on the Competitive Effects of Mergers in the Gasoline Industry," unpublished draft, July 16, 2003

testimony, briefings, legal opinions, and other products and services that are timely, accurate, useful, clear and candid." *Id.*

⁴ The Federal Trade Commission Act provides that the Commission take action when it determines that such action would be in the public interest. 15 U.S.C. § 45(b). For provisions specifically addressing data quality, *see also* Data Quality Act, i.e., Treasury and General Government Appropriations Act for Fiscal 2001, Pub. L. No. 106-554, § 515 (Dec. 21, 2000); 67 Fed. Reg. 8452 (Feb. 22, 2002) (Office of Management and Budget guidance); <http://www.ftc.gov/opa/2002/08/fyi0242.htm> (FTC guidelines implementing the Data Quality Act).

Enclosure 1

August 25, 2003

Enclosure 1:
Discussion of Deficiencies in Chapter 5 of the GAO Report "Effects of Mergers and Market Concentration in the U.S. Petroleum Industry in the 1990's"¹

I. Background

These are comments of the Federal Trade Commission's (FTC) Bureau of Economics (BE) Staff on Chapter 5 and the related appendix of the Report. These comments reflect the limited ability that FTC staff had to review the draft.² Additional time would have allowed a more comprehensive set of staff comments. Nevertheless, even within the short time we have had to review the report, we have identified fundamental methodological flaws with the econometric analyses that we will discuss in detail. These fundamental flaws mean that the Report cannot provide a reliable basis for addressing the issues it claims to study.

BE staff provided six pages of comments to the GAO in a December 20, 2002 letter as well as verbal comments at a subsequent meeting with GAO staff on the draft methodology for the GAO study, dated December 2, 2002.³ Most of the FTC staff comments on the draft econometric model were not incorporated into the Report. In addition, changes in the methodology in the Report and the methodology provided in December 2002 raise additional

¹ In the memorandum, we will refer to the GAO report as the "Report."

² FTC staff was only given access to the Report on August 4, 2003 (a few weeks prior to when we understand that GAO plans to publish the report) and was not allowed to retain copies of the report.

³ FTC staff also forwarded the attached report by Professor John Geweke on "Empirical Evidence on the Competitive Effects of Mergers in the Gasoline Industry," unpublished draft, July 16, 2003, to GAO staff. Professor Geweke is one of the most widely respected econometricians in the United States.

concerns.⁴ This comment includes the comments that the FTC staff previously gave to the GAO as well as comments on the changes made from the December 2002 methodology and problems that were not evident in that methodology.

II. Summary

Chapter 5 of the Report contains two econometric analyses. The first attempts to estimate the wholesale price effects of eight petroleum mergers on conventional, reformulated (RFG), and CARB gasoline (separately for branded and unbranded). The second attempts to estimate the relationship between wholesale prices and concentration (HHI) for the three main formulations of gasoline.⁵ The Report purports to show: (1) significant price increases for at least one gasoline formulation for six of the eight mergers reviewed; and (2) a significant positive relationship between price and concentration. As discussed below, there are fundamental flaws with both the merger event analysis and the price/margin concentration analysis. Given the severity of the flaws, the results of the statistical analysis cannot be used to make reliable inferences about the price effects from the mergers analyzed or the relationship between price and concentration.

The flaws with the analyses stem both from the underlying theoretical models and the implementation of the models used by the GAO. Of key importance is the lack of controls for important factors that affect the price of gasoline. Without such controls, the analyses cannot be used to reliably isolate the impact of the mergers or concentration on price. There are also

⁴ Given the important differences between the December 2002 methodology and the methodology used in the current Report, the opinions of external reviewers may not be the same for the Report as for the December methodology. If the external reviewers have not reviewed the final Report, they should be given that opportunity or the representation of their views should state that fact.

⁵ Both of these analyses use posted rack (wholesale) pricing data, which will not reflect any discounts distributors may receive off posted prices. Gasoline is also sold to dealers at a dealer tank wagon price or is transferred to refiner owned and operated outlets at some internal transfer price. The portion of transactions that are sold on a rack basis varies considerably across the country, and as a result, any inferences about effects on retail prices from predictions regarding rack prices must be carefully qualified.

serious problems with the interpretation and characterization of the results. In several instances the results in the Report are not as robust as represented in the discussion of the results. Also, the results of some of the specifications reveal some of the underlying methodological problems. Finally, the size of the estimated price effects, both positive and negative, of the mergers is implausibly large in the context of this industry.⁶

An additional problem is that the documentation of the regressions and methodology is incomplete. The discussion in the Report should be sufficient so that a researcher with access to the same data set could replicate the results. Unfortunately, the discussion does not satisfy this basic requirement.

III. Merger Event Study Methodology

There are several generally recognized methodologies that might be used to analyze the price effects of a merger through an event study. These approaches emphasize the need to control for factors unrelated to a merger that may influence prices. Controlling for such factors is important because the goal of an event study is to isolate the effect of the event on the variable of interest such that any changes in the variable after the event can be attributed to the event. GAO's preferred model, however, excludes important control variables, and as a result, can not reliably isolate any price effect—be it positive or negative—that can be attributed to a merger.

⁶ Many of the estimated merger effects from the preferred estimations are over a nickel a gallon. When determining what is the relevant antitrust market in which to assess a proposed merger, the FTC asks the question: what is the smallest relevant market in which a hypothetical monopolist could impose a small but significant and non-transitory price increase? See U. S. Department of Justice and the Federal Trade Commission, Horizontal Merger Guidelines, April 1997. The FTC has generally applied a one cent per gallon price increase as the standard for evaluating market definition in refined petroleum products, including gasoline. While a one cent per gallon increase is smaller than that generated under the 5% price increase rule that the FTC typically uses in evaluating mergers, a one cent per gallon increase is extremely significant in this industry. The petroleum industry is characterized by large volumes and relatively thin margins. For example, data from the Energy Information Administration indicate that the net refinery margins (which reflect crude costs as well as refinery operating costs) were on the order of four cents per gallon during the late 1990's.

Among the factors not adequately controlled for in GAO's specifications are supply shocks, changes in fuel specifications, and seasonal effects. The Report's use of instrumental variables to deal with endogeneity issues is also problematic. Finally, it is well known that achieving reliable results in an event study requires that the event's pre- and post-periods be appropriately specified: unfortunately, the Report is flawed on this dimension as well.

The remainder of this section discusses: (1) general methodologies employed in event studies regarding the effects of mergers; (2) why the approach chosen by the GAO is inferior to potential alternatives; and (3) why the implementation of the approach chosen by the GAO is fundamentally flawed.

A. General Approaches in Merger Event Studies

The GAO Report correctly cites a number of published economic studies that examine the price effects of consummated mergers.⁷ In most merger event studies that examine the price of products before and after a merger, one of two types of regressions has been estimated. In the first type of regression (*see* Barton and Sherman (1984) and Kim and Singal (1993)), the price of the product affected by the merger is compared to a substitute product or the same product in another market that faces similar demand and cost conditions before and after the merger. Specifically, the analysis is a reduced form regression of the price of the product of the merged firm relative to control product(s) on various time trends and a merger dummy variable. To implement this approach for oil mergers, the dependent variable would be the price of gasoline in

⁷ Relevant economic papers include Barton, D.M., and R. Sherman, "The Price and Profit Effects of Horizontal Merger: A Case Study", *Journal of Industrial Economics*, 33(2), December 1984, pp. 165-77. Kim, E.H., and V. Singal, "Mergers and Market Power: Evidence from the Airline Industry," *American Economic Review*, 83(3), June 1993, pp. 549-69. Schumann, L., J. Reitzes, and R. Rogers, "In the Matter of Weyerhaeuser Company: The Use of a Hold-Separate Order in a Merger with Horizontal and Vertical Effects", *Journal of Regulatory Economics*, 11(3), May 1997, pp. 271-89. Schumann, L., R. Rogers, and J. Reitzes, *Case Studies of The Price Effects of Horizontal Mergers*, Federal Trade Commission, April 1992. Vita, M. and S. Sacher, "The Competitive Effects of Not-for-Profit Hospital Mergers: A Case Study," *Journal of Industrial Economics*, 49(1), March 2001, pp. 63-84.

a city where the merger reduced the number of competitors and the independent variables would include the price of gasoline in a nearby city (or set of cities) that arguably has the same supply and demand characteristics but is not affected by the merger. The choice of control cities, *i.e.*, the cities where the merger should not affect prices, has to be made carefully and should be subjected to sensitivity analysis.

In the second type of regression (*see* Schumann *et al.*, (1992) and (1997)), the price of the merged firm's product (or market price) is regressed on demand and supply/cost shifters plus a merger dummy. The researcher is trying to model how prices are determined in the markets at issue, and the merger is one of the factors potentially affecting price. This approach can be problematic due to the lack of available demand and supply variables that have sufficient variation over time and over geography to capture adequately the factors impacting price, thereby isolating the effect of the merger. This problem is particularly acute in gasoline markets because there are few variables that are available on a weekly basis at the city level to help explain rack price variation.

A third approach, used in Vita and Sacher (2001), combines elements of both of these approaches. In their study of a hospital merger, they examined the price of the merged firm relative to the price of a control group of firms unaffected by the merger that should be affected by the same demand and supply factors and regressed these relative prices on demand shifters, cost shifters, and the merger event to gauge the effect of the merger.

B. Approach Used By GAO

The methodology described in the GAO December 2002 draft suggested that the regressions would use control cities as well as supply and demand shifters. This would have been similar to Vita and Sacher (2001). The Report, however, does not use a control city methodology. This is surprising given the following quote from page 123 of the Report and the related footnote 90, "... prices at the nearest rack could influence prices at the rack city. We did not however, incorporate this variable directly in our model because there is co-movement between the nearest price variable and prices since both variables are likely to be generated by

the same set of independent variables.” Footnote 90 describes, “In fact in our preliminary estimation we found that the estimated coefficient on the nearest prices were not statistically different from 1.”

The GAO appears to use these explanations as a reason for not including control cities in the analysis. On the contrary, these findings strongly suggests that prices from nearby racks (not affected by the merger) are an important, if not the important, control variable. As the quotation from the Report suggests, the price at the nearest rack and the price at the merger rack are determined by the same independent variables, *i.e.*, the demand and supply variables. Using the price at an appropriate nearby (control) rack is particularly important because demand and supply variables that are specific to control racks are not readily available and therefore unobservable. All of these unobservable demand and supply effects are measured jointly in the control rack price. We thus believe that, assuming control racks were selected appropriately⁸ and additional supply and demand parameters were included to measure any price changes between the control racks and the merger racks unrelated to the merger as appropriate, the control rack approach would be superior. The GAO, however, chose not to use the control rack approach but rather to try to control for supply and demand factors directly.

C. Methodological Problems with the GAO’s Approach

1. GAO’s Model is Under Specified

When using the approach of controlling for supply and demand factors directly without using control cities, the independent variables in the regressions, as noted in the Report, should consist of market structure and regulatory factors, cost/supply factors, and demand factors. (P.130-31) This analysis is most similar to the approach taken in the Schumann *et al* studies. If

⁸ The control rack should not be directly or indirectly affected by the merger. Both firms should not be posting at the control rack and the control rack should not be so close to the merger rack that arbitrage is likely. A control rack would have the same demand, bulk supply and fuel specifications as the merger rack.

the GAO staff had been able to develop control variables for all the market structure, regulatory factors, cost/supply factors and demand factors, this approach would have been useful. The Report's regressions contain no measures of market structure factors, other than the merger dummy, no measures of regulatory factors, minimal cost/supply factors and no demand factors. The relatively low amount of variation explained by the regressions estimated in the Report (with R² of less than 20%) suggests that important factors explaining pricing were excluded from the regression. As a result, the approach employed by the GAO is not viable for estimating merger effects. Thus, the results are of no value for studying the effects of the mergers analyzed.⁹

The basic equation estimated by the GAO is as follows¹⁰:

$$PRack_{i,t} - Crude_t = \beta_0 + \beta_1 DMerger_{i,t} + \beta_2 InventoryRatio_{j,t} + \beta_3 Utilization_t \quad (1)$$

Prack (i,t)	= average rack price in city i week t
Crude (t)	= price of WTI at week t
Dmerger(t)	= merger dummy 0,1
InventoryRatio(j,t)	= inventory/estimated demand for PADD j week t
Utilization(t)	= national refiner utilization rate at week t

For the GAO preferred specification, the above equation is estimated using data for just the racks directly affected by the merger. In addition, the refinery utilization variable is typically dropped in the preferred specification. The merger effect, β_1 , is simply a comparison of the average price after the merger compared to the average price before the merger, controlling for the inventory and in some cases utilization. This model is under-specified. There are any

⁹ The following quotation from page 85 of the Report is incorrect with respect to the economics literature and information on the petroleum industry, and it even conflicts with other parts of the Report, demonstrated by the lack of explanatory power of the variables used in the regressions: "However, we believe that our model specification have captured all the relevant variables that could affect wholesale gasoline prices. Moreover, we believe that our economic methodology is sound and generally consistent with previous studies."

¹⁰ The preferred specification used by GAO included city fixed effects or random effects as well as a correction for autocorrelation.

number of additional supply, demand, and regulatory variables that will differ before and after the merger. We discuss these further below.

Given all of these issues, the regression estimates clearly suffer from omitted variable bias. The effects of omitted variable bias are well documented in the economics literature.¹¹ Symptoms of omitted variable bias are clearly evident in the results presented in the Report. For example, footnote 87 states, "Also, the estimates with years variable (dummy variable for years) appear unreasonable. Furthermore, we did not have a good economic reason for including the years variable." The reason for including this variable is there are omitted variables that are correlated with years. It is likely that this was a highly significant variable in the regressions and was serving as a proxy for other variables. Events correlated with years include formulation changes, supply disruptions, refinery closures and, in some parts of the country, changes in imports.¹² The effect of including the years variable can be seen in Tables 15-17. As discussed more fully below, the inclusion of the years variable changes the estimated results for the two transactions affecting CARB gasoline, Tosco-Unocal and Shell-Texaco I, from finding positive effects on prices from the merger to negative effects. While not as dramatic as the change in the CARB results, many of the other estimated mergers effects change, either to increase or decrease the effect, when the years variable is added to the regression.

The rationale provided by the GAO for dropping the years variable raises another concern. Classical statistical inference – like that used in the GAO report – assumes that the regression specification is chosen independently of the results it generates. If not, then the researcher injects his beliefs into the estimation process. The consequence of such "specification searching," or, less charitably, "regression fishing" is that the estimated coefficients are biased, as

¹¹ For a discussion of omitted variable bias and its effects see the literature review on returns to education in Ehrenberg, R.G. and Smith, R.S., *Modern Labor Economics*, Fourth Edition, Harper Collins, (1991), pp. 320--330.

¹² In California, an additional complication is the need to model or include a variable for a new CARB specification in 1996.

are the standard errors.¹³ It appears that, in at least one instance, the GAO report did indeed discard a specification because the "estimates appear unreasonable" (Footnote 87).

In what is deemed a sensitivity check, the prices of all the racks having the same gasoline specification are pooled, those racks affected by the merger and those presumed to be unaffected by the merger, and the effect of the merger is estimated using equation (1). When the regression is estimated using data from the merger and non-merger affected racks, the merger effect, β_1 , is calculated as a comparison of the average price in the merger racks post-merger compared to the average rack price of the non-merger racks pre- and post- merger and the merger racks pre-merger. Thus, the control group price is an average of the rack price in merger cities pre-merger and the rack price in non-merger cities pre- and post-merger. This is not an appropriate control price. Since the merger effect is being calculated based on a control price equal to the average price of all the cities during different time periods in the data, this analysis provides neither a meaningful pre- and post- merger comparison nor a good sensitivity check.

Both merger affected and non-affected racks could have been used to calculate a difference in difference estimator, if implemented correctly.¹⁴ One way for the estimator to be implemented correctly would be to add an additional variable to equation (1) as follows:

$$PRack_{i,t} - Crude_t = \beta_0 + \beta_1 DMerger_{i,t} + \beta_2 DMergtime_{i,t} + \beta_3 InventoryRatio_{j,t} + \beta_4 Utilization_t \quad (2)$$

Prack(i,t)	= average rack price in city i week t
Crude(t)	= price of WTI at week t
Dmerger(i,t)	= for racks affected by the merger 0 pre merger, 1 post merger, non-merger racks 0
Dmergtime(i,t)	=for all racks 0 pre merger 1 post merger
InventoryRatio(j,t)	= inventory/estimated demand for PADD j week t
Utilization(t)	=national refiner utilization rate at week t

¹³ See E.E. Leamer, *Specification searches: ad hoc inference with nonexperimental data*, Wiley, New York (1978).

¹⁴ The differences-in-differences model is discussed in depth in Angrist, J.D. and Krueger, A.B, "Chapter 23: Empirical Strategies in Labor Economics," *Handbook of Labor Economics*, Vol. 3, 1999, pp. 1296--1299.

The estimate of β_1 would be the merger effect controlling for the time-specific effects of gasoline prices unrelated to the merger.

Even with the difference-in-difference estimation, it is again crucial to pick the non-affected racks carefully. The control racks should have the same supply, demand and regulatory characteristics. Given that finding control racks can be difficult it is important to also include additional variables measuring supply, demand and regulatory changes to measure possible differences in the merger and control racks.

2. Key Control Variables Important to Wholesale Gasoline Prices Were Excluded

There are a number of factors that affect price not currently included in the model. The FTC staff comment in December stated:

“In addition to the number of supply outages that need to be included in the model, all areas of the country do not have the same reformulated gasoline, the formulation changes between winter and summer happen at different times of the year in various parts of the country, there have been a large number of changes in gasoline formulations in addition to RFG (reformulated gasoline) and CARB (California Air Resources Board gasoline) and phase II of RFG began in 2000. In addition, the price of conventional gasoline in any state or city may be affected by the existence of reformulated gasoline. For example, conventional gasoline on the West Coast may be higher priced than conventional in the rest of the country since conventional gasoline in the west is a substitute in production for CARB gasoline.”

FTC staff also gave the GAO staff copies of a map showing all the various formulations of gasoline within the country, including CARB and RFG as well as low Reid vapor pressure conventional gasoline that must be sold in various parts of the country, usually during the summer months. There is no mention in the Report of these other formulations and whether changes in these formulations possibly affected the regression results. All of these issues suggest the need for additional variables in the analysis.

In addition, there are several important variables that were removed from the GAO December methodology. The original December 2002 methodology included a number of dummy variables that may have removed most of the meaningful variation from the data.

Therefore, excluding some of these variables may have been warranted.¹⁵ But *completely* omitting measures of supply disruptions, seasonal variables, or year dummies is not appropriate.

Consider first the impact of supply disruptions on the estimation of price effects. Clearly the effect of the supply shocks that caused Midwest Gasoline price increases of the summer of 2000 and West Coast outages in various years will influence the results of the merger effect regressions because these outages were in the post-merger period for some of the mergers.¹⁶ To the extent that prices were higher in the after period as a result of these outages, not controlling for these variables would result in an observed “effect” from the merger where there may have been a smaller effect or none at all.

While the inventory ratio variable may control for these effects to some degree as the Report suggests in footnote 87, preliminary work by FTC staff suggest that inventories do not change dramatically when there are supply shocks. Our regressions of rack prices in Midwest cities on the PADD II inventory level and a supply disruption variable for the Midwest Gasoline episode shows that the supply disruption variable has a large positive coefficient and is highly significant. In December, FTC staff gave the GAO a copy of a recently published paper documenting supply outages on the West Coast and discussed the need for the GAO staff to research other supply outages.¹⁷

The results in the current Report do not incorporate any information on supply outages either in the regressions (where they should be) or in interpreting the results. Given the high number of supply outages in the year 2000 and that the GAO data set ends in December of 2000,

¹⁵ Including week, season, year and PADD level dummy variables in the same regression may remove much of the meaningful variation while controlling for seasonal effects. If the year dummies are to proxy for supply disruptions such as refinery outages, a change in the level of gasoline imports, gasoline formulation changes, or a demand change, it would be better to measure these effects directly.

¹⁶ For at least some of the West Coast mergers, the after-merger period ends before the major supply disruptions in 1999-2000.

¹⁷ Taylor, C., and J. Fischer, “A Review of West Coast Gasoline Pricing and the Impact of Regulations,” *International Journal of the Economics of Business*, Vol. 10(2), 2003, pp. 225--243.

estimating the effects of the Exxon-Mobil merger and the Marathon-UDS transaction, which occurred in 2000 and late 1999 respectively, will be very difficult even with a control methodology. Isolating the effect of the merger from the effect of supply disruptions will be difficult because both occurred during the same time period. As FTC staff noted in its December 2002 comments on the GAO's draft model (footnote 8):

“Some of the mergers will be especially difficult to model. The Marathon Ashland purchase of Ultramar Diamond Shamrock assets in Michigan occurred in late 1999 and the data ends in 2000. The summer of 2000 includes the Midwest gasoline episode as well as major pipeline problems in Michigan. In addition a refinery closed in Michigan right before the purchase. The same lack of sufficient length of the data set applies to other mergers as well depending on which areas of the country are being examined.”

Another type of supply shock involves changes in fuel specifications. In 2000 the implementation of RFG phase II (one of the prime causes of the Midwest gasoline spike) began, which likely increased costs for the industry. As mentioned above, multiple fuel specifications, such as RFG and CARB, or changes in fuel specifications, such as RFG phase I to RFG phase II can cause the price of conventional gasoline to be higher in a given region because all of these products are substitutes in production.¹⁸

Another glaring omission in the Report is the lack of controls for seasonal effects. A simple graph of the difference between the spot or rack price of gasoline and the price of crude oil will show that this margin is highly seasonal: the margin is generally wider in the summer than in the winter.¹⁹ Given the short merger windows used by the GAO and the lack of controls for seasonal affects, some of the merger results (such as Marathon-Ashland, Shell-Texaco II (Motiva) and UDS-Total) are likely being driven by seasonal effects, not by the mergers themselves. For example, the time periods examined for the Marathon-Ashland joint venture

¹⁸ In order to produce additional reformulated gasoline a refiner has to produce less conventional gasoline, all else being equal. A refiner will decide on how much reformulated and conventional gasoline to make based on the relative margins. Since making reformulated gasoline requires additional capital investment, the margin on reformulated, and hence conventional, must be higher to cover the capital investment.

¹⁹ The price-concentration analysis in the Report did use seasonal dummy variables.

were four years before the merger and six months after. The six months after included the summer of 1998. The finding that prices were higher during this period is not surprising because prices generally rise in the summer. The time periods examined for the Shell-Texaco II joint venture (Motiva) were six months before the merger and six months after. In this case the comparison was between the summer of 1998, before the merger, and the winter of 1998 after. Finding that prices decreased during this time period also is not surprising. The time periods and effects for the UDS-Total merger are similar.

3. Report's Use of Instrumental Variables is Problematic and Incomplete

Another methodological problem involves the use and choice of instrumental variables in the Report. There is no discussion in the Report of why the instruments used²⁰ are valid instruments for the given endogeneity problem.²¹ Footnote 11 of the December FTC staff comments said:

“If the controls are correlated with the HHI, then using weak (or inappropriate) instruments may give worse estimates than not using instruments. The final Report needs to explain why instruments for control variables are needed, show that the instruments are sufficiently powerful as to improve the results and discuss the changes in the results when using instrumental variables.”

²⁰ The instrumental variables include: number of suppliers at the rack, state level unemployment rate, the previous period's state level unemployment rate, the previous period's inventory ratio, the previous period's utilization, level, a time trend, and a time trend squared. The number of suppliers at the rack changes for a number of reasons including mergers. There is also no discussion in the Report of how the number of suppliers at the rack was determined. If the number of suppliers at the rack is determined by counting firms posting at the rack, this raises other issues, including: one firm posting multiple brands at a rack and traders who post infrequently at the rack and do not have their own supply but are merely reselling gasoline from another supplier to that rack.

²¹ The endogeneity problem in this case is that prices, inventories, and utilization (a gross measure of quantity) may be jointly (simultaneously) determined. Nevertheless, the test for endogeneity assumes that the instruments being used in the test are valid.

While this remark was made in the context of endogeneity in the price-concentration analysis, it also applies to endogeneity issues in the merger event analysis.

The current Report does test for endogeneity, *i.e.*, whether there is the need for instruments, but does not discuss why these are valid instruments or show that the instruments are sufficiently powerful to improve the results. There is a well known economic literature on the impact of using inappropriate or weak instruments in instrumental variables regression.²² As a general matter, the first stage instrumental variable results should be reported to show the effects of using the instruments. The goodness of fit of the first stage regressions as well as the coefficient estimates and their significance are important in evaluating the use of instrumental variables. The Report does not provide these results, making it difficult to assess the validity of these instruments. (The suggestions that the instrumental variables estimation is appropriate because the results do not change very much when estimating the regression with and without instrumental variables is discussed in Section D below.)

4. Specification of Before and After Periods of Merger Events is Problematic

Another issue with the “event” study methodology used by the GAO involves the specification of the before and after periods when multiple mergers affect a rack in close succession. The point of an event study is to isolate the effect of a given event, in this case a merger, from all the other events that have occurred. The best way to isolate the effect in this case would have been to concentrate on the racks that did not have mergers in rapid succession. The second best way to isolate the events would be to specify overlapping merger dummy variables. In effect, the approach would be to assess whether the first merger has a price effect and then test whether the price effect of the second merger would be a price increase or decrease on top of the first merger effect. Instead, in the Report, the effect of the second merger is

²² For an example see, Staiger, R.W. and J.H. Stock, “Instrumental Variables Regressions with Weak Instruments,” *Econometrica*, Vol. 65, 1997, pp. 557--586 and Bound, J., D. Jaeger and R. Baker, “The Cure Can Be Worse Than the Disease: A Cautionary Tale Regarding Instrumental Variables,” NBER Technical Paper 137, June 1993.

calculated by comparing the prices after the second merger to the prices before the first merger. This procedure will give misleading results, because the price effects of the second merger will reflect competitive conditions at the time it occurs, not competition prior to the first merger.

The specification of the merger dummy variable is of crucial importance. The effect of the merger is estimated by comparing the average price before and after the merger. As mentioned earlier, in some cases the before-merger period is multiple years and the after-merger period is only six months. It is an important question whether six months of data is sufficient to reasonably calculate a merger effect. With respect to this issue, the Report states: “When mergers closely followed each other, it tended to shorten the before and after merger time periods that we could model, especially when more than one merger affected the same rack cities. Nonetheless we believe we had enough data.” There is no discussion of how this conclusion was reached. The ability to do sensitivity testing on the size of the merger windows is not as problematic as the Report suggests. Additional data could have been used in the post-merger period to see if the results changed. In addition, the pre-merger period could have been shortened to see if that had any effect as well. Sensitivity tests on the duration of the pre- and post-merger periods should have been conducted; unfortunately, they were not.

D. Reports Results are Not Robust in Many Cases

Because the basic methodology is fundamentally flawed, exhaustively discussing the results at length is unwarranted. One important point, however, that further undermines the reliability of the results is that many results are not robust to the different estimations used in the Report. Different specifications and estimation methodologies can frequently be used to estimate a given relationship. To the extent estimates vary significantly across specifications, assessing what is the “true” relationship is difficult (unless one has good reasons to pick a particular specification or group of specifications as more reliable and clearly explains the choice).

There are several examples of lack in robustness in the results. For example, consider Tables 15-17 of the Report, which summarize the estimated merger effects from various specifications. In Table 16, the estimated price effect of the BP-Amoco merger ranges from no

statistically significant price effect to 3.5 cents a gallon among the three regression specifications. In Table 17, the estimated price effects of both the Tosco-Unocal and the Shell-Texaco I transactions shift from being large, negative and statistically significant to being large, positive and statistically significant among the regression specifications. These results show that small changes in specification lead to large changes in results. The inclusion of year effects and using data on all the racks, albeit incorrectly,²³ has sizeable effects on the regression results for a number of the mergers.

Other examples concerning the lack of robustness of the regression results include the results with and without instrumental variables. The Report mentions that the results using the instrumental variable estimation techniques are not very different from the fixed and random effects regression results. Examination of the regression results shows that this is not true. Tables 27-32 show the individual merger regression results. The first few columns of these tables show the results without instrumental variables and the last few columns show the results with instrumental variables. The regressions estimating the effects of the MAP-UDS merger on the price of conventional gasoline show that without instrumental variables there would have been no estimated price effect. The same is true for the estimates of the effect of the BP-Amoco merger on conventional gasoline, the BP-Amoco merger on reformulated gasoline, and Shell-Texaco I joint venture on branded CARB gasoline.

Moreover, when the instrumental variables approach is used, the relationship between price and refinery capacity utilization goes from positive and significant to negative and significant. The relationship between refinery utilization and prices is strongly expected to be positive. This is yet another reason to suspect that the use of these particular instruments was problematic.²⁴

²³ As discussed on page 9, the pooling of all racks was incorrectly done. The merger effect is being calculated based on a control price that is an average of non-merger affected racks before and after the merger period and the merger racks before the merger.

²⁴ The rationale for dropping the utilization variable from the preferred regression specifications does not make sense. The fact that utilization is correlated with the inventory ratio variable is not relevant. Eliminating variables from a regressions because of multicollinearity when both variables are independently significant is not appropriate. There are a number of

The current results are also not robust across different racks. Comparing Tables 37 and 38 shows that the Tosco-Unocal merger did not have a statistically significant effect in the instrumental variables specification with the three racks used in the estimation in Table 38, but had a very large effect when using the six racks in California. Because the effects estimated for all six racks include the three racks where the merger had no effects, the Tosco-Unocal merger had a very large effect on those additional three other racks. The sensitivity of the results to which racks are included is an important robustness check and helps the reader to judge the quality of the results. The Report fails to report results for subsets of the racks or indicate whether such sensitivity checks were conducted.

Another problem is that the very strong statement in the Report that all known variables that affect the wholesale price of gasoline have been included in the regressions cannot be supported. Few of the regressions explain more than 15-20 percent of the variation in the dependent variable. If any additional control variables that would increase the explanatory power of the regressions are correlated with the merger time periods and/or cities, which is highly likely, the merger results would change.²⁵

IV. Price Concentration Methodology

The Report's second econometric analysis seeks to describe the relationship between wholesale gasoline prices (adjusting for crude oil costs) and state-level concentration. While price-concentration studies were once a focus in the economics literature on market structure and industry competitiveness, these studies have been largely abandoned in favor of analyses like merger event studies that attempt to model more directly and with more precision the effects of structural change (such as mergers) upon prices. There are a number of widely recognized methodological issues with price concentration studies. These issues were highlighted by the

significant changes in the results when the utilization variable is excluded.

²⁵ As mentioned earlier, variables measuring seasonality, supply outages and formulation changes will likely be correlated with the merger variable.

FTC staff in their communications with GAO staff last December. Surprisingly, the Report does not acknowledge these problems. The results of the Report's price-concentration analysis also suffers from a lack of robustness, and there is little discussion comparing or reconciling the results from the price-concentration study and those of the merger event study.

A. Methodological Issues in Price Concentration Studies

There is a large literature on the problems with obtaining meaningful estimates from price or margin on concentration regressions.²⁶ Meaningful estimates are estimates that can be interpreted as causal.²⁷ In other words, the estimated relationship can show that increases in concentration *result* in higher prices or margins. The key issues were summarized in our December 2002 letter and are still relevant to the Report:

The wholesale antitrust market(s) for gasoline are not likely at the state level; some markets are smaller and some larger. Unless the GAO has evidence that the changes in state level HHI's are closely correlated with changes in the concentration of relevant markets, the analysis is unlikely to provide meaningful results. The effect of aggregation on the estimated relationship is not easily predicted.

The reduced form model of price/margin as a function of HHI has a number of theoretical problems. The major problem is that the coefficient on HHI can not be estimated consistently. Articles in the *Handbook of Industrial Organization* (1989) by Schmalensee and Breshnahan discuss this issue. If the GAO is going to estimate this type of relationship, the large literature on the problems with this

²⁶ Examples of this literature include: Evans, W., L. Froeb, and G. Werden, "Endogeneity in the Concentration-Price Relationship: Causes, Consequences, and Cures" *Journal of Industrial Economics*, Vol. 41, 1993, pp. 431--438 and Breshnahan, T.F. "Empirical Studies of Industries with Market Power," in Schmalensee R. and Willig R. (Eds), *Handbook of Industrial Organization*, Vol. II, Ch.17, 1989, (North-Holland, Amsterdam).

²⁷ Given the number of well documented theoretical problems with estimating and interpreting price concentration regressions, the following statement on p. 146 of the Report cannot be supported: "Also we used the market concentration model because market concentration better represents overall market conditions than mergers."

type of model needs to be acknowledged and addressed and/or the results should be given with appropriate caveats acknowledging these problems. The problems with estimating this type of relationship include:

(1) HHI is a function of individual firm price and quantity decisions which are affected by supply and demand shocks. The error term in the regression is also a function of these shocks. Therefore the HHI and the error term are likely to be correlated. While the GAO staff are using an instrument, number of suppliers in a state, for HHI to mitigate endogeneity issues, it is unclear why this instrument will allow the identification of the effects of competition or efficiency but is not correlated with other variables of interest. While HHI in a state and the number of suppliers will be correlated, it is likely that the number of suppliers will be correlated with barriers to entry, supply shocks, exit, etc.

(2) HHI may also be correlated with omitted variables that affect price such as various measures of fixed and variable costs and barriers to entry.

It is important to understand the source(s) of variation in the HHI both in formulating and in interpreting the model. The annual or monthly HHI's may be changing due to mergers, entry, exit, or relative price changes caused by supply disruptions or other factors. Few of the changes in the HHI's will be caused by mergers. This point needs to be made clear in interpreting the results. In addition, there is evidence that HHI may only matter past a critical point. The model as currently written is testing for a linear relationship. Alternative specifications should be used to test other functional forms.

There is no basis for the suggestion in the Report that, to the extent racks are close to each other and would tend to have similar market characteristics, the available state level data on market concentration is a reasonable variable. Although the GAO staff have added additional instruments to the December 2002 methodology, the same criticisms apply. These instruments – the unemployment rate and lags of the dependent variables – are unlikely to solve the multiple endogeneity problems nor do they solve the market mismeasurement problem. In addition, the criticisms about omitted variable bias and instrumental variable issues outlined when discussing the problems with the event study methodology are applicable to the price-concentration regressions as well. No functional form other than a linear relationship between price and concentration was tested. Given this linear relationship, a 100 point change in the HHI between 900 and 1000 is treated the same as between 4900 and 5000. Because any relationship between

price and concentration may not be linear (for instance, there may be only a relationship above a certain threshold), forcing a linear relationship may give misleading results. In addition, the use of year and HHI interaction terms should be discussed and justified. There is no discussion of why an interaction between the concentration and year variables for years when there was a sizeable change in the HHI is appropriate. No interpretation of these estimates is offered.

B. The Report's Price-Concentration Results are Not Robust

Abstracting from the general problems with estimating a price-concentration relationship, the price-concentration relationship as estimated in the Report is not robust in many circumstances. Tables 39 and 40 show that the relationship between price and concentration for conventional gasoline, both branded and unbranded, is negative and significant in the first fixed effect instrumental variable regressions. In each of the next two estimations a variable is dropped and the relationship becomes positive and significant. There is no reason to prefer the last two specifications. As discussed earlier, there is no reason to drop the utilization or the inventory variable because of multicollinearity. The Report could just as easily have reached the conclusion that the relationship between the price of conventional gasoline and concentration is negative based on the results presented in the Report.

While many of the specifications showed a positive and significant relationship between price and concentration, the range of the effects is very large. The effects shown on Tables 39 to 46 give a range of the price effect of a 100 point change in the HHI of 0 to approximately 3 cents per gallon. The results in Tables 18 and 19 show much larger effects. By adding seasonal dummy variables the price effect of a 100 point change in the HHI ranges from 0 to 4 cents per gallon.²⁸ With year interaction terms, the rationale for which is not clearly explained, the effect of a 100 point HHI change can be as large as 10 cents per gallon.

²⁸ It is not clear why seasonal dummy variables are included in the price-concentration regressions but not in the merger event studies. Any variable that explains variation in the price variable should be included in both studies. If seasonal effects were significant in the price-concentration analysis (and they must have had sizeable effects given the change in the results), then they should have been included in the merger event analysis.

The results presented in Tables 41 and 42 show estimates of the effects of market concentration on branded and unbranded conventional prices by regions. Results are not reported for the specifications of fixed or random effects instrumental variables regressions with both control variables, utilization and inventory ratio. The results on these two tables do show that without using the instruments there would be no relationship between concentration and price in conventional gasoline in the eastern half of the United States.

There is also no discussion in the Report comparing the results of the merger analysis and the price-concentration analysis. Because concentration is affected by factors other than mergers, such as entry, exit, and expansion, the sum of the estimated merger effects should be less than the effect from the change in concentration.²⁹ In at least some cases, this is not true. Table 19 shows that the aggregate effect of increased concentration in California has raised the price of branded CARB by 5 cents per gallon. The combined estimated effect of Shell-Texaco I (Equilon) and Tosco-Unocal mergers on the price of branded CARB given in Tables 37 and 38 is 8.5 cents a gallon. Because these two mergers caused only a portion of the change in concentration in California, a price change associated with the mergers larger than that associated with the change in concentration is puzzling. The same comparison can be made from other mergers and regions. The GAO preferred estimate of the price effect of the Exxon-Mobil merger is approximately a nickel. The entire change in concentration in PADD's I-III is calculated to have a similar effect on price.³⁰

²⁹ This would be true unless the merger effects were very short lived.

³⁰ Even as a purely descriptive matter, the Report does a poor job in linking concentration changes to mergers. In Chapter 3, however, the Report offers a statistical correlation analysis to associate the degree or connections between merger activity and concentration. Data used for this correlation are the HHI estimates for domestic crude oil, refining capacity, and the Herold data set on merger and acquisitions. This analysis shows positive correlations between HHI and merger activity. Correlation analysis does not establish causation, however, and we suspect that similar results would have been obtained had HHI estimates been correlated with the overall merger activity in the economy or stock market indices. As far as we can tell, merger transaction value from the Herold data set are not separated out by industry segment or geographic area: thus it appears that any functional level (crude, refining, or wholesaling) and geographic area that might be affected by a given merger appears to have the same weight in the correlation against the relevant HHI. The data for crude oil and

V. The Report's Documentation is Poor

In any well-performed study, the descriptions of the data work and the econometrics are sufficient to allow the reader to understand fully how the calculations were done. Such descriptions enable interpretation and replication of the results. The current Report does not include complete descriptions of a number of calculations and regressions. For example, while the merger retrospective regressions do list the number of racks included in each regression there is no list of which racks are included in each regression. By itself, this omission makes interpreting the results and replication impossible.

Moreover, the Report contains no discussion of how divestitures were handled in the estimation. For example, it is unclear from the text whether the racks affected by the Exxon-Mobil divestitures are included in the regressions analyzing the impact of the Exxon-Mobil merger. The description in the Report about the Exxon-Mobil merger is simply unclear. On the one hand, there are suggestions that there was an estimated increased price from the merger in markets in PADD I where divestitures of wholesale and retail assets eliminated the overlaps. On the other hand, the report does not estimate the effects of the Exxon-Mobil merger in California where there were also divestitures. It is similarly unclear what procedure was followed for the BP-Amoco merger with divestitures in the Eastern United States or the Shell-Texaco merger with divestitures in California.

Another problem is that the information in the text or in the tables on occasion seems to be contradictory. One example is the description in Table 14 that suggests that any given merger variable has a value of one at the time of the merger and stays one throughout the rest of the data

refining does permit estimation of the concentration changes associated with individual mergers because the data is reported by firm. This would be the usual approach in connecting mergers with changes in HHI. While the wholesale data does not identify firm-specific data because of confidentiality restrictions imposed by EIA, this data is available monthly. Because the month when transactions are consummated is known, fairly strong inferences can be made about the effect of particular mergers on HHI. It is not clear why such an analysis was not done in the Report.

set. This procedure would lead to overlapping merger variables. The text, however, discusses the need to avoid overlapping merger variables.

Finally, there is no discussion of how price is measured. Although the prices are rack prices from the Oil Price Information Service, there is no discussion of the exact calculation of the rack averages and which formulations are used, such as reformulated with MTBE or ethanol. How the racks were used to calculate state level averages for the price concentration analysis is also omitted.

VI. Conclusion

Omitted control variables in both of the Report's econometric analyses are fundamental flaws. Both studies: (1) fail to control adequately for exogenous factors that impact wholesale gasoline prices; (2) suffer from endogeneity problems that are not adequately addressed; and (3) have results that are not robust.

There are other problems specific to each study such as questionable pre- and post-event periods in the merger event study and the assumption that state-level wholesale markets are economically meaningful in the price-concentration study. The Report's documentation of its methodology is inadequate, particularly in view of the potential significance of the Report's findings for public policy in the petroleum industry. These flaws make the Report unable to isolate reliably either the effects of mergers or of concentration on wholesale gasoline prices. Thus, the Report cannot be used to inform public policy.

Enclosure 2



Bureau of Economics

**Enclosure 2:
Comments on the GAO methodology for "Econometric Analysis of Effects of Market
Concentration and Mergers on U.S. Wholesale Gasoline Prices in the 1990's"**

December 20, 2002

General Comments: These are Federal Trade Commission (FTC) Bureau of Economics Staff preliminary comments on the GAO draft model.¹ In order for FTC Staff to provide complete comments the GAO needs to expand and clarify the models in several respects as discussed below. FTC staff expect that they will have further opportunities to review additional drafts of the proposed models. The draft methodology is not a complete description of the analysis. The methodology does not discuss any underlying theoretical model which motivates the reduced form model(s). The reduced form model(s) estimated are not fully described, *i.e.* the description does not include the estimating equations or a complete/clear description of the variables and how they are aggregated. The discussion of the estimation technique is terse and is not completely correct.

Because the description of the methodology is incomplete, it is impossible to provide a comprehensive list of issues that need to be addressed. FTC staff will need to review a complete and revised draft model, the underlying data, the programs used to estimate the model, and the results from the analysis to assess fully the validity of the model.

¹ The views expressed in this memorandum are of Bureau of Economics staff members. These comments do not necessarily reflect the views of the Commission, the Commissioners or other members of the staff.

Confidential

General Comments on Model 1 - Market Concentration Model - Using State Level Data

- The wholesale antitrust market(s) for gasoline are not likely at the state level; some markets are likely smaller and some larger. Unless the GAO has evidence that the changes in state level HHI's are closely correlated with changes in the concentration of relevant markets, the analysis is unlikely to provide meaningful results. The assumptions in the 1986 report that state level aggregation are acceptable, since the state data is an average of relevant markets, is not correct. The effect of aggregation on the estimated relationship is not easily predicted.²
- The reduced form model of price/margin³ as a function of HHI has a number of theoretical problems. The major problem is that the coefficient on HHI can not be estimated consistently. Articles in the handbook of Industrial Organization(1989) by Schmalensee and Breshnahan discuss this issue. If the GAO is going to estimate this type of relationship, the large literature on the problems with this type of model needs to be acknowledged and addressed and/or the results should be given with appropriate caveats acknowledging these problems. The problems with estimating this type of relationship include:

² Another assertion in the 1986 report, that monthly HHI's are less meaningful than annual HHI's is not necessarily correct. Monthly HHI's are available. The difference between using monthly HHI's and annual HHI's should be discussed and tested. The use of yearly dummy variables along with annual HHI's may also be problematic. There are also conceptual problems with using annual observations on HHI (and other variables) to explain weekly margins. This type of estimation will likely create residuals that are correlated over time and therefore give incorrect results.

³ The use of West Texas Intermediate (WTI) as the cost variable in the margin calculation may affect the results. A better measure of crude costs would be the crude acquisition costs collected by EIA. In any event, sensitivity analysis on this issue is warranted.

Confidential

(1) HHI is a function of individual firm price and quantity decisions which are affected by supply and demand shocks. The error term in the regression is also a function of these shocks. Therefore the HHI and the error term are likely to be correlated.⁴

(2) HHI may also be correlated with omitted variables that affect price such as various measures of fixed and variable costs and barriers to entry.

- It is important to understand the source(s) of variation in the HHI both in formulating and in interpreting the model. The annual or monthly HHI's may be changing due to mergers, entry, exit, or relative price changes caused by supply disruptions or other factors.⁵ Few of the changes in the HHI's will be caused by mergers. This point needs to be made clear in interpreting the results. In addition, there is evidence that HHI may only matter past a critical point.⁶ The model as currently written is testing for a linear relationship. Alternative specifications should be used to test other functional forms.

Comments on Model 2 - Specific Mergers Model Using Terminal City or Rack Data

⁴ While the GAO staff are using an instrument, number of suppliers in a state, for HHI to mitigate endogeneity issues, it is unclear why this instrument will allow the identification of the effects of competition or efficiency but is not correlated with other variables of interest. While HHI in a state and the number of suppliers will be correlated, it is likely that the number of suppliers will be correlated with barriers to entry, supply shocks, exit, etc.

⁵ The current document does not mention or include variables on the myriad of supply outages that will affect the results. It is important to understand the supply relationships between various cities and states and control for supply outages in neighboring states. The Midwest Crisis variable in the GAO model for the summer 2000 price spike is only for June in PADD II. This is incorrect. The number of additional supply outages in California, Washington, and the upper Midwest within the data set is substantial. All of these supply shocks need to be included as control variables.

⁶ Carlton & Perloff, Modern Industrial Organization, 1994, p358.

Confidential

- This model is being used to examine eight mergers from 1997 to 2000 in one estimation.⁷ This is an extremely ambitious project since it will be necessary to understand all the supply issues in each of these areas of the country for the entire data set. This one size fits all approach is unlikely to work well. To validate this approach any number of pooling tests will need to be done.⁸ A better approach would be to examine each merger and the areas affected by the merger separately. Looking at each merger separately would make it easier to control for supply disruptions and other confounding effects.

- There are a sizeable number of other factors that affect price not currently included in the model. In addition to the number of supply outages that need to be included in the model, all areas of the country do not have the same reformulated gasoline, the formulation changes between winter and summer happen at different times of the year in various parts of the country, there have been a large number of changes in gasoline formulations in addition to RFG and CARB and phase II of RFG began in 2000. In addition, the price of conventional gasoline in any state or city may be affected by the existence of reformulated gasoline. For example, conventional gasoline on the West Coast may be higher priced than conventional in the rest of the country since conventional gasoline in the west is a substitute in production for CARB gasoline.⁹

⁷ There needs to be sensitivity testing with respect to the date of the merger. Most of these mergers had divestitures after the consummation of the merger.

⁸ Some of the mergers will be especially difficult to model. The Marathon Ashland purchase of Ultramar Diamond Shamrock assets in Michigan occurred in late 1999 and the data ends in 2000. The summer of 2000 includes the Midwest gasoline episode as well as pipeline problems in Michigan. In addition a refinery closed in Michigan right before the purchase. The same lack of sufficient length of the data set applies to other mergers as well depending on which areas of the county are being examined.

⁹ The use of nearby margins as a control factor is complicated by the number of gasoline specifications and other factors. The choice of the nearby margins will be difficult, needs to be discussed fully and multiple margins should be used to check the robustness of the results.

- It is unclear why the GAO is not also examining retail prices in these econometric models. A merger might raise rack prices but retail prices may go down. In other words a rack price increase is not a necessary and sufficient criteria for determining possible anticompetitive effects of a merger.

Comments Both Models

- Optimally the GAO will perform multiple estimation techniques so the readers of the report will understand the robustness of the results. Picking one technique and presenting one set of results does not make sense. There are multiple estimation techniques that could be used. A concern with the Newey-West estimator is that in small samples it may perform poorly. In addition the Newey-West estimator without fixed effects is likely to have residuals for each city/state with a permanent city/state fixed effect.¹⁰
- The reasons given for not using fixed/random effects estimation do not make sense. While fixed effects models do eliminate time invariant variables (although they can be recovered) this does not make the model less reliable. The fixed effect estimation captures the effect of all time invariant variables. Given the limited number of control variables proposed in this study, not using fixed effects is problematic.¹¹

¹⁰ If there is a city/state fixed effect that is not controlled for, then the correlation between residuals from distant time periods will not go to zero. This could lead to inaccurate standard error estimates as the Newey-West procedure assumes that correlation between residuals of greater than n lags is sufficiently small to be ignored.

¹¹ Also omitting potentially important control variables because there are not sufficient instruments to test for endogeneity is not appropriate. If the important variables are excluded, the remaining control variables will be correlated with the error term. Further, instrumenting for control variables can be very problematic. If the controls are correlated with the HHI, then using weak instruments may give worse estimates than not using instruments. The final report needs to explain why instruments for control variables are needed, show that the

Confidential

- There needs to be discussion of exactly which prices are being used. Are the models being estimated using the low rack price, the average of the lowest three prices, or the average price?¹² It is also not clear how the OPIS rack prices were aggregated to a state level price. There may be considerable aggregation bias in aggregating to state level prices. It will also be important to use different prices in multiple specifications of the model to check the robustness of results.

instruments are sufficiently powerful as to improve the results and discuss the changes in the results when using instrumental variables.

¹² It is important to discuss the prices gathered from the Census Bureau and how these prices match OPIS rack prices.

Enclosure 3

Empirical Evidence on the Competitive Effects of Mergers in the Gasoline Industry

**John Geweke
University of Iowa
Revised July 16, 2003**

Introduction

Horizontal mergers that significantly increase concentration can increase the ability of the merging firms to restrict their output and increase their prices and profits, even though competitors respond with their own output increases – a phenomenon sometimes termed “unilateral effects” since there is no explicit collusion. They can also improve the profitability of coordinated output restrictions. On the other hand, horizontal mergers can result in efficiencies (like scale economies) that reduce costs and prices. All three aspects may be present in a pending merger.

Vertical mergers do not increase concentration but change the behavior of the merging parties in both the upstream and downstream markets in which they compete. If an upstream and a downstream firm merge, it may be in the interest of the new firm to raise wholesale gasoline prices charged to downstream rivals. Depending on the structure of the wholesale market, this strategy could ultimately lead to foreclosure of that market to some rivals. In contrast, if a vertical merger reduces transactions costs or eliminates double marginalization then retail prices may fall, to the benefit of consumers.

The competitive effects of mergers in the gasoline industry has been a topic of great interest and controversy. For example, participants at the FTC Refined Petroleum Product Price hearings differed in their opinions about the effects of mergers on gasoline prices. There was general agreement that there has been substantial consolidation of the gasoline industry in the past two decades, but the extent to which this is attributable directly to mergers rather than the elimination of marginally competitive firms was less clear. There was much more disagreement about the competitive effects of mergers. Some participants believed mergers had reduced competition in an industry where there was too little competition to begin with. Others argued that mergers were part of an efficient, competitive response to changing cost and regulatory conditions.¹

The competitive effects of mergers in the gasoline industry bear heavily on current policy-making. At the national level there have been several large mergers or acquisitions in recent years (British Petroleum and Amoco, British Petroleum and Arco, Exxon and Mobil, Chevron and Texaco, and Phillips and Conoco) and scores of smaller such transactions. Many of these were approved only with FTC-mandated divestiture

¹ The Senior Assistant Attorney General for California, Tom Greene, saw a “striking increase in concentration” and former Senator and Chairman of the Consumer Federation of America, Howard Metzenbaum, cited “the lack of meaningful competition in the oil industry” and stated, “a wave of mergers drove this consolidation and concentration.” However, industry expert Phillip Verleger commented, “Petroleum products (prices) are more volatile and higher. Mergers in the industry are not—I repeat not—the primary cause. The proliferation of blends ordered by the EPA has reduced storage capacity and increased volatility.” John Cook, the director of the Energy Information Administration of the Department of Energy, observed, “When people observe that prices seemingly increase faster than they decrease, this generally leads to the speculation that perhaps market forces don’t explain all the variation so there must be some type of anticompetitive behavior at work. Our analysis suggests otherwise.”

requirements in markets where the merger was deemed to have anti-competitive effects. At a more local level, nearly every state, and some large cities, have at least considered divorcement legislation that prohibits ownership of retail gasoline outlets by refineries. Such laws are currently in force in six states and the District of Columbia. Informed policy-making requires reliable evidence on the impacts of alternative policy choices on prices, and (ideally but more ambitiously) on consumer welfare.

This report summarizes nine studies providing such evidence and assesses the reliability of the empirical work in each. Nine is not a large number, but taken together this work addresses a wide variety of issues using different kinds of data and methodology. Four of these studies examine horizontal concentration or mergers, and the other five are concerned with vertical integration. Four of the studies utilize national, cross-section time-series data, while the other five examine specific states and events. With respect to methodology, five of the studies set out to measure the impact of past changes in structure, either mergers or divorcement legislation; three examine the relationship between measures of horizontal or vertical concentration, on the one hand, and gasoline prices, on the other; and one utilizes a structural model of consumer, retailer and refiner behavior to predict the impact of hypothetical but specific mergers. Each approach has its strengths and limitations, and studies vary in the degree to which they realize the potential of the methodology used.

Evidence on horizontal mergers

U.S. General Accounting Office Accounting Office (1986)

In 1986 the U.S. General Accounting Office (GAO) studied the relationship between price and horizontal concentration in the wholesale gasoline market. The work was undertaken because of concerns about the simultaneous increase in gasoline prices, decrease in crude oil prices and two major mergers in 1985. The mergers—Texaco's purchase of Getty and Chevron's purchase of Gulf—were allowed to proceed after the FTC required divestiture of refineries and wholesale and retail outlets in geographic regions with significant overlap.

Given the price volatility in petroleum markets, it would be very difficult at best to determine the impact of one or two specific mergers on wholesale or retail gasoline prices, and the GAO study does not take this approach. Instead, it models the wholesale gasoline price in each state as a function of that state's horizontal concentration in wholesale gasoline. Concentration is measured by the Herfindahl-Hirschman Index (HHI).² The study employs a regression model³ and uses aggregate data for each of the 48

² This is the sum of squared percentage market shares of each wholesaler; see Carlton and Perloff (2000) p 247.

³ The study also presents estimates of a structural model that distinguishes between demand and supply. This was done in an effort to isolate possible supply or demand shifts in different years, a question distinct from that of the relationship between HHI and price. The findings in these models were consistent with

continental states and the District of Columbia, for each of the 32 months from February 1983 through September 1985. (The report was issued in September 1986.) Thus the GAO data set has a classic cross-section time-series structure.

Assessment of the technical work in the GAO report is hampered by the fact that the report's documentation of data and estimation methods does not meet generally accepted academic standards.⁴ State gasoline prices each month correspond to the first sale of gasoline in a state for consumption in that state, as compiled by the Energy Information Administration (EIA).⁵ The report provides no indication of how prices were actually constructed, but it is clear that there are distinct prices for each state and month, and that they were seasonally adjusted by GAO.⁶ HHIs were based on EIA wholesale market share data, but were prepared only annually—not monthly—by state.⁷ In the study's regression model, wholesale gasoline price is also a function of the quantity of wholesale gasoline sales in the previous month, per capita income, the price of crude oil, and the difference between the price of home heating oil and the price of crude oil.⁸ The report provides no sources for these data, nor does it indicate which of these measures are specific to a state and which are not, nor which are truly monthly measures and which are annual or are interpolated to monthly values. It appears, but is never stated explicitly, that many variables (including HHI and gasoline wholesale price) are transformed to logarithms. All of these details are critical in interpreting the implications of the reported estimates for the question of the relationship between concentration and price. In addition to these variables, the model in the GAO study also includes indicator variables for summer and for winter, and for the years 1984 and 1985.

The study estimates two variants of the model. The second variant includes, in addition to these explanatory variables, indicator variables for each state. Consistent with the econometrics literature, the study refers to this as an "error components" model. As this literature recognizes,⁹ the relationship in an error components model differs in a fundamental way from that in a regression model. A regression model captures the impact of persistent geographical differences in the explanatory variables across states (like the HHI) on the outcome variable (wholesale gasoline prices, in the GAO study). In an error components model these persistent differences are explained away by the state indicator variables: the model captures only the systematic effects of explanatory variable changes over time on outcome variable changes over

those in the regression and error components models, and they do not overcome the difficulties with those models discussed here.

⁴ This may be due, in part, to time limitations imposed on the preparation of the GAO report; see GAO (1986), p 38.

⁵ GAO (1986), p 39.

⁶ GAO (1986), p 37.

⁷ This fact emerges only in GAO's response to one of the FTC comments on their study; see GAO (1986), p 62.

⁸ The last variable is included as a measure of the profitability of producing refined products other than gasoline (GAO (1986), p 41).

⁹ Stock and Watson (2003), Section 8.3, provides a clear and simple exposition. For more technical details see Greene (2003), Sections 13.2-13.5.

time, in each state. Thus the coefficient on HHI in the regression model provides a summary answer to the question, "Controlling for other relevant factors, how has variation in HHI over states and time systematically affected the wholesale price of gasoline?" whereas in the error component model the question answered is, "Controlling for other relevant factors, how have variations in HHI over time within each state systematically affected the wholesale price of gasoline?" The advantage of an error components model is that it controls for omitted relevant factors that differ across states (but not over time). The disadvantage is that there is less variation in the data available to accurately estimate the impact of variables of interest (like HHI) on the outcome (wholesale gasoline price).

These features of the model, by themselves, render any inference from the GAO study about the relationship between concentration and wholesale price rather difficult. The GAO regression model provides a coefficient estimate of 0.013 for HHI (Table I.1, p 46). Presuming, as the text of the report appears to do, that HHI and prices are measured in logs, this implies an "elasticity" of 1.3% of wholesale price with respect to HHI: for example, if HHI rises from 2000 to 3000 points, prices increase by 0.53%. This is an extremely small effect; according to the GAO report it is statistically significant, a point to which we shall return shortly. In the error components model the HHI coefficient is 0.041, over three times larger but still quite small. According to the reported standard errors, this value is statistically significantly greater than the estimate of 0.013 in the regression model. The interpretation of these results, however, is at best clouded by the nature of both the regression and error components models. Recall that both models remove systematic differences in the calendar years 1983, 1984 and 1985 by means of indicator variables, and that the HHI values are forced to be the same within each calendar year in each state. Thus both models exclude any nationwide effect of changes in HHI on changes in price in the 1983-1985 period. For reasons already discussed, the error components model also excludes systematic relations between HHI and wholesale prices across states. Thus, by construction, the GAO error components model focuses on the impact of year-over-year changes in state HHIs that are uncorrelated with year-over-year nationwide changes in HHIs. One would expect these effects to be small compared to the more important systematic changes in concentration and price over years or across states. Thus the very small impact of concentration on price found with the GAO model can therefore be ascribed primarily to the statistical methodology of that model.

All of the coefficient estimates in the GAO model are statistically significant (save one, the seasonal "winter" indicator variable in the error components model) and most are highly significant. For example, the elasticity of wholesale gasoline price with respect to income per capita is estimated to be 0.066 in the regression model, with a standard error of 0.007 (a t-ratio of almost 10), and in the error components model it is found to be 0.257 (standard error 0.017, t-ratio 15). Genuine statistical significance connotes informative data, but informative data appear lacking here. It is not even clear whether GAO was using

monthly or annual per capita income by state; there are no reliable monthly state data in any event.¹⁰ Given the included indicator variables for state, calendar year, and season, the model has already eliminated most variation in income, and what is left is precisely that which cannot be measured with any accuracy. Yet the results claim to have estimated elasticity accurately to within a few percent.¹¹

In interpreting the GAO findings, the possibility that the reported high statistical significance is spurious must be taken into consideration. The GAO report used conventional least-squares standard errors, which are valid if the error terms in a regression equation are uncorrelated. As in most cross-section time-series models, this assumption is presumptively untrue: an unusually high error term for one state in a given month will typically be accompanied by unusually high ones in other states (cross-section correlation), and for the same state error terms in adjacent months are almost certainly serially correlated (autocorrelation). The GAO report indicates that serial correlation exists, but dismisses the problem, noting that coefficient estimates are about the same when the problem is taken into account in a more sophisticated model.¹² As all econometrics texts point out, however, the issue is not one about coefficient estimates (which remain unbiased) but rather about standard errors¹³, which are typically and often dramatically underestimated in these circumstances.¹⁴ If, as is likely the case here, standard errors are understated by a factor of four or more, then a proper restatement of the empirical findings in the GAO report would indicate no significant relation between HHI and wholesale gasoline prices. The extent of the understatement in the GAO model could only be assessed beginning with the data used in the GAO report, and applying econometric methods that cope with cross-section correlation and autocorrelation. The outcome of that exercise would indicate whether there is, in fact, a significant relation between HHI and wholesale gasoline price in the GAO model.

An overriding reservation in any interpretation of the findings in GAO (1986) is that concentration is the endogenous outcome of the same economic forces—some measured and others not—that determine the wholesale price of gasoline. The impacts of this problem are difficult to disentangle from the way the model limits the impact of systematic geographic and persistent changes over time. One might expect, for example, that smaller and more isolated markets would attract fewer wholesalers, leading to higher concentration ratios, but that the characteristics of these markets themselves imply higher costs and prices not captured in the model. Given that the error components model sweeps away such changes, and that the geographical unit of the state is not the appropriate market in the first place, it is difficult to assess either

¹⁰ State level income data are published quarterly, not monthly, by the Bureau of Economic Analysis.

¹¹ The fact that the estimates are dramatically different in the regression and error components models implies that there are omitted, time-invariant variables that are strongly correlated with persistent differences in state per capita incomes.

¹² GAO (1986), p 47.

¹³ See, for example Stock and Watson (2003), pp 253-254, or Greene (2003), Section 10.2.

¹⁴ See Davidson and MacKinnon (1993), Section 10.2, and references therein.

the extent or direction of the impact of endogenous HHI on the GAO model. The econometric problems posed by endogeneity can be addressed using instrumental variables methods. These methods presume the existence of other variables (the “instruments”), not yet included in the model, that are related to concentration but not to other unmeasured economic forces that determine the wholesale price of gasoline. These characteristics are essential to the efficacy and reliability of instrumental variables methods. However, neither the GAO report nor any of the other papers reviewed here have suggested appropriate instruments for concentration in a gasoline pricing model.

Chouinard and Perloff (2002)

Hayley Chouinard and Jeffrey Perloff utilize a cross-section time-series data set to investigate the impact of 8 producer mergers in 5 states and 27 retail mergers in 19 states on wholesale and retail gasoline prices. Like the GAO report, this study employs a reduced-form regression model estimated with monthly data (but for a longer period, January 1989 through June 1997) for the 48 contiguous states and the District of Columbia. The study utilizes only the error components form of the model.¹⁵ that is, it incorporates an indicator variable for each state.

This study incorporates a more extensive list of explanatory variables than does the GAO investigation. There are variables driving demand (per capita income, vehicles per capita, share of population in metropolitan areas, miles per gallon, speed limits), costs (crude oil current and two lags), taxes, indicators for three kinds of pollution laws, and the fraction of stations leased and company operated. The last two variables are available only at the national level, and per capita income is interpolated from annual state data and national monthly data. Most important, there is an indicator variable for each of the 35 mergers. For a given merger and any “affected state”¹⁶ the indicator variable is assigned the value zero before the completion of the merger and one in the month the merger is completed and for all months thereafter through the end of the sample; for all unaffected states the value is always zero.

The study includes these variables because its objective is to determine the relative importance of demand, costs, and market power in explaining variation in retail and wholesale gasoline prices over time and across geographic locations.¹⁷ One of its principal conclusions is, “Tax variations and mergers contribute substantially more to geographic price differentials than do price discrimination, cost factors or pollution

¹⁵ The study uses the term “fixed effects model,” because of the way the model is estimated. Error components models may be estimated as fixed effects or random effects models (Greene (2003), Sections 13.3 and 13.4). Both the GAO and Chouinard-Perloff studies use error components models with a fixed effects estimator. A fixed effects model permits error components correlated with explanatory variables whereas a random effects model does not. A fixed-effects model is presumptively more appropriate; the Chouinard-Perloff study conducts a test that confirms this fact (Chouinard and Perloff (2002), p 7 fn 7).

¹⁶ Chouinard and Perloff (2002) p 7; no exact definition of “affected” is given.

¹⁷ Chouinard and Perloff (2002) abstract.

controls.”¹⁸ But given the error components structure included in its regression model—i.e., an indicator variable for each state¹⁹—the study can say nothing about variation over geographic locations and its conclusions about geographic price differentials are unwarranted. All systematic, time-invariant differences between states are absorbed by these indicators, leaving the model to explain only the impacts of changes in explanatory variables on changes in retail and wholesale gasoline prices in each state.²⁰ The study repeatedly fails to recognize this property of error components of models.²¹

Despite this limitation, the study could, in principle, yield some useful information about the impact of changes in its demand, cost and market power measures (including mergers) on retail and wholesale gasoline prices. For example, it indicates that crude oil price changes are passed through one-for-one to both wholesale and retail prices.²² As currently written, however, the report does not list the mergers or the states affected, but provides the following summary²³ of the estimated impacts of 8 producers mergers and 27 retail mergers.

Table 1

	Sign and statistical significance of estimated merger effect				Range (Cents)
	Positive		Negative		
	Significant	Insignificant	Insignificant	Significant	
Wholesale prices					
Producer merger	1	2	3	2	20.0
Retail merger	4	17	4	2	12.0
Retail prices					
Producer merger	1	2	4	1	9.8
Retail merger	8	9	5	5	16.4

If the standard errors are reliable then these results provide some basis for concluding that mergers may have had some impact on prices. If this is the case, then it is also true that some mergers had positive impacts, and others negative, but the study provides no systematic analysis of characteristics of mergers or the market environment that account for the differing effects. Establishing the reliability of the standard errors must confront the same issues of serial and cross-section correlation that arose in the GAO study,

¹⁸ Chouinard and Perloff (2002) abstract.

¹⁹ Chouinard and Perloff (2002) p 4.

²⁰ There are a number of specific examples from Chouinard and Perloff (2002) that illustrate this point. For instance, “State specific taxes explained up to 22.2 cents of the difference in retail prices across states.” Suppose, contrary to fact, that states had different gasoline taxes but never changed them. In a fixed effects model, like the one Chouinard and Perloff use, it would then be impossible to include state tax as a covariate because it would be perfectly collinear with the indicator variables for the states. Only if some of the states changed their taxes (as they in fact did) is it possible to include state tax and estimate a coefficient for it. What the model is then capturing is not the impact of differences in taxes across states, but the impact of changes in state taxes over time.

²¹ See the discussion of the effect of different market power and demand conditions across states, p 4; the discussion of retail price effects, p 14; the discussion of mergers and retail price differentials across states, p 15; and Table 4, p 29, predicting price effects across states.

²² This may seem unsurprising; by contrast, however, the GAO study found a 183% elasticity of gasoline wholesale price with respect to crude oil price (GAO, 1986, p 46).

²³ Information taken from Chouinard and Perloff (2002) pp 7, 8, and 17.

and when these issues are resolved it is entirely possible the conclusion would be that mergers have no detectable impacts on prices. In the interim, this study provides little information about these effects.

Hastings (undated)

Justine Hastings analyzes evidence of a very different kind about the relationship between horizontal concentration and prices in the retail gasoline market. Early in 1997 Atlantic Richfield Oil Company (ARCO) announced the long-term lease of 260 Los Angeles and San Diego service stations from the independent, unbranded retail chain Thrifty. Thrifty was by far the largest independent gasoline retailer in the two metropolitan areas, and the lease involved nearly all of its stations. As an independent, unbranded retail chain, Thrifty bought from many different wholesalers, including wholesalers of branded gasoline, but product was always sold under the "Thrifty" name. The takeover by ARCO significantly increased ARCO's already major presence in southern California. About two thirds of the Thrifty stations were converted to company-operated ARCO sites, and the others became ARCO lessee-dealer operations, or dealer-owned outlets supplied either by the company or by jobbers. Regardless of ownership details, all appeared the same to the consumer.

Hastings (undated) interprets this change in ownership as a natural experiment. The study supports this interpretation by noting that station locations and characteristics were determined prior to the ARCO long-term lease agreement, and that the transition was completed in a 60-day period following its announcement with no remodeling, expansion or other facility improvements. The study controls for station-specific fixed effects, and separate time effects for each city. It utilizes the retail census of gasoline stations to determine physical and ownership characteristics of stations, and combines these data with readings of sales volumes (from pump meters) and prices (from postings). The data set used includes, in addition to the 260 leased (former) Thrifty stations, about 670 stations that compete with the former Thrifty stations, but are not themselves (former) Thrifty stations. (A station is defined as competing with a (former) Thrifty station if it is one mile or less by roadway from that station. Hastings (undated) reports the results are substantively the same with a one-half mile criterion.) There are price and volume readings from four time periods for each of these stations.

The treatment group consists of the subsample of these stations in a (former) Thrifty retail market. All others—those not in a (former) Thrifty retail market are the control group. Hastings (footnote 12) reports that the distribution of brands in the treatment and control groups is the same, consistent with the interpretation of the ARCO long-term lease as a natural experiment. However, the paper does not present the actual distributions or formal tests, as is conventional in reporting the results of natural experiments. The only other similar information is a map, showing the distribution of (former) Thrifty stations in Los

Angeles. In particular, no information is provided about the distribution of ownership in the treatment and control groups.

The study finds that, before the lease agreement and rebranding, prices at Thrifty-competing stations were from two to three cents lower than in non-competing stations, and that, after rebranding, prices at these treatment stations were from three to four cents higher than at the control stations. Controlling for time and city effects, station and market characteristics, the point estimate of the difference is 5.0 cents, and the accompanying standard error of estimate is one cent. Whether or not the newly branded ARCO station is company- or lessee-operated has no discernible impact on local competitors' prices (Table IV, p 22 and Table V, p 26). The number of competitors in a market decreased if an ARCO station was already in that market, but this does not affect price changes systematically, either (Table VII, p 29). There was a statistically significant tendency for conversion of Thrifty to ARCO to raise prices more for low-share brands than for high-share brands, in the treatment group (Table VIII, p 32). Hastings interprets these results as supporting "the hypothesis that independent competitors decrease prices through increased price competition. When they are replaced with branded competitors, in a market with consumer brand-loyalty, price competition will be softened, and equilibrium prices will increase" (Hastings (undated), p 33).

Hastings (undated) is attractive on many grounds. It identifies very specific retail markets and a very specific change in market structure. It carefully measures price²⁴ and accounts for the most important covariates. There are no apparent complications with the behavior of error terms to compromise standard errors as there are in the cross-section time-series studies. Its finding of a five-cent change in price differentials is reliable, despite the fact that it is substantial relative to retail margins. At the same time, the specificity of the study limits its implications for larger questions about horizontal competition and prices. As the author has noted, the natural experiment simultaneously removed an independent brand that traditionally competes on price, and introduced a brand that traditionally builds loyalty. It is impossible to separate the effects of these two changes in Hastings' data. In addition, it would be difficult to argue that the price change discovered in Hastings (undated) is a difference between two equilibria. The ARCO takeover was a major event in many local retail gasoline markets, and there could be further responses in those markets, including modification of capital stock, further changes in ownership, and perhaps entry and exit. These market responses might ameliorate the large price change found in Hastings (undated). A re-examination of these same retail markets, several years after the ARCO takeover, would shed valuable light on these questions.

²⁴ Posted prices were observed at specific points in time, in December 1996 and February, June, October and December 1997. Since posted retail prices can change weekly (or even daily) the Hastings (2000) data set does not contain good information on average prices at any one station. But this is incidental to the main point of the study, which is to examine differences across stations. For this comparison, it is having price observations at the same time that is important.

Manuszak (2002)

Mark Manuszak constructs a structural model in which a small number of upstream refiners sell to a large number of downstream retailers, each of whom faces less than perfectly elastic demand because consumers have preferences over retailer locations. Retailers take wholesale price as given and maximize profits. Taking into account retailers' derived demand as a function of wholesale prices, the latter are then determined in a Nash-Bertrand equilibrium. The model can be extended to allow some retail outlets to be refiner-owned.

The empirical study examines detailed retail outlet data (including station locations and characteristics, as well as prices and quantities sold broken down by station, level of service, and grade of gasoline) in Maui and Kauai (two of the Hawaiian Islands) over the period 1990 to 1995. There are no wholesale price data.

Using these data, the study estimates the parameters of the structural model. This requires fairly specific assumptions about the distribution of preferences. The paper uses state-of-the-art procedures in choosing these assumptions: many are required in order to produce demand functions that are suited to estimation, but little if anything is known about the sensitivity of results to these choices since there are few practical alternatives. The preference distributions contain unknown parameters, which the study estimates using generalized method of moments (GMM) procedures.

The structural model also incorporates the costs of the upstream refiners, which in turn partially determine wholesale prices. These costs are estimated employing an analysis of variance structure with factors for island, time, grade, service, and refiner. Estimated cost differences are substantial (those for branded wholesale on Kauai average almost twenty cents per gallon less than those for unbranded Kauai wholesale and all wholesale on Maui²⁵) and are "implausible" according to the study.²⁶

The GMM estimation procedure utilized in the paper has been employed widely in the past two decades and is well understood. It requires that the econometrician find secondary data (known as instruments) that are correlated with the observed random variables in the model, but uncorrelated with the model's unobserved random disturbances. In this study the observed random variables in the model are prices and quantities, and the unobserved random disturbances are the product-specific costs, and mean consumer valuations, of retail gasoline products.²⁷ (Each combination of station, level of service, and grade of gasoline constitutes a different gasoline product.) For example, the study assumes that the number of stations selling the same brand or product as a given station or the number of stations within a certain distance of a given station, is uncorrelated with consumers' mean preference for that station as well as that

²⁵ Manuszak (2002) Table 9.

²⁶ Manuszak (2002) p 30.

²⁷ Manuszak (2002) p 21.

station's unobserved costs.²⁸ It assumes that the level of service offered (e.g., two different service levels for the same grade of gasoline) is uncorrelated with consumer preferences.²⁹ These assumptions are highly questionable; indeed, the study emphasizes the endogeneity of location characteristics,³⁰ including the fact that locations are chosen by refiners in large part on the basis of aspects of consumer demand that are unobserved. There are conventional tests for the validity of instruments³¹, but this study did not apply them.

A compelling advantage of structural models, like this one, as opposed to reduced form models, like the other studies of horizontal market structure reviewed here, is that they provide predictions of changes in prices and consumer welfare in response to any proposed merger. Manuszak (2002) provides predicted price increases of from 1.9 cents to 3.4 cents per gallon on Maui for the merging firms in various combinations of two refiners, with negligible price changes for non-merging refiners. Point estimates of offsetting compensation for reduction in consumer welfare range from 23 to 46 cents per month per registered vehicle. (There are no standard errors presented with these estimates.) A limitation common to even the best structural models is that results like these are contingent on correct specification of the model and reliable estimates. A thorough study can mitigate this limitation by investigating the sensitivity of the main conclusions to alternative specifications of the model, and by carrying out specification tests when possible. This study, which appears to be a work in progress, has not yet taken those steps. In the interim, its conclusions—which will always need to be qualified by the fact that it pertains to a specific and arguably unique American retail gasoline market—must be regarded as quite tentative.

Evidence on vertical integration

Hastings-Gilbert (2002, national study)

Justine Hastings and Richard Gilbert attack the question of the effect of vertical integration on wholesale and retail prices using two different data sets. The hypothesis entertained in each approach—described as the “raising rivals’ costs model”—is that a vertically integrated firm that also sells to independent suppliers will post a higher wholesale price than one that is not vertically integrated. The posted wholesale price is merely an accounting entry to the vertically integrated entity, but a higher wholesale price will drive up the costs of competing retailers. As the market share of retailers integrated back to the refining stage increases, the spread between wholesale and retail prices should narrow.

²⁸ Manuszak (2002) p 23.

²⁹ Manuszak (2002) p 23.

³⁰ Manuszak (2002) p 20.

³¹ Greene (2003), Section 18.4.1.

The first approach taken in Hastings and Gilbert (2002) is one of two studies (the other being Vita (2001)) using national cross-section time-series data to investigate the impact of vertical integration on gasoline price. This study uses quarterly data from January, 1993 through June, 1997, for 26 major metropolitan areas. It employs as its dependent variable the spread between the average unbranded wholesale price and the crude oil spot price.³² The explanatory variables include four measures of vertical integration – the number and market share, respectively, of vertically integrated suppliers, and of unintegrated suppliers. These measures are assembled from retail census data, but the study provides only a sketchy description of how the variables are actually constructed. It documents considerable variation in these measures from one metropolitan area to the next.

Unfortunately, these are the only explanatory variables used in the Hastings and Gilbert (2002) cross-section time-series study.³³ This is in marked contrast to other longitudinal studies of wholesale or retail gasoline prices, and there are two notable consequences of this omission. One is that the goodness of fit of the Hastings-Gilbert model³⁴ is poorer than in other longitudinal studies. The second, and more important, consequence is that omitted variables (for example, those that drive demand) may well be correlated with the measures of vertical integration, resulting in biased estimates of the vertical integration variable coefficients.

The Hastings and Gilbert (2002) study provides regression model estimates with and without error components (the latter incorporated as fixed effects). Without error components, all four measures of vertical integration are statistically significant: increases in the number of vertically integrated and unintegrated suppliers both decrease the price spread, while increases in market share increase price spread in the case of integrated suppliers and decrease price spread for independent retailers. The effect appears to be two cents or less, moving from the first to third quartile.³⁵ With error components, three of the four coefficients are reduced in magnitude and are statistically insignificant. The components are jointly highly significant, indicating probably omitted variables bias in the model estimated without error components. The longitudinal study in Hastings and Gilbert (2002) does not, therefore, provide much evidence on the impact of vertical organization on price spreads.

³² If changes in crude oil prices are passed through to wholesale prices one-for-one within a three month period (as is the case in Chouinard and Perloff (2002)) then the results in Hastings and Gilbert (2002) should be about the same using wholesale price as the dependent variable so long as crude oil price is included as an explanatory variable.

³³ Although the study offers no explanation, this may be due to the difficulty of organizing covariates by metropolitan area.

³⁴ See Hastings and Gilbert (2002), Table 4.

³⁵ This calculation entails matching variables in Tables 1 and 4 of Hastings and Gilbert (2002), which is not straightforward due to different nomenclature in the two tables.

Hastings and Gilbert (2002, local study)

The second study in Hastings and Gilbert (2002) examines the impact of the sale of Unocal's west coast refining and marketing assets to Tosco in November 1996. This transaction had negligible impact on horizontal concentration at the retail level, because there were no cities in which both companies had a significant presence. Since Tosco sold wholesale gasoline in all west coast markets whereas Unocal sold only in some, concentration increased in some wholesale markets but not in others. The raising rivals' costs model implies that Tosco should increase wholesale prices more in markets where it gains a greater downstream market share, and in markets where there is heavier competition from independent retailers.

Using retail census data, Hastings and Gilbert (2002) constructs the product of the increased downstream market share from the acquisition of Unocal retail outlets (i.e., Unocal's pre-acquisition share of the downstream market), and the percentage of those outlets that are geographically within one mile of an independent retailer.³⁶ This measure is one covariate in a regression equation explaining Tosco's weekly average wholesale price of unbranded gasoline. The other covariates are the number of refiners selling unbranded gasoline, the percentage of stations that are independent retailers, and Tosco's weekly average wholesale price in Phoenix. The latter variable is included as a proxy for costs: Phoenix terminals are supplied by common carrier pipeline from Los Angeles, and Phoenix was unaffected by the Unocal acquisition – in fact it had no change in upstream or downstream structure during the time period considered.

The structure of the data set is cross-section time-series, employing 12 cities and 128 weeks (July, 1996 through December, 1998). It utilizes an error components structure, and the usual tests favor fixed effects rather than random effects. The model also incorporates first-order autocorrelation. It does not accommodate correlation between shocks to wholesale prices in different cities, and examination of this feature would render standard errors more reliable – but, as it stands, the model represents a reasonably careful application of statistical inference using cross-section time-series.

The only significant covariate is the product of increased downstream market share and percentage of outlets competing with independents, with a coefficient estimate of about 0.4. The study interprets this finding as implying that “an integrated refiner's [wholesale] price is an increasing function of its competition with independent retailers.”³⁷ This interpretation must be made cautiously, however, because

³⁶ It appears that this covariate is zero before the acquisition in all cities (see Hastings and Gilbert (2000) pp 15-16 and Table 5) but the study is not completely explicit on this point. This definition of the variable imparts changes in time that vary substantially across cities (Table 5) and are not absorbed in the variance components of the model.

³⁷ Hastings and Gilbert (2002), p 20. For example, given an increase in market share of 0.08, the median value in the sample, Tosco would increase wholesale prices by 0.7 cents if 20% of its acquired retail outlets were within a mile of an independent competitor, and by 1.7 cents if 50% of its retail outlets were so

in the sample there is greater variation in the increase in market share than there is in competition with rival independent retailers.³⁸ The raising rivals' costs hypothesis points directly to the interaction of Tosco's market share and the percentage of Tosco's outlets competing with independents as the driver of Tosco's wholesale prices. In the specification of the regression model used in Hastings and Gilbert (2000) this is the only way that Tosco's market share can affect wholesale prices, and one could interpret the results as simply indicating that Tosco's wholesale prices are an increasing function of its market share – i.e., the results simply reflect the change in Tosco's wholesale demand schedule. By entering market share separately in the regression Hastings and Gilbert could have used a simple *t*-test to sort out these competing interpretations.

Vita (2000)

State gasoline “divorcement” statutes restrict and in some cases proscribe the vertical integration of gasoline refiners and retailers. Divorcement laws are currently in effect in six states (Connecticut, Delaware, Hawaii, Maryland, Nevada and Virginia) and the District of Columbia, and have been proposed, at one time or another, in most state legislatures. To the extent these laws can be regarded as natural experiments, there are rich possibilities for empirical investigation of the impact of vertical integration on retail prices.

Historically, franchised retail gasoline dealers have supported divorcement legislation as a means of preventing predation by refiner-owned service stations. It is indeed the case that retail prices at refiner-owned stations are systematically lower than those at lessee or independent dealers,³⁹ but there is no economic case for predation of refiners upon their efficient dealers.⁴⁰ Simple microeconomics suggests that divorcement statutes may indeed increase retail prices, because of the efficiencies of integration and double marginalization in the determination of the wholesale prices paid and retail prices set by lessee or independent branded retailers. The case for divorcement statutes lowering retail prices is relatively weak, relying on specific forms of imperfect competition and obtaining at most ambiguous impacts on retail prices.⁴¹

Vita (2000) undertakes a cross-section time-series study utilizing monthly data for all 50 states, January 1995 through December 1997. The dependent variable is the retail price of unleaded regular gasoline net of taxes. The covariate of interest is an indicator for a divorcement law. Because such laws were in effect,

situated. The largest predicted price increase in the sample is for Santa Barbara, amounting to about 3.6 cents per gallon.

³⁸ Hastings and Gilbert (2002), Table 5.

³⁹ See Barron and Umbeck (1984) and Shepard (1993), discussed subsequently.

⁴⁰ See Vita (2000) pp 217-218.

⁴¹ See Salinger (1988), Ordovery et al. (1990), Hart and Tirole (1991), Reiffen and Vita (1995).

or not, for the entire 1995-1997 period in each of the states, it is impossible to use a variance components model and estimate the impact of divorce. This being the case, it is important to specify as completely as reasonably possible a set of covariates explaining systematic state-by-state differences in retail gasoline price. To this end the study incorporates demand shifters (income, driver and vehicle characteristics), cost shifters (wages, transportation prices, crude oil price, gasoline characteristics arising from environmental legislation, heating degree days, and regional transportation cost indicators for the northeast, US west of the divide, Alaska and Hawaii), yearly and monthly indicators. It also includes as covariates an indicator for the presence of a sales-below-cost law and the percentage of sales through self-service. The exclusion of the District of Columbia and the presence of the cost indicator for Hawaii implies that inference about the impact of divorce laws is being drawn from five states – Connecticut, Delaware, Maryland, Nevada and Virginia.

The base model is estimated by least squares correcting for first-order autocorrelation. The main finding is that divorce laws increase retail prices by 2.6 cents (95% confidence interval 1.2 cents to 4.0 cents). The estimation method presumes that all covariates, including divorce laws, are exogenous. The suitability of this assumption is typically a central question in “natural experiment” studies like this one. As a check on assumptions, Vita (2000) also estimates the model taking both divorce and sales-below-cost laws to be endogenous, and uses as instrumental variables the presence of state anti-takeover or minimum wage laws, the ADA rating of the state Congressional delegation, and an indicator for whether a large gasoline refiner is located in the state. Under this alternative set of assumptions the study finds that divorce laws increase retail prices by 3.1 cents (95% confidence interval 0.7 cents to 5.5 cents). Prohibition of self-service sales (Oregon and Nevada) increases price by about 3.5 cents per gallon, while sales-below-cost laws have no statistically significant impact on price.

Vita (2000) is a well-executed time-series cross-section study that focuses on the impact of divorce laws on the retail gasoline price in the states where they have been instituted. It finds this impact has been two-and-one-half cents, and it substantiates its interpretation of these laws as a natural experiment from which we may learn about the impact of the degree of vertical integration on retail gasoline prices. This study, and studies like it that utilize differences across states to identify the impact of horizontal or vertical organization on price, are reliable to the extent that they are able to identify all of the relevant factors accounting for state price differentials. Vita (2000) explains about 85% of the variation in gasoline prices, which is better than the GAO (1986) regression model that explained about 80%, but falls short of the variance components models in Chouinard and Perloff (2002) at 90% and GAO (1986) at 96%.

The increased explanatory power in the latter studies comes from the inclusion of state indicator variables. These indicator variables account for state-specific covariates that are important for gasoline prices but have not been included in the regressions in Chouinard and Perloff (2002), GAO (1986) or Vita (2000).

(Vita (2000) cannot include state indicators because they would be perfectly collinear with the divorce indicator.) The implication is that there are unaccounted factors underlying variance across states in all of these studies, factors that may be correlated with the presence of divorce laws and consequently bias the estimate of the impact of these laws. An important goal in future econometric cross-section time-series studies of this topic is reduction of the variation in prices that must be laid to unexplained differences in states (as is done in variance components models). Vita (2000) shows that one way to do this is through more deliberate accounting of cost differences

Barron and Umbeck (1984)

John Barron and John Umbeck utilize the Maryland divorce law, which took effect in 1979, as a natural experiment whose outcome provides the impact of divorce on retail gasoline prices. The theoretical foundation for the expected impact is a comparison of prices and hours of retail outlet operation when retail stations are company-owned, with these decisions when retail outlets purchase from refiners and set their own retail prices and hours of operation. The textbook double marginalization model implies that prices will be higher and hours shorter, given separation of upstream and downstream decision making.

As a consequence of the Maryland divorce legislation, seven refiners were forced to sell or franchise a total of 170 stations, mostly during June or July 1979. Barron and Umbeck (1986) collected primary data from the seven refiners on hours of operation and retail prices at the affected stations, and at stations the refiners perceived as competitors of these stations. (Typically three or four such stations were identified for each formerly owned station.) They received responses for 144 stations and utilized 99 franchised stations, and their corresponding competitors, in their analysis.⁴²

The study measures retail price relative to the BLS U.S. monthly gasoline price index, separately for self-serve and full-service gasoline. Table 2 summarizes the empirical findings in Barron and Umbeck (1986). The top panel indicates comparisons using the price and hours data directly; prices are expressed in real 1981 cents. The bottom panel provides comparisons controlling for station characteristics (service bay, convenience store, car wash, both full- and self-service, number of hoses, and acceptance of credit cards) and the number of stations identified by the leasing refiner as being in the market. The results are all consistent with the implications of the double marginalization model that company owned stations price lower than franchise or independent stations, and that divorce will increase prices at both affected and competing stations. Affected stations reduce hours while competing stations have little or no discernible

⁴² The alternative to franchise was outright sale, which changed the station brand. This consideration, plus the fact that many companies did not retain data for stations sold, led Barron and Umbeck (1986) to confine the study to franchised stations.

change in hours; however the comparison of hours between affected and competing stations depends on whether covariates are included.⁴³

Table 2^a

	Affected station before	Affected station after	Competing station
No controls:			
Full service price	-9.51	-2.86	+1.01
Self-serve price	-3.92	-2.53	+0.03
Hours	+13.76	+4.54	+2.99
With controls:			
Full service price	-5.55	-0.26	+1.01
Self-serve price	-2.42	-0.71	+0.30
Hours	-3.93	-12.31	+1.17

^aPrices and hours are all shown relative to unaffected competing stations before divorcement. Standard error of estimate ranges from 0.25 to 0.40 for prices and 2 to 3 for hours.

Vita (2000) found that divorcement laws raise prices 2.6 cents per gallon, give or take 1.4 cents for a 95% confidence interval. In Maryland, 170 stations were affected by divorcement, there were about 600 competing stations⁴⁴ as defined in Barron and Umbeck (1984), and there were roughly 930 unaffected non-competing stations – a total of 1700 stations state-wide. We do not know the fraction of Maryland sales at company-owned stations; if it was only ten percent, then the study implies an increase of about 1 cent in full service and 0.2 cents in self-service, but if it was half then the increase is about 3 to 3.5 cents. Given that company owned stations are typically substantially larger than other stations,⁴⁵ and taking into account estimation error, the results in Vita (2000) and Barron and Umbeck (1984) are consistent. This is notable in view of the fact that Vita (2000) is based on state differences in the mid-1990s, whereas Barron and Umbeck (1984) is grounded in changes over time in one state in the late 1970s.

Shepard (1993)

Andrea Shepard provides some evidence on the relationship between the ownership of a gasoline station and retail gasoline price.⁴⁶ There are three kinds of ownership: open dealer contracts in which the station is owned and operated by an independent dealer who contracts with a refiner for gasoline supply; lessee dealer stations in which the capital at the station is owned by the refiner and leased by the self-employed dealer; and company outlets at which the capital is owned by the refiner and the station operator is employed by the refiner. The operator sets prices at open and lessee dealerships, the refiner under company ownership; the operator controls station characteristics in an open dealership, the refiner under a lessee dealership or company ownership.

⁴³ Barron and Umbeck (2000) takes no note of this discrepancy in findings.

⁴⁴ Barron and Umbeck (2000) p 321,

⁴⁵ Blass and Carlton (2001).

⁴⁶ An earlier summary version of this paper is Shepard (1990).

The principal focus of Shepard (1993) is on the relationship between station characteristics and ownership form. This is part of a larger literature (including Slade (1998) and Taylor (2000)) that uses classical principal-agent theory to optimize the contract between owner and operator. A secondary focus of this study is the relationship between ownership form and price, which falls within the purview of this survey. It is based on a primary data set collected from a cross-section of 924 branded gasoline stations in eastern Massachusetts in the first quarter of 1987. Of these stations, 38 were company owned, 452 were lessee-dealer and 434 were open dealer. In the empirical work on price, Shepard (1993) distinguishes only between company-owned and all other stations.

The study utilizes a conventional regression equation. The covariate of interest is an indicator variable for company ownership. The other covariates capture station characteristics (repair services, convenience store, number of cars that can be serviced simultaneously, whether both full and self-service are offered, whether the station was recently remodeled, whether it offers "mini-service", and an indicator for outlying area) and a "nearby capacity" variable that sums the number of cars that can be served simultaneously at other stations located within a one-mile radius. Six variants of the regression model are presented, corresponding to the six combinations of full- and self-service, and to leaded, unleaded regular, and unleaded premium gasoline. Across these six equations, five of the company ownership coefficients are statistically insignificant (ranging from -3.17 to 0.73); only that for full-service unleaded premium, -5.47 cents, is barely significant. The median estimate is -1.5 cents.

Shepard (1993) also presents a variant on this equation that is intended to control for local geographic effects. The documentation of this procedure is too sketchy to permit critical evaluation. The pattern of inconsistent and insignificant results is repeated in this variant.

The findings in Shepard (1993) are consistent with the findings in Vita (2000) and Barron and Umbeck (1984) that retail prices are lower at stations vertically integrated back to the refining stage. However, the estimates in the study are sufficiently imprecise (most likely due to the small data set) that it adds little information about the magnitude of the difference.

Conclusion

The empirical evidence on the competitive effects of mergers in the gasoline industry consistently supports the proposition that retail prices are lower with vertical integration than with separation of refining and retailing. State divorce laws provide a plausible natural experiment measuring this price differential, and two well-executed independent studies utilizing different data and methods (Barron and Umbeck (1984) and Vita (2000)) indicate that it is substantial. Taken together, these studies support the proposition that retail gasoline prices at vertically integrated stations are from 1.5 cents to 5.0 cents lower than at leased

or independent stations, other things equal, and that prices at competing stations are also lower. The evidence from local posted prices in Shepard (1993) is much weaker, but is consistent with these estimates (as well as a much wider range). The Hastings and Gilbert (2002) national study examines the broader question of the relationship between various measures of vertical integration and retail gasoline prices, but fails to control for known, important retail price covariates, and obtains inaccurate estimates in any event. The Hastings and Gilbert (2002) local study addresses the change in wholesale prices following a single acquisition that increased vertical integration with little effect on horizontal retail concentration. It raises the possibility of slightly increased wholesale prices, but does not examine retail prices.

There is little, if any, reliable evidence on the competitive effects of horizontal mergers or concentration. One study (Hastings (undated)) plausibly interprets the conversion of independent to branded retail outlets in Southern California as a natural experiment. It finds that at stations competing with the stations that changed hands, prices increased—by three to seven cents—relative to prices generally, and without regard to ownership status of stations, in the five months following the transaction. This is a very large change, relative to retail-wholesale price differentials. However, Hastings (undated) does not track subsequent price changes or market adjustments. Another study (Manuszak (2002)) projects the impacts of mergers in one of the Hawaiian Islands in the context of a structural model that makes specific assumptions about preferences and costs. These assumptions could have been tested but were not, and the study itself regards the findings for costs as implausible.

Two studies utilize cross-section time series reduced form models. GAO (1986) examines the impact of the HHI measure of horizontal wholesale concentration on wholesale prices, using less than three years of monthly state data. Chouinard and Perloff (2002) estimates the impact of 35 different mergers on both wholesale and retail prices, using over eight years of monthly state data. The statistical methods used in both cases assume that variations in concentration (across both states and time) can be regarded as a “natural experiment” – for example, GAO (1986) implicitly assumes that concentration is uncorrelated with differential distribution costs that are not recorded, and Chouinard and Perloff (2002) requires that mergers not be related systematically to unexplained wholesale price differentials or movements over time. Both studies control for demand and cost shifting covariates that would plausibly enter a reduced form price model. In any event, taken at face value, Chouinard and Perloff (2002) finds no systematic relationship between mergers and prices. GAO (1986) reports a very weak link between concentration and prices, with wholesale prices rising by 0.5% (1.7% in the error components model) if the HHI increases from 2000 to 3000.

Given the importance and prominence of the motivating issues of economic policy, the dearth of reliable evidence on the competitive effects of horizontal mergers or concentration is unfortunate. Retail and wholesale gasoline markets differ substantially across time and place with respect to horizontal

concentration, entry, exit, and the impact of mergers, but to date no study has utilized good data or state-of-the-art methodology to examine this evidence. GAO (1986) and Chouinard and Perloff (2002) utilize data organized by state. State data are more readily available, but they are inconsistent with markets that tend to be organized around metropolitan areas, and around terminal clusters in metropolitan areas in particular.⁴⁷ Hastings and Gilbert (2002) constructed wholesale gasoline prices by metropolitan area, but apparently did not construct (and certainly did not use) most of the covariates that turn out to be important in other studies. Organization of data along market rather than state lines is an important topic in future research.

The failure of any study to appropriately utilize econometric methods for cross-section data that could be found in textbooks even at the time of GAO (1986) is harder to understand. Both GAO (1986) and Chouinard and Perloff (2002) draw implications that have no foundations given those studies' use of error components models, as detailed above. An ambitious yet worthy goal is to construct a cross-section time-series data base sufficiently rich in covariates that it accounts for geographic differentials in wholesale and retail prices of gasoline, so that cross-section time-series studies do not have to resort to error components methods (either fixed effects or random effects). Vita (2000) takes a step in this direction, and such a data base would be useful in studying not only the competitive effects of horizontal concentration and vertical integration, but also the vexing and politically relevant question of seemingly persistent differences in prices in different geographic regions. The implications of the petroleum physical distribution network for costs, incorporated if at all only with a few regional dummies in these studies, are well worth working out.

Better data and appropriate use of models notwithstanding, it is unlikely that relationships between horizontal concentration and vertical integration, on the one hand, and retail or wholesale prices of gasoline, on the other, will be self-evident. At the end of the day many questions will come down to the reliability of standard errors and hypothesis tests: an estimated price increase of five cents plus or minus one cent is not the same as an estimated price increase of five cents plus or minus ten cents, and so it is important to know whether the "one cent" or the "ten cents" (if either) is appropriate. This requires taking seriously properties of statistical models (like serial correlation) that are either ignored or simply go unreported in most of these studies. The natural sciences have long-established standards for full documentation and replicability. There is no reason why the same standards of evidence should not apply here.

⁴⁷ GAO (1986, p 40) recognizes this problem.

References

- Barron, J.M. and J.R. Umbeck (1984), "The Effects of Different Contractual Arrangements: The Case of Retail Gasoline Markets," *Journal of Law and Economics* 27: 313-328.
- Blass, A.A. and D.W. Carlton (2001), "The Choice of Organizational Form in Gasoline Retailing and the Cost of Laws that Limit that Choice," *Journal of Law and Economics* 44: 511-524.
- Carlton, D.W., and J.M. Perloff (2000), *Modern Industrial Organization*. Addison-Wesley.
- Chouinard, H. and J. M. Perloff (2002), "Gasoline Price Differences: Taxes, Pollution Regulations, Mergers, Market Power, and Market Conditions," working paper (September).
- Davidson, R., and J.H. MacKinnon (1993), *Estimation and Inference in Econometrics*. Oxford University Press.
- General Accounting Office (1986) "Energy Prices: Gasoline Price Increases in Early 1985 Interrupted Previous Trend," GAO/RCED-86-165BR (September).
- Greene, W.H. (2003), *Econometric Analysis* (Fifth Edition). Prentice-Hall.
- Hart, O., and J. Tirole (1990), "Vertical Integration and Market Foreclosure," *Brookings Papers on Economic Activity* 205-286.
- Hastings, J.S. (undated), "Vertical Relationships and Competition in Retail Gasoline Markets," working paper (<http://www.dartmouth.edu/~jhasting/Hastings1.pdf>)
- Hastings, J.S. and R. Gilbert (2002), "Market Power, Vertical Integration, and the Wholesale Price of Gasoline," working paper (<http://www.dartmouth.edu/~jhasting/Hastings2.pdf>)
- Manuszak, M.D. (2002), "The Impact of Upstream Mergers on Retail Gasoline Markets," September working paper.
- Ordober, J., et al. (1990), "Equilibrium Vertical Foreclosure," *American Economic Review* 80:917-942.
- Reiffen, D., and M. Vita (1995), "Is There New Thinking on Vertical Mergers?" *Anitrust Law Journal* 63: 917-941.
- Salinger, M. (1988), "Vertical Mergers and Market Foreclosures," *Quarterly Journal of Economics* 103: 345-356.
- Shepard, A. (1990), "Pricing Behavior and Vertical Contracts in Retail Markets," *American Economic Review* 80 (proceedings issue): 427-421.
- Shepard, A. (1993), "Contractual Form, Retail Price, and Asset Characteristics in Gasoline Retailing," *The RAND Journal of Economics*, 24: 58-77.
- Stock, J.H. and M.W. Watson (2003), *Introduction to Econometrics*. Addison-Wesley.
- Taylor, B.A. (2000), "Retail Characteristics and Ownership Structure," *Small Business Economics* 12: 157-164.
- Vita, M.G. (2000), "Regulatory Restrictions on Vertical Integration and Control: The Competitive Impact of Gasoline Divorcement Policies," *Journal of Regulatory Economics* 18:3 217-233.