

Using Ex Post Evaluations to Improve the Performance of Competition Policy Authorities

William E. Kovacic*

I. INTRODUCTION	505
II. THE EVALUATION OF COMPETITION AGENCIES: TWO FORMS OF PERFORMANCE MEASUREMENT	506
III. THE RATIONALE FOR EX POST EVALUATION IN MAKING COMPETITION POLICY ...	508
A. <i>Uncertainty: Competition Policy as Experimentation</i>	508
B. <i>Limited Transparency: The Impact of Policymaking by Settlement</i>	510
C. <i>The Evolutionary Quality of Antitrust Jurisprudence and Enforcement Policy</i>	511
D. <i>Institutional Multiplicity: Concurrent Oversight</i>	512
E. <i>Institutional Multiplicity: Benchmarking Operational Processes</i>	513
F. <i>The Value of Competition Agency Participation in Performance Management</i>	513
IV. EXPERIENCE WITH PERFORMANCE EVALUATION OF COMPETITION AGENCIES	515
A. <i>General External Studies of Competition Agency Performance and Evaluation Efforts</i>	516
1. <i>Studies by Individual Researchers or Small Teams of Researchers</i>	516
2. <i>Studies by Other Government Institutions</i>	517
3. <i>Studies by Blue Ribbon Panels</i>	517
4. <i>Agencywide or Systemwide Evaluations by Multinational Organizations</i>	518
5. <i>General Assessments of Competition Agencies by Outsiders: Strengths and Weaknesses</i>	519
B. <i>Case-Specific Evaluations of Competition Agency Intervention</i>	520
1. <i>Studies Performed by Outside Researchers Without Competition Agency Involvement</i>	520
2. <i>Studies by Other Government Bodies: GAO Study of U.S. Petroleum Mergers</i>	521

* Commissioner, Federal Trade Commission. The author thanks Paul Maylon, Joe Phillips, and Patricia Heriard-Dubreuil for many helpful comments and suggestions about the preparation of this Article. The author also is grateful to the Organization for Economic Cooperation and Development, which funded the preparation of an earlier version of this Article while the author was a member of the faculty of the George Washington University Law School. The views expressed here are the author's alone and not necessarily those of the Federal Trade Commission or any individual commissioner.

3. <i>Case-Specific Reviews Performed by Multinational Bodies</i>	522
4. <i>Case-Specific Reviews Undertaken or Sponsored by Competition Agencies</i>	522
5. <i>The FTC Vertical Restraints and Abuse of Dominance Impact Evaluations</i>	524
C. <i>Process-Oriented Evaluations of Competition Agency Operations and Organization</i>	527
1. <i>Process-Based Evaluations by Competition Agencies</i>	527
2. <i>Process-Oriented Evaluations by Multinational Bodies</i>	531
3. <i>The OECD Performance Management Initiative</i>	531
V. <i>CONSTRUCTING A METHODOLOGY FOR CONDUCTING PERFORMANCE EVALUATIONS</i>	533
A. <i>Constructing an Evaluation Methodology</i>	535
1. <i>Choosing a Process for Designing a Performance Measurement Methodology</i>	535
a. <i>Preliminary Discussions within the Competition Agency</i>	535
b. <i>Consultations with Expert Outsiders</i>	535
c. <i>Budgeting</i>	537
d. <i>Identifying Restrictions on the Use and Dissemination of Confidential Records</i>	537
2. <i>Conducting Ex Post Assessments: Using Insiders, Outsiders, or Both</i>	538
3. <i>Disclosure of Results</i>	539
B. <i>Special Considerations for Process-Oriented Performance Measurement</i>	540
1. <i>Mechanism for Strategic Planning</i>	540
2. <i>Maintenance and Disclosure of Databases</i>	541
3. <i>Explanation of Actions Taken and Not Taken</i>	541
4. <i>Assessment of Human Capital</i>	541
5. <i>Investments in Competition Policy Research and Development</i>	542
6. <i>Recognition of Policymaking Interdependencies</i>	542
7. <i>Benefits of Comparative Study</i>	542
C. <i>Special Considerations in Case-Specific Evaluations</i>	543
1. <i>Reviewing a Collection of Matters Rather than Analyzing Single Cases</i>	543
2. <i>Understanding Enforcement Choices in Context</i>	543
3. <i>Sensitivity to Institutional Capabilities and Long-Term Consequences</i>	544
D. <i>Case-Specific Evaluations: Building Mandates into Settlement Decrees</i>	544
VI. <i>CONCLUSION</i>	545

I. INTRODUCTION

Competition policy is a work in progress. Since the late 19th century, the history of competition law has featured a continuing search for optimal statutory commands, institutional designs, and operational techniques.¹ Recent decades have witnessed extraordinary change as existing competition systems seek better practices for defining substantive commands, setting priorities, choosing cases, and selecting remedies. This Article suggests that charting the future course of competition policy can benefit heavily from looking back and asking two fundamental questions. First, did the agency's interventions produce good results? Second, did the agency's managerial processes help ensure that the agency selected initiatives that would yield good outcomes? By assessing the quality of its substantive interventions and internal procedures, a public agency can gain valuable insights about how to improve its performance. Even the effort to define performance measures can impose valuable discipline on an agency's allocation of resources and help identify possibilities for improvement.

Encouraging government bodies to conduct performance evaluations is not only good public policy, it is likely to be a key ingredient of future attempts by competition authorities to demonstrate the value of competition law to broader audiences. In existing competition systems, as well as in jurisdictions considering the adoption of a competition law, a number of observers decline to assume as a matter of theory, or accept as an article of faith, that the enforcement of a competition law yields socially useful results. An a priori presumption of efficacy is a weak substitute for a systematic assessment of outcomes.²

This Article discusses how government competition authorities might use ex post evaluations of enforcement decisions, operational mechanisms, and organizational design to improve the quality of their work. The Article examines the value of ex post analyses of previous public enforcement actions as follows. Part II describes two basic forms of performance measurement that competition agencies might undertake to improve the quality of their work. Part III identifies the rationales for establishing and applying performance measures to evaluate the effectiveness of the substantive interventions and operational procedures of competition policy agencies. Part IV discusses the past experience of competition agencies in conducting ex post analyses of substantive outputs and operational procedures. Although not a comprehensive account of agency experience, Part IV uses noteworthy examples to provide a context for the proposals in Part V that present methodologies for performing evaluations.

1. JONATHAN BAKER, ANDREW I. GAVIL & WILLIAM E. KOVACIC, *ANTITRUST LAW IN PERSPECTIVE: CASES, CONCEPTS AND PROBLEMS IN COMPETITION POLICY* 923-30 (2002).

2. See Robert Pitofsky, *Roundtable Conference with Enforcement Officials*, 68 *ANTITRUST L.J.* 581, 593-95 (2000) (Phillip A. Proger, Moderator) (stating that "unexamined enforcement policy is not an appropriate way to proceed" and observing that antitrust agencies "virtually have an obligation, not just to enforce the law, but to take a look at the consequences of [their] enforcement").

II. THE EVALUATION OF COMPETITION AGENCIES: TWO FORMS OF PERFORMANCE MEASUREMENT

This Article addresses two basic ways that competition agencies can measure the quality of their performance. The first is to evaluate the contribution of the agency's outputs, such as the prosecution of cases or the commitment of resources to competition advocacy, to the attainment of the goals embodied in the jurisdiction's competition laws. Suppose, for example, that the jurisdiction's competition law seeks to improve economic performance by proscribing anticompetitive mergers, improper exclusion by dominant firms, or naked agreements among rivals to set prices or other terms of trade. An outcome-oriented program of evaluation would seek to determine how much the agency's pursuit of specific initiatives helped achieve these ends.

A second approach to evaluation focuses on process. In lieu of, or in addition to, evaluating outcomes, an evaluation program might seek to assess the quality of the competition agency's internal operations—the mix of managerial methods and organizational choices that determine how the agency allocates and applies its resources. This approach treats management and organization as critical inputs into the implementation of competition policy and seeks to identify improvements in how the competition agency operates. The logic is that progress toward superior managerial and organizational techniques will increase the likelihood that the agency's substantive outputs generally promote the realization of the competition law's objectives.

In striving to cope with the crisis of the moment or the press of new business, public (and private) institutions might be tempted to regard performance evaluation as a purely optional and discretionary undertaking. The need for a competition agency to address current exigencies—for example, to cope with an unusually large number of merger filings—can encourage acceptance of the view that the expenditure of effort to evaluate substantive outputs or managerial inputs is an unaffordable luxury or a costly “diversion” of resources away from the treatment of immediate operational needs. In this frame of mind, an agency's managers may reason that they can *either* support current operations *or* conduct ex post assessments, but not both.

To perceive that performance measurement comes at the expense of fulfilling current operational obligations misconceives the role of evaluation. As a large and growing literature demonstrates, performance measurement is a necessary ingredient of good practice for public institutions.³ From this perspective, strong mechanisms for evaluating outcomes and decisionmaking processes are not detached from, but instead are integral to, the creation of sound policy outputs.⁴ Outlays for performance measurement by a public institution are no more “discretionary” than the effort that physicians exert after performing surgery to meet with their patients and determine whether the surgery improved the patient's health.

3. See JODY ZALL KUSEK & RAY C. RIST, TEN STEPS TO A RESULTS-BASED MONITORING AND EVALUATION SYSTEM 12-22 (2004) (discussing the importance of results-oriented evaluation to effective public administration); JAMES Q. WILSON, BUREAUCRACY 373-75 (1989) (describing the policy benefits of having government bodies evaluate the results of their programs).

4. See, e.g., Deborah Platt Majoras, Remarks Before the Charles River Associates: A Dose of Our Own Medicine: Applying a Cost/Benefit Analysis to the FTC's Advocacy Program 3 (Feb. 8, 2005), available at <http://www.ftc.gov/speeches/majoras/050208currebttopics.pdf>.

Properly understood, performance measurement is not an institutional burden but instead is a valuable asset. It enables an agency to tap its base of experience and borrow from the experience of peer institutions to improve the quality of its work. A routine of retrospective assessments can help agency managers engage in what two scholars of public administration, Richard Neustadt and Ernest May, have called “thinking in time streams,” which is applying “the kind of mental ability that readily connects discrete phenomena over time and repeatedly checks connections.”⁵ Robust performance measurement tools can help answer the three fundamental questions that Neustadt and May have identified as the core concerns of public agency decisionmaking: “‘Will it work?’ ‘Will it stick?’ ‘Will it help more than it hurts?’”⁶

As noted above, two important focal points for ex post evaluation are the substantive outputs of the competition agency (cases and advocacy interventions) and the operational processes (strategic planning and case handling) that generate such outputs. The literature on performance evaluation notes the tendency of public and private institutions to conduct detailed ex post assessments solely or largely as a desperate and necessary response to an operational calamity. Such evaluations unmistakably play a valuable role in improving agency performance in the short and long term. For example, a detailed reconstruction and analysis of a serious operational failure can identify weaknesses in specific substantive projects or in an institution’s decisionmaking processes, and can foster adjustments in the institution’s projects or the establishment of formal rules and cultural norms that reduce the possibility of future disasters.⁷

A custom of conducting ex post assessments only in response to grave operational failures is a seriously incomplete form of performance measurement. A superior approach is to examine programs on an ongoing basis and not simply in response to dramatic failures. A well conceived evaluation mechanism would study programs that appear, at any single moment, to be working well. Such initiatives can serve to identify possible improvements that will make an acceptable or commendable program truly excellent. From a more defensive perspective, a routine of unflinching assessment can spot flaws that, if uncorrected, could lead to a major policy breakdown. When success is achieved by narrow margins or gained by sheer luck, an organization may overestimate its skill and shave the margin of error to a dangerous thinness that eventually begets failure. To make no effort to assess outcomes and the causes of outcomes is a form of flying blind.⁸

5. RICHARD NEUSTADT & ERNEST MAY, THINKING IN TIME 252-53 (1986). Professors Neustadt and May observe that this type of thinking “is a special style of approaching choices, more the planner’s or the long-term program manager’s than the lawyer’s or judge’s or consultant’s or trouble-shooter’s.” *Id.*

6. *Id.* at 270.

7. One example is the effort by De Havilland to determine the cause of crashes of the world’s first commercial jet airliner, the Comet, in the early 1950s. RAY WILLIAMS, IMAGES OF AVIATION: COMET AND NIMROD 15-16 (2000). De Havilland conducted extensive research on a Comet in its inventory, identified a latent design flaw that caused the crashes, and made the results of its research available to the world’s aviation community. *Id.* De Havilland’s work in identifying the cause of the disasters and in publicizing the results of his evaluation facilitated significant improvements in the design of commercial jet transports. *Id.*

8. In his recent analysis of the business practices of a professional sport league, Michael Lewis quotes one observer who captures the problems of an organization that cannot, or will not, develop good measures of performance: “[The sport franchise executives] aren’t equipped to evaluate their own systems. They don’t have the mechanism to let in the good and get rid of the bad. They either keep everything or get rid of everything, and they rarely do the latter.” MICHAEL LEWIS, MONEYBALL 241 (2004).

Few institutions, public or private, enjoy subjecting themselves to rigorous assessments of their work. Good performance measures can be hard to design, and the cost and indeterminacy of trying to formulate truly informative measures can be daunting. When pressed to devise performance measures, some institutions have embraced badly flawed criteria simply because the activities they chose to observe were measurable—an approach no less delusional than disdaining all measurement until the roof caves in. Ask the wrong questions, and you can get spurious answers.

Selecting good performance measures unquestionably can be a difficult task, and the worth of an evaluation exercise depends heavily on the validity of the chosen criteria. At the same time, one should be skeptical of claims that a given activity is so resistant to measurement that one must rely entirely upon the actor's untested (and, the actor might say, untestable) representations about the quality of the actor's work. Even if measurement problems are significant and turn out to be overwhelming, the very process of seeking to define what constitutes good performance can add rigor to an institution's decisions about what it should do. Among other results, the effort to define good performance can press an institution to improve its classification and tracking of programmatic inputs and outputs.

The apprehension about performance measurement goes beyond problems associated with choosing the correct yardstick. Even if supplied a perfect methodology, some institutions might shrink from applying it. On a good day, an evaluation can show that an institution's actions or processes had good consequences. On a bad day, the evaluation can reveal harmful effects, or no consequences at all. An institution might perceive that two of these three outcomes, no effects or bad effects, will be unflattering, and it may lack confidence that an ex post assessment will uncover good results. These anxieties may be understandable, but few objective observers would regard such fears as reasonable grounds to forgo performance measurement. If the institution's discomfort with conducting evaluations is based on a strong sense that rigorous measurement will reveal serious error, that, by itself, is reason for an institution to perform such assessments.

III. THE RATIONALE FOR EX POST EVALUATION IN MAKING COMPETITION POLICY

A number of rationales support the commitment of resources by competition agencies to develop and apply performance measures to evaluate substantive interventions and operational procedures. These rationales go beyond the general considerations of sound public administration described above in Part II. The need for performance measurement in competition policy stems from both the distinctive qualities of competition policy as a form of government regulation and from modern institutional developments within and across jurisdictions with competition laws.

A. Uncertainty: Competition Policy as Experimentation

The formulation of competition policy frequently requires antitrust authorities to make difficult judgments amid uncertainty about the competitive significance of business conduct.⁹ Will a merger retard or increase competition? Are the business justifications

9. Ken Heyer, *A World of Uncertainty: Economics and the Globalization of Antitrust*, 72 ANTITRUST L.J.

offered to support a refusal to deal or an exclusive contract genuine or contrived? These and other decisions can be difficult even in “routine” matters, and they can be especially challenging when rapid technological change, deregulation, or other dynamic forces complicate the analysis of competitive effects.¹⁰

Making policy amid uncertainty gives a substantial experimental element to government enforcement. Enforcement decisions can be viewed as experiments in which public authorities test the efficacy of different hypotheses about business behavior.¹¹ Over time, antitrust agencies reach a policy equilibrium by expanding and contracting the zone of enforcement. Testing the validity of different hypotheses involves making enforcement decisions that take chances with either intervening too aggressively or not intervening enough. Without experiments that sometimes intervene too much or sometimes intervene too little, it would be impossible for enforcement authorities to determine the correct mix of policies.¹²

The role of experimentation in competition policy is evident in the development in the United States since the 1930s of a policy against cartels. The story of modern U.S. anticartel policy can be told as a series of interrelated experiments with initiatives that primarily sought to increase the rate of detection for collusive schemes and to boost the punishments for violators.¹³ The enhancement by the Department of Justice (DOJ) of its leniency program in 1993 and 1994 is perhaps the best known and most widely emulated of these experiments.¹⁴ The successes in the past decade of high-powered leniency programs in eliciting the revelation of illegal cartels have obscured the skepticism within and without the DOJ that accompanied the reforms of the early 1990s. The only way to know whether a policy of stronger assurances about immunity from criminal prosecution would attract more disclosures was to try it out. It was not self-evident at the time that the program would make a genuine difference.

The importance of uncertainty and the experimental quality of competition policy is apparent in several dimensions of government decisionmaking. One analytical exercise that vividly underscores the uncertainty related difficulties of diagnosing past behavior or

375, 375 (2005).

10. See William E. Kovacic, *Transatlantic Turbulence: The Boeing-McDonnell Douglas Merger and International Competition Policy*, 68 ANTITRUST L.J. 805, 825-39 (2001) (reviewing how the Federal Trade Commission (FTC) and the competition directorate of the European Commission analyzed the likely competitive effects of Boeing’s acquisition of McDonnell Douglas amid conditions of technological dynamism and regulatory complexity).

11. See generally William E. Kovacic, *Evaluating Antitrust Experiments: Using Ex Post Assessments of Government Enforcement Decisions to Inform Competition Policy*, 9 GEO. MASON L. REV. 843 (2001) (discussing the experimental features of competition policy).

12. See William E. Kovacic, *The Modern Evolution of U.S. Competition Policy Enforcement Norms*, 71 ANTITRUST L.J. 377, 472-76 (2003) (emphasizing the cumulative nature of competition policymaking and the importance of experimentation as a source of learning for competition authorities).

13. See *id.* at 416-25; Jonathan B. Baker, *A Preface to Post-Chicago Antitrust*, in POST CHICAGO DEVELOPMENTS IN ANTITRUST LAW (Antonio Cucinotta et al. eds., 2002) (discussing how economic learning is absorbed into antitrust jurisprudence and enforcement policy).

14. The experimentation did not come to rest with the leniency reforms of the 1990s; it continues at a robust pace in many jurisdictions today. See, e.g., R. Hewitt Pate, *Current Issues in International Antitrust Enforcement*, in INTERNATIONAL ANTITRUST LAW & POLICY (Barry Hawk ed., 2005) (discussing the rationale for U.S. reforms in 2004 that increased maximum criminal fines for Sherman Act violations and provided partial dispensation from treble damages for certain leniency applicants).

predicting the impact of proposed transactions is the formulation of remedies.¹⁵ In negotiating settlements or preparing a request for relief in a case, competition agency officials often must make problematic judgments about which remedies will resolve competitive problems. The choice of cures in any specific case—for example, imposing limits on conduct to remedy an abuse of dominance—to some degree inevitably involves experimentation.¹⁶ As much as any element of competition policy, the selection of remedies calls for efforts to examine and learn from past experience as a guide to future decisions.

In the sciences, an indispensable element of the process of experimentation is the systematic evaluation of the experiment's results; so it should be with experimentation in competition policy. Without ex post testing, it will rarely be possible to determine whether the assumptions and hypotheses that motivated a competition agency's decision to prosecute were sound. Ex post assessments can also reduce the uncertainty associated with future decisions by illuminating how well various theories diagnose business conduct or predict competitive effects, and by informing judgments about the impact of various remedies.

B. Limited Transparency: The Impact of Policymaking by Settlement

A widely accepted principle of public administration is that government authorities should make the rationale for public policies and the processes for establishing such policies transparent. Transparency promotes clarity in forming public competition policy, increases the understanding of legal commands by affected parties, and disciplines the exercise of discretion by public officials by subjecting their actions to external review and criticism.¹⁷ Transparent policymaking methods that inform external observers (especially business operators) about the content of, and rationale for, specific decisions help ensure the regularity and honesty of public administration.¹⁸ Common transparency enhancing measures include publishing decisions in law enforcement matters, issuing guidelines, and using speeches to articulate the basis for specific initiatives.

The enforcement of competition laws relies significantly upon settlements.¹⁹ One

15. See William E. Kovacic, *Designing Antitrust Remedies for Dominant Firm Misconduct*, 31 CONN. L. REV. 1285, 1314 (2000) (describing how technological dynamism complicates the design of antitrust remedies).

16. In his discussion of the remedy accepted by the FTC in 1975 to resolve its monopolization case against Xerox, Willard Tom uses the language of experimentation. Willard K. Tom, *The 1975 Xerox Consent Decree: Ancient Artifacts and Current Tensions*, 68 ANTITRUST L.J. 967, 979 (2001). After recounting the origin of the remedy, Tom asks: "What are we to make of this naked, but apparently highly successful, experiment in social engineering?" *Id.*

17. See Mark A. Dutz & Shyam A. Khemani, *Instruments of Competition Policy and Their Relevance for Economic Development*, in REGULATORY POLICIES AND REFORM: A COMPARATIVE PERSPECTIVE 28 (Claudio R. Frischtak ed., 1995) (emphasizing the importance of transparent decisionmaking processes in formulating competition policy).

18. Weak transparency regimes can undermine the quality of public administration. See Joseph E. Stiglitz, *Knowledge for Development: Economic Science, Economic Policy, and Economic Advice*, in ANNUAL WORLD BANK CONFERENCE ON DEVELOPMENT ECONOMICS 9 (Boris Pleskovic & Joseph E. Stiglitz eds., 1999) ("Governments in many countries have a strong proclivity for secrecy. . . . Secrecy provides more scope for the work of special interest groups, greater cover for corruption, and greater opportunities for hiding mistakes.").

19. See Spencer Weber Waller, *Prosecution by Regulation: The Changing Nature of Antitrust Enforcement*, 77 OR. L. REV. 1383, 1408-17 (1998) (discussing how U.S. enforcement agencies often use

important source of this development is the establishment of mandatory notification mechanisms and waiting periods for merger control.²⁰ Premerger notification systems have expanded recourse to settlements involving divestitures or conduct-related undertakings to resolve competitive concerns with individual transactions.

In the typical settlement, it may be difficult for those other than the parties to the negotiations to accurately assess the basis for or significance of the settlement. In announcing settlements, the official statements of competition agencies tend to favorably portray the enforcement agency's decision to prosecute and its resolution of the matter. In their formal public pronouncements, competition agencies ordinarily do not express doubts about whether the solution contained in the settlement will cure apparent competition problems with the transaction. In the hospital of public competition law enforcement, surgeries usually are portrayed as successes.

One way to address transparency problems associated with using settlements to resolve government competition cases is to adopt policies that commit the competition agency to undertake periodic ex post assessments to evaluate the soundness of the decisionmaking process and to consider the effects of substantive interventions such as settlements. The performance of ex post evaluations, either by the competition agency itself or by outsiders with the competition agency's cooperation, and the publication of results would help determine the value of specific remedies.

C. The Evolutionary Quality of Antitrust Jurisprudence and Enforcement Policy

The development of competition policy is inherently evolutionary.²¹ In some jurisdictions, the evolutionary character stems partly from the institutional design of the competition policy system. In adopting the principal U.S. antitrust laws, for example, Congress created general statutory commands and gave the federal courts responsibility for interpreting their operative terms and adjusting the content of doctrine over time.²² The U.S. courts have recognized that fulfilling their assigned role in this consciously evolutionary scheme requires an awareness of how past interpretations of the antitrust laws have affected commerce.²³

A more universal source of competition policy's evolutionary nature is its interdisciplinary foundation. In antitrust practice, economic analysis plays a central role in resolving such key antitrust issues as delineating the relevant market and assessing the efficiency consequences of various forms of business behavior.²⁴ Economics is a

settlements to resolve antitrust concerns).

20. See William Blumenthal, *Overenforcement in the Hart-Scott-Rodino Second Request Process*, in *THE ECONOMICS OF THE ANTITRUST PROCESS* 15 (Malcolm B. Coate & Andrew N. Kleit eds., 1996) (describing the U.S. system for notification and review of proposed mergers).

21. See Melanie Johnson, *Keynote Speech*, in 2 *CURRENT COMPETITION LAW* 3, 7 (Mads Andenas et al. eds., 2004); see generally Kovacic, *supra* note 12 (describing how history-based narratives have informed federal antitrust policy).

22. See William E. Kovacic, *The Influence of Economics on Antitrust Law*, 30 *ECON. INQUIRY* 294 (1992) (describing congressional delegation of interpretational authority to the courts in the U.S. antitrust laws).

23. See *State Oil Co. v. Khan*, 522 U.S. 3, 20 (1997) (describing the Supreme Court's distinctive role under the antitrust laws "in recognizing and adapting to changed circumstances and the lessons of accumulated experience").

24. See Baker, *supra* note 13, at 60-76 (discussing how economic learning is absorbed into antitrust jurisprudence and enforcement policy); see also William E. Kovacic & Carl Shapiro, *Antitrust Policy: A*

dynamic discipline, and the history of industrial organization economics has featured considerable change and refinement in the understanding of commercial phenomena.²⁵ Changes in economic learning can alter antitrust legal doctrines that are informed by economic insights. Empirical research, including the analysis of past antitrust cases, has supplied a major impetus for the development of new industrial organization ideas and an important stimulus for alterations in antitrust doctrine and enforcement policy.²⁶

D. Institutional Multiplicity: Concurrent Oversight

The modern competition policy environment features numerous public and private actors with authority to enforce antitrust commands. Multiplicity has a significant international dimension and, in some jurisdictions, an important domestic element. Globally, the development of new competition policy systems and the enhancement of older regimes have meant that individual mergers or other forms of commercial activity are likely to attract attention from a variety of national or regional competition bodies.²⁷ In some jurisdictions, such as Brazil, France, and the United States, responsibility for formulating and executing competition policy is shared by two or more public competition authorities.²⁸ In many jurisdictions, the public competition authorities also share authority with sectoral regulators to scrutinize mergers or other forms of conduct.²⁹ The slow but continuing development of private rights of action in jurisdictions that previously had relied solely upon public enforcement provides for still further

Century of Economic and Legal Thinking, 14 J. ECON. PERSP. 43 (2000) (describing the influence of economic analysis on the development of legal antitrust doctrines in the United States since 1890).

25. See Dennis W. Carlton, *A General Analysis of Exclusionary Conduct and Refusal to Deal—Why Aspen and Kodak are Misguided*, 68 ANTITRUST L.J. 659, 680 (2001) (“[A]s the literature in economics shows, economists often take decades to understand certain business practices.”).

26. See *State Oil*, 522 U.S. at 15-18 (reviewing how “a considerable body of scholarship discussing the effects of vertical restraints” influenced judicial reassessment of the per se ban on maximum resale price maintenance); see also Timothy J. Muris, *GTE Sylvania and the Empirical Foundations of Antitrust*, 68 ANTITRUST L.J. 899, 903-07 (2001) (describing the impact of developments in industrial organization economics on merger enforcement policy).

27. See Symposium, *International Merger Control*, ANTITRUST MAG., Spring 2001 (reviewing developments in merger control in various national jurisdictions); William E. Kovacic, *Merger Enforcement in Transition: Antitrust Controls on Acquisitions in Emerging Economies*, 66 U. CIN. L. REV. 1075 (1998) (examining opportunities for multiple review of mergers resulting from the growth in the number of competition policy systems in transition economies).

28. The decentralization of prosecutorial authority across public and private entities in the United States is unmatched. See William E. Kovacic, *Toward a Domestic Competition Network*, in COMPETITION LAWS IN CONFLICT 316 (Richard A. Epstein & Michael S. Greve eds., 2004).

29. See William J. Kolasky, *The FCC’s Review of the Bell Atlantic/NYNEX and SBC/Ameritech Mergers: Regulation Overreach in the Nature of Promoting Competition*, 68 ANTITRUST L.J. 771 (2001) (describing parallel review by the DOJ and Federal Communications Commission of mergers in the telecommunications sector). In a number of jurisdictions, government regulatory bodies not ordinarily considered to have a competition policy portfolio make decisions that significantly influence the competitive process and affect, at least indirectly, the programs of public competition authorities. See FED. TRADE COMM’N, TO PROMOTE INNOVATION: THE PROPER BALANCE OF COMPETITION AND PATENT LAW AND POLICY (2003), available at <http://www.ftc.gov/os/2003/10/innovationrpt.pdf> (discussing how the U.S. process for granting patent rights can affect competition); William E. Kovacic & Andreas Reindl, *An Interdisciplinary Approach to Improving Competition Policy and Intellectual Property Policy*, 28 FORDHAM INT’L L.J. 1062 (2005) (discussing how decisions of institutions that grant intellectual property rights can have a major impact on competition).

decentralization of prosecutorial authority.

The fragmentation of policymaking authority requires business managers to account for a wide array of substantive competition policy commands and varied interpretations of the same commands by different public enforcement bodies. In recent years a number of observers have drawn attention to complications that arise when individual episodes of business behavior are subject to numerous parallel enforcement efforts within and across jurisdictions. Most attention has focused on the possibility that multiple reviews of mergers by national competition authorities can unduly raise the cost of executing such transactions.³⁰ An emerging area of concern is the extent to which diverse remedial schemes interact in the prosecution of specific instances of misconduct. Ex post studies of specific enforcement episodes would assist in clarifying the impact of having numerous public antitrust authorities examine the same business conduct and informing debate about possible adjustments.

E. Institutional Multiplicity: Benchmarking Operational Processes

For all of the complexity it introduces into business planning and interagency coordination, the development of new competition systems and the reform of existing systems have created numerous opportunities for competition authorities in one jurisdiction to benchmark their operational procedures with their counterparts. There is considerable variation in the manner in which individual authorities organize their institutions into operational units, devise strategies, and allocate resources. The diversification in approaches provides numerous comparative yardsticks by which an agency can evaluate the soundness of its own organizational choices and procedures.

Among other possibilities, comparative study supplies a basis for a competition agency to decide how to establish internal quality control—for example, by establishing an independent unit of economists with authority to report directly to the agency leadership or by creating “devil’s advocate” panels to test the assumptions and evidence of agency case handlers. As suggested below, diversification also has supplied models for agencies to consider in building their own measures of performance and conducting ex post performance reviews.

F. The Value of Competition Agency Participation in Performance Management

Competition agencies are not the only possible source of performance assessments of competition policy. Researchers have created a substantial body of literature that analyzes specific cases, enforcement programs, or procedures without participation by or cooperation from the government agencies responsible for the matters in question.³¹ In a

30. See INT’L COMPETITION POLICY ADVISORY COMM., ANTITRUST DIV., DEP’T OF JUSTICE, FINAL REPORT 3-18 (2000), available at <http://www.usdoj.gov/atr/icpac/finalreport.htm> (summarizing possible costs associated with multiple reviews of individual transactions conducted by national competition authorities and by different competition bodies within a single country).

31. For a representative collection of papers examining the outcomes of specific cases, see Symposium, *Antitrust at the Millennium (Part I)*, 68 ANTITRUST L.J. 1 (2000); and Symposium, *Antitrust at the Millennium (Part II)*, 68 ANTITRUST L.J. 621 (2001). See also JOHN FINGLETON ET AL., *COMPETITION POLICY AND THE TRANSFORMATION OF CENTRAL EUROPE* (1995) (written by researchers who enjoyed extensive access to the competition agencies and who examined in detail the enforcement programs in several nations in Central

number of instances, such assessments have influenced antitrust policy by changing the views of enforcement officials and courts about the validity of certain doctrines or enforcement practices.³² Government bodies outside the competition agency, such as entities entrusted with audit responsibilities, also have conducted performance assessments.³³ In recent years, government bodies such as Canada's International Development Research Centre (IDRC) have funded projects to encourage local researchers in transition economies to evaluate the actual effects of competition law enforcement and to identify barriers to competition.³⁴

As these examples suggest, a competition policy system might rely solely on the initiative of researchers or institutions other than the competition authority to obtain ex post assessments of individual cases or enforcement programs without the participation of the competition agency. Several considerations indicate the importance of formal participation by the competition agencies. The competition agency is likely to be the main repository of information about the decision to prosecute and about internal deliberations concerning the management of individual cases. The competition authority also may have a unique capability to collect information relating to the effects of enforcement programs.³⁵

The most important factor favoring performance measurement within the competition agency is the likelihood that a routine of evaluation will have a greater effect in promoting regular and timely adjustments in the agency's practice. One ideally would prefer to have the competition agency actively seek improvements in its operations and to immediately and directly incorporate knowledge from past experience into future operational decisions. George Stigler said as much 35 years ago in discussing whether it was wise for an institution to rely entirely on periodic reviews by outsiders to evaluate programs and identify areas for adjustment:

No organization can achieve or maintain efficiency in structure or operation by having a critical review made by expert outsiders once each five or ten years—even if . . . the recommendations of the review panel are unfailingly adopted. A good organization must have built into its very structure the incentives to its personnel to do the right things.³⁶

Europe); Joe Sims & Deborah P. Herman, *The Effect of Twenty Years of Hart-Scott-Rodino on Merger Practice: A Case Study in the Law of Unintended Consequences Applied to Antitrust Legislation*, 65 ANTITRUST L.J. 865 (1997) (examining in detail the merger review process in one jurisdiction).

32. For example, efforts in the United States to establish mandatory notification requirements and waiting periods for proposed mergers drew heavily upon the research of Kenneth G. Elzinga. See Kenneth G. Elzinga, *The Antimerger Law: Pyrrhic Victories?*, 12 J.L. & ECON. 43 (1969).

33. For example, the United Kingdom's National Audit Office recently completed an evaluation of the operations of the Office of Fair Trading. See NAT'L AUDIT OFFICE, ENFORCING COMPETITION IN MARKETS (2005), available at http://www.nao.org.uk/publications/nao_reports/05-06/0506593.pdf (analyzing the effectiveness of the United Kingdom's Office of Fair Trading).

34. See Simon J. Evenett, Promoting Competitive Markets in Developing Countries (Mar. 2004), http://web.idrc.ca/en/ev-32149-201-1-DO_TOPIC.html (describing the IRDC research project).

35. See 15 U.S.C. § 46 (2000) (granting the FTC power to gather company data to perform economic studies, such as ones examining the effects of specific enforcement measures).

36. BLUE RIBBON DEFENSE PANEL, REPORT TO THE PRESIDENT AND THE SECRETARY OF DEFENSE ON THE DEPARTMENT OF DEFENSE 198 (1970).

IV. EXPERIENCE WITH PERFORMANCE EVALUATION OF COMPETITION AGENCIES

Beginning in the late 1970s and continuing to the present, three strands of commentary have focused attention on the adequacy of competition agency efforts to evaluate their performance. The first strand has taken the form of recommendations by government bodies and individual commentators that competition agencies expand the amount of resources allocated to evaluating the effects of past enforcement decisions.³⁷ The FTC hearings in 1995 on innovation and globalization were a formative event that identified an emerging consensus on this point.³⁸ The desirability of devoting greater resources to ex post evaluation of completed matters was a major theme of competition policy experts who testified at these and subsequent FTC proceedings.³⁹

A second and distinct strand of noteworthy commentary has come from observers who question the value of past competition policy initiatives and argue that many competition policy programs failed to achieve their aim of improving economic performance.⁴⁰ Some analysts focus on the lack of evidence suggesting that antitrust intervention had yielded measurable benefits. Others argued that some or many interventions had diminished consumer welfare. The mildest policy implication of these

37. See FINGLETON ET AL., *supra* note 31, at 171 (discussing competition policy enforcement in Visegrad countries and urging “greater evaluation and refinement of the criteria used in making decisions”); ROBERT A. KATZMANN, *REGULATORY BUREAUCRACY: THE FEDERAL TRADE COMMISSION AND ANTITRUST POLICY* 204 (1980) (commenting that studies by the FTC “could help resolve current public-policy debates about the role of antitrust in economic policy” and that “[m]ore work needs to be done about the effects of market concentration; there is a paucity of information about the impact of antitrust action on various economic indices”); U.S. GEN. ACCOUNTING OFFICE, *CLOSER CONTROLS AND BETTER DATA COULD IMPROVE ANTITRUST ENFORCEMENT* 15 (1980) (recommending that the U.S. Attorney General ensure that the Antitrust Division of the DOJ “provide[] for a continuing assessment and evaluation of the effectiveness of enforcement efforts in promoting and restoring competition”); William E. Kovacic, *Failed Expectations: The Troubled Past and Uncertain Future of the Sherman Act as a Tool for Deconcentration*, 74 IOWA L. REV. 1105, 1147 (1989) (stating that ex post evaluations could improve the design of government monopolization cases); William E. Kovacic, *Federal Antitrust Enforcement in the Reagan Administration: Two Cheers for the Disappearance of the Large Firm Defendant in Nonmerger Cases*, 12 RES. L. & ECON. 173, 187 (1989) (advocating ex post evaluations to inform the selection of cases and litigation techniques); Armando E. Rodriguez & Mark D. Williams, *Recent Decisions by the Venezuelan and Peruvian Agencies: Lessons for the Export of Antitrust*, 43 ANTITRUST BULL. 147, 178 (1998) (suggesting that multinational bodies and individual donors should increase monitoring and analysis of enforcement decisions of transition economy competition agencies); J. Luis Guasch, *Competition Policy Advocacy and Regulatory Reform* (Dec. 13, 1998) (paper presented at the World Bank’s First International Training Program on Competition Policy) (“It is imperative that competition agencies . . . periodically evaluate the impact (efficiency and distribution) of their decisions and widely disseminate their findings.”).

38. The results of these proceedings are reported in FED. TRADE COMM’N, *ANTICIPATING THE 21ST CENTURY: COMPETITION POLICY IN THE NEW HIGH-TECH, GLOBAL MARKETPLACE, VOLUME I* (1996), available at http://www.ftc.gov/opp/global/report/gc_v1.pdf.

39. See Kovacic, *supra* note 11, at 855-56 (collecting authorities). At an FTC roundtable of industrial organization economists in September 2001, a central theme of discussion was the desirability of additional empirical research to analyze the effects of past antitrust enforcement decisions. See Fed. Trade Comm’n, *Empirical Industrial Organization Roundtable* 115-41 (Sept. 11, 2001), available at <http://www.ftc.gov/be/empiricalioroundtabletranscript.pdf>.

40. See, e.g., Robert W. Crandall, *The Failure of Structural Remedies in Sherman Act Monopolization Cases*, 80 OR. L. REV. 109 (2001); Robert W. Crandall & Kenneth G. Elzinga, *Injunctive Relief in Sherman Act Monopolization Cases*, 21 RES. L. & ECON. 277 (2004); Robert W. Crandall & Clifford Winston, *The Breakdown of ‘Breakup,’* WALL ST. J., Mar. 8, 2006, at A14.

critiques is that selected forms of antitrust enforcement, such as efforts to control abusive behavior by dominant firms, had generally been counterproductive and warranted dramatic retrenchment or abandonment. The bolder proposition in this line of commentary is that because antitrust laws more often than not retard economic progress, nations are better off without them.

The third line of commentary has developed in discussions about the appropriate foundations for economic growth in countries undertaking the transition from central planning toward greater reliance on market processes. In discussions about economic law reform, commentators sometimes have questioned the value of placing competition policy on the reform agenda or have argued that transition economies would do best by forgoing the enactment of antitrust laws.⁴¹ Both variants of this perspective, those who doubt the benefits of competition policy for transition economies and those who emphatically oppose the adoption of antitrust laws in transition environments, suggest that the proponents of competition policy bear the burden, as yet unmet, of showing that the enactment and implementation of competition laws yield positive results. In effect, the skeptics and the outright opponents have underscored the value of assembling evidence that shows the value of competition policy in emerging markets.

The discussion below addresses three types of evaluation related activities concerning competition agencies: general assessments of competition agency performance and evaluation efforts, evaluations of the effects of individual interventions such as cases or advocacy initiatives, and evaluations of agency processes or programs. This Article describes each type of evaluation activity and provides detailed treatments of more notable initiatives that illustrate important methodological strengths and weaknesses of specific approaches.

A. General External Studies of Competition Agency Performance and Evaluation Efforts

On many occasions, competition agencies have been the subject of studies by external observers who seek to appraise the quality of competition agency performance. In a number of instances, the authors of these studies have assessed or commented upon the effectiveness of competition agency efforts to make performance evaluations a component of routine operations. The general and wide-ranging studies of agency performance and performance measurement initiatives fall into the four categories described below.

1. Studies by Individual Researchers or Small Teams of Researchers

One category of inquiry consists of studies that examine the full scope of a competition agency's operations or major elements of an agency's work. Some of these studies have been performed by individual scholars,⁴² while others have been assembled

41. See William E. Kovacic, *Institutional Foundations for Economic Legal Reform in Transition Economies: The Case of Competition-Policy and Antitrust Enforcement*, 77 CHI.-KENT L. REV. 265, 286-90 (2001) (surveying and discussing this literature).

42. Notable examples of this type include KATZMANN, *supra* note 37; and SUZANNE WEAVER, *THE DECISION TO PROSECUTE: ORGANIZATION AND PUBLIC POLICY IN THE ANTITRUST DIVISION* (1977) (studying the DOJ's Antitrust Division).

by teams of researchers.⁴³ The authors of this type of study ordinarily, but not invariably, enjoy some degree of formal or informal cooperation from insiders at the competition agency in performing research on the institution.⁴⁴ The studies tend to focus most intensely on the quality of the agency's internal decisionmaking processes, but some authors have presented detailed evaluations of specific interventions.⁴⁵

2. *Studies by Other Government Institutions*

A second category of broad-ranging study consists of works prepared by government agencies outside the competition authority. The most noteworthy of these works have been prepared by national authorities with responsibility for auditing government agencies. Due to their status as government institutions, and by reason of their information-gathering powers, these auditing authorities typically have extensive access to the internal records and personnel of the competition agency in performing their research.

In the United States, Congress occasionally instructs the Government Accountability Office (GAO) (formerly known as the Government Accounting Office) to study various features of the operations of the U.S. antitrust agencies. One noteworthy example of this type of inquiry took place in 1980, when the GAO conducted a detailed assessment of the management systems of the FTC and the Antitrust Division of the DOJ. Among other findings, the GAO study urged the federal antitrust agencies to devote more resources to conducting ex post assessments of individual competition matters and to improve internal management systems for monitoring the progress of investigations and other case-related activities within the agencies.⁴⁶ In 2005 the United Kingdom's National Audit Office published an assessment of the Office of Fair Trading that involved, among other focal points, the discussion of performance measures for evaluating the Office's competition policy activities.⁴⁷

3. *Studies by Blue Ribbon Panels*

Competition agencies occasionally have been the subject of assessments prepared by special committees or blue ribbon panels. In many instances, blue ribbon panels have been assembled under the auspices of nongovernmental organizations (NGOs), either at the request of government officials or on the independent initiative of the NGO.⁴⁸ Other

43. Important contributions of this type include FINGLETON ET AL., *supra* note 31 (assessing the implementation of competition policy in the Visegrad countries); THE FEDERAL TRADE COMMISSION SINCE 1970: ECONOMIC REGULATION AND BUREAUCRATIC BEHAVIOR (Kenneth W. Clarkson & Timothy J. Muris eds., 1981) (studying the FTC); and EDWARD COX ET AL., THE NADER REPORT (1969) (evaluating the consumer protection activities of the FTC).

44. In some jurisdictions, researchers can take advantage of laws that compel public agencies to disclose certain types of records.

45. See THE FEDERAL TRADE COMMISSION SINCE 1970: ECONOMIC REGULATION AND BUREAUCRATIC BEHAVIOR, *supra* note 43 (providing detailed assessments of selected competition cases).

46. See U.S. GEN. ACCOUNTING OFFICE, *supra* note 37.

47. See *supra* note 33 and accompanying text.

48. In 1969, at the request of President Richard Nixon, the American Bar Association (ABA) assembled a blue ribbon panel to evaluate the FTC. In the late 1980s, at its own initiative, the ABA convened blue ribbon panels to examine the DOJ Antitrust Division and the FTC.

blue ribbon panels have been established on an ad hoc basis without any attachment to an existing government or NGO.⁴⁹

For the most part, blue ribbon panels have devoted most of their attention to the management of competition agencies, and have focused less upon the wisdom of specific interventions such as cases or advocacy measures. While performing these process-oriented assessments, the blue ribbon inquiries have generally enjoyed substantial cooperation from the competition agencies, particularly if the blue ribbon panel originated with a request from the chief of state or the national legislature. The impact of these inquiries has varied considerably, but one can identify a number of instances in which the recommendations of the blue ribbon panel have spurred significant reforms in the management of the competition agency.⁵⁰

4. Agencywide or Systemwide Evaluations by Multinational Organizations

A fourth category of general assessment has taken the form of studies of national competition systems or individual competition agencies performed by multinational organizations. The Organization for Economic Co-Operation and Development (OECD) in particular has engaged in two types of broad-based inquiries. In the first type of inquiry, consultants to the OECD or members of the OECD Competition Committee Secretariat sometimes conducted studies of the enforcement activities of individual jurisdictions and published reports of their findings. Although such reports have not attempted to provide a comprehensive profile of activity, they have provided useful perspectives on the enforcement programs and operations of the competition authorities in question.⁵¹

The second form of agencywide or systemwide assessment has consisted of “peer reviews,” a format pioneered by the OECD’s competition committee. In order to build a foundation for the peer review exercises, OECD consultants or professional staff prepare reports based on a detailed study of the competition agency’s enforcement patterns, interviews, and written questionnaires, and conversations with competition policy experts and business officials inside the country.⁵² In conducting the investigation, the OECD researchers typically have enjoyed substantial cooperation from the competition agency under examination. The reports supply the chief basis for peer review sessions at the

49. A recent example is the Antitrust Modernization Commission (AMC), which Congress established to evaluate various aspects of U.S. competition policy. The AMC received an appropriation of funds from Congress to hire a professional staff and conduct its operations. For a discussion of the AMC and earlier antitrust blue ribbon commissions, see Albert A. Foer, *Putting the Antitrust Modernization Commission into Perspective*, 51 BUFF. L. REV. 1029 (2003).

50. See William E. Kovacic, *The Federal Trade Commission and Congressional Oversight of Antitrust Enforcement*, 17 TULSA L.J. 587, 630-71 (1982) (describing how recommendations of a 1969 ABA blue ribbon panel elicited major reforms of the FTC).

51. See John Clark, *Competition Law and Policy in the Baltics*, OECD J. COMPETITION L. & POL’Y, Sept. 1999, at 144 (providing a representative example of OECD Competition Committee Secretariat reports).

52. See Michael Wise, *Review of Competition Law and Policy in Hungary*, OECD J. COMPETITION L. & POL’Y, June 2000, at 57 (providing a small representative sample of OECD peer review studies); Michael Wise, *Review of Competition Law and Policy in Italy*, OECD J. COMPETITION L. & POL’Y, Sept. 2001, at 59 (same); Michael Wise, *Review of Competition Law and Policy in Korea*, OECD J. COMPETITION L. & POL’Y, Oct. 2000, at 121 (same); Michael Wise, *Review of United States Competition Law and Policy*, OECD J. COMPETITION L. & POL’Y, Feb. 1999, at 3 (same).

regularly scheduled meetings of the OECD competition committee and the OECD's Global Competition Forum. The results of these audits often are made available to the public. OECD's peer review efforts have been supplemented by contributions from other multinational bodies such as the World Trade Organization (WTO). The WTO has used conferences and seminars to promote the benefits of peer review as a means for newer competition authorities to improve operations and benefit from the transfer of know-how from more experienced institutions.

5. General Assessments of Competition Agencies by Outsiders: Strengths and Weaknesses

External observers have made important contributions to our understanding of the operation of competition agencies and, in important cases, have served to improve agency management and operations. The agency's awareness that it will periodically undergo review by outside researchers and institutions provides an incentive to engage in internal initiatives to improve operations and to measure performance. Where the external examiner or examining body is an expert in the operations of the competition agency and enjoys extensive access to the agency's internal records and personnel, it is possible to offer an informative diagnosis of the agency's work and to suggest useful avenues for improvement.

However, reliance on external observers to perform general assessments of competition agency performance can suffer from four basic problems. One weakness in the general assessments generated by outsiders, particularly evident in the work of blue ribbon panels, is the lack of consistent and commonly accepted criteria for measuring agency performance. The external researcher or research body may fail to carefully specify its evaluative standards and, in reviewing the findings of previous evaluative research, may fail to identify differences in the standards applied by earlier studies.⁵³ The rush to prepare the instant report often denies the blue ribbon panel time to do the more discriminating and historically accurate assessment that is vital to making sound assessments about the quality of past performance and to make sensible recommendations about future improvements.

A second problem concerns the knowledge base of the external analysts. In some cases, the external observers lack adequate access to agency records and personnel to develop a confident understanding of the agency's current operations. This is less problematic for researchers, such as national audit bodies or panels operating with the endorsement of the chief executive or national legislature, with strong tools to compel cooperation than it is for individual researchers seeking, without official government sponsorship, to conduct broad-ranging studies of agency operations. Another form of knowledge deficiency is the researcher's limited background in the agency or in the competition policy field generally. Without a well-informed lens for studying the agency, it is unlikely that the researcher will make accurate or wise interpretations of what he observes.

A third problem concerns the incentives of the external observer to provide a fully

53. Blue ribbon panels have a habit of reciting the conclusions of earlier and related blue ribbon studies without accounting for differences in the evaluative standards applied by previous panels. See Kovacic, *supra* note 50, at 599-602 (discussing the work of blue ribbon panels examining FTC performance).

candid assessment of the agency's operations. A researcher, or member of a research team, who wishes to have a future relationship or to work in cooperation with the competition agency may be reluctant to give a truly unvarnished assessment of the agency's flaws. When an agency cooperates with an external reviewer by choice and not by compulsion, the agency may decline to participate in exercises in which the reviewer refuses, implicitly or explicitly, to moderate harsh conclusions, at least for public consumption. This is an inherent limitation of the voluntary peer review exercises conducted by OECD and other multinational bodies. If the peer review report is absolutely unflinching in its criticism of a competition agency, and if discussants at OECD meetings pull no punches in their interrogatories to agency officials, agencies may be unwilling to subject themselves to such assessments.⁵⁴

A fourth problem concerns the passivity that may permeate an agency that relies chiefly on outsiders to conduct occasional reviews of the agency's performance. Reviews by outsiders can complement, but cannot substitute completely for, the agency's own efforts to engage in a continuous and routine assessment of its interventions and operations. As noted in Part III.F above, the aim should be to encourage each competition agency to devise its own mechanisms to evaluate and improve the quality of its work.

B. Case-Specific Evaluations of Competition Agency Intervention

The antonym of the macroscopic and highly generalized assessment of agency performance described above is the narrowly focused and microscopic evaluation of the effects of specific interventions. This subsection discusses the evaluation of individual initiatives. This form of inquiry has been the province of individual researchers, competition agencies, and multinational bodies, as blue ribbon panels generally have not made detailed assessments of individual matters. The discussion below reviews experience with case studies and examines the most ambitious effort to date by a competition agency to evaluate case outcomes.

1. Studies Performed by Outside Researchers Without Competition Agency Involvement

Most of the case-specific evaluations of the substantive outcomes of competition agency interventions have been performed by academics working alone or in small teams. With some notable exceptions, discussed below, academics have performed their evaluations of individual cases solely on the basis of publicly available data and without access to the competition agency's internal records. These studies are "independent" of the competition agency in the sense that they neither depend on the competition agency as a source of information nor do they require the researchers to make explicit or implicit concessions to the agency as a condition for performing the research. The main limitation of such studies is the lack of access researchers ordinarily have to internal agency records that reveal the thought process that led to the decision to prosecute.

54. A compromise of sorts that may address this problem is for the peer review researcher to provide highly sensitive criticisms privately to the agency under review and to moderate or omit such criticisms in reports or presentations made to a larger audience.

2. *Studies by Other Government Bodies: GAO Study of U.S. Petroleum Mergers*

Government agencies with audit responsibilities have occasionally conducted detailed studies of individual competition agency enforcement decisions. In 2004, the GAO published the most ambitious effort to date by a government audit authority to evaluate the effects of a competition agency's interventions.⁵⁵ The GAO sought to measure the competitive effects of eight petroleum industry mergers completed between 1997 and 2000. The FTC had reviewed and approved all eight mergers. However, it had required the merging parties to make significant divestitures in some of the transactions. The mergers the GAO analyzed included Exxon's acquisition of Mobil, which the FTC approved following the largest divestiture of retail assets in the history of U.S. merger control.

To perform its study, the GAO developed econometric models to identify how the mergers and supplier concentration affected prices of gasoline. In formulating its model and research plan, the GAO relied mainly on its own staff of economists to develop the models, but also consulted external advisors, including several expert economists in petroleum industry competition issues. The GAO acquired data purchasing information on wholesale prices from a private firm, the Oil Price Information Service, which monitors petroleum industry activities. The GAO shared an initial draft of its report with the FTC, which offered extensive critical comments on the study's methodology. The GAO largely declined to accept the FTC's methodological suggestions and published the report in May 2004. The GAO study concluded that six of the eight mergers it reviewed resulted in price increases, while two caused gasoline prices to fall.

Publication of the petroleum merger study triggered an extended and intense public debate between the GAO and the FTC about the soundness of the GAO's work. The FTC contended that the GAO report contained serious methodological errors that negated the reliability of its results.⁵⁶ The FTC argued that three fundamental flaws undermined the GAO study. First, in performing econometric analyses of specific mergers, the GAO allegedly failed to account properly for many factors that affect gasoline prices. Second, the FTC contended that the GAO's assessment of how concentration affects prices did not use the properly defined relevant markets that sound industrial organization analysis requires. Third, the FTC said that the GAO failed to consider critical facts about individual transactions, such as Exxon's merger with Mobil, that were vital to assessing their impact on prices.

In January 2005, the FTC and GAO jointly convened a conference to discuss the work that the two agencies had done to evaluate the effects of petroleum industry mergers.⁵⁷ Academic economists, including several experts in econometrics, critiqued the

55. U.S. GOV'T ACCOUNTABILITY OFFICE, *ENERGY MARKETS: EFFECT OF MERGERS AND MARKET CONCENTRATION IN THE U.S. PETROLEUM INDUSTRY* (2004).

56. *Market Forces, Anticompetitive Activity, and Gasoline Prices: FTC Initiatives to Protect Competitive Markets: Hearing on the Status of the U.S. Refining Industry Before the H. Comm. on Energy and Air Quality* (July 15, 2004) (prepared statement of the FTC, presented by William E. Kovacic, FTC General Counsel), available at <http://www.ftc.gov/os/2004/07/040715gaspricetestimony.pdf>.

57. Deborah Platt Majoras, Chairman, Fed. Trade Comm'n, Remarks at the FTC and GAO Joint Conference on Petroleum Industry Mergers: Estimating the Price Effects of Mergers and Concentration in the Petroleum Industry: An Evaluation of Recent Learning (Jan. 14, 2005), available at <http://www.ftc.gov/speeches/majoras/050114oilmergerconferenceremarks.pdf> (describing the conference to

work of the two agencies. Two major limitations of this exercise were the GAO's refusal to fully specify its econometric methodology and its refusal to make available all the data used to run its models. A number of commentators pointed out that without the full revelation of the model and all the data used to run the model, it would be impossible to conduct a meaningful test of the GAO results.

3. Case-Specific Reviews Performed by Multinational Bodies

A number of multinational organizations have conducted seminars or workshops whose focus was on the analysis of specific enforcement measures undertaken by new competition authorities. OECD has carried out several forms of evaluation exercises since the early 1990s.⁵⁸ OECD has conducted regular case seminars that have provided detailed examination of selected enforcement initiatives of transition economy competition agencies.⁵⁹ In the case seminar format, small teams of foreign experts meet for several days with case handlers and other officials from the competition authorities to review specific cases. The outside advisors examine the competition agency's prosecution files and hear presentations by the competition agency officials, who recount their decision to prosecute and discuss the responses of the defendant firms. The foreign advisors then critique the agency's work. Variants of the case seminars have taken place in regional conferences organized by the United Nations Conference on Trade and Development (UNCTAD) individually or in cooperation with other donor bodies.

4. Case-Specific Reviews Undertaken or Sponsored by Competition Agencies

Over the past 25 years competition agencies on some occasions have conducted studies of individual cases. As described below in Part IV.B.5, the FTC in the late 1970s and early 1980s sponsored evaluations by outside academics of FTC decisions involving vertical restraints⁶⁰ and dominant firm behavior.⁶¹ Among other assessments completed in the past decade, the FTC and its professional staff published studies of the agency's interventions to control mergers in the soft drink industry,⁶² the implementation of divestiture decrees in merger cases,⁶³ the effect of mergers in the hospital sector,⁶⁴ and

examine methodologies to evaluate the effects of petroleum industry mergers).

58. See William E. Kovacic, *Lessons of Competitive Policy Reform in Transition Economies for U.S. Antitrust Policy*, 74 ST. JOHN'S L. REV. 361, 394-95 (2000) (discussing OECD programs to evaluate antitrust enforcement).

59. The author thanks Sally Van Siclen for many useful conversations about the OECD case seminars.

60. See U.S. FED. TRADE COMM'N, *IMPACT EVALUATIONS OF FEDERAL TRADE COMMISSION VERTICAL RESTRAINTS CASES* (Ronald N. Lafferty et al. eds., 1984) (presenting the results of assessments of FTC vertical restraints cases).

61. See Timothy Bresnahan, *Post-Entry Competition in the Plain Paper Copier Market*, 75 AM. ECON. REV. 15 (1985) (presenting an evaluation of the relief obtained in 1975 by the FTC through settlement of its abuse of dominance complaint against Xerox).

62. U.S. FED. TRADE COMM'N, *TRANSFORMATION AND CONTINUITY: THE U.S. CARBONATED SOFT DRINK BOTTLING INDUSTRY AND ANTITRUST POLICY SINCE 1980* (1999).

63. BUREAU OF COMPETITION, FED. TRADE COMM'N, *A STUDY OF THE COMMISSION'S DIVESTITURE PROCESS* (1999), available at <http://www.ftc.gov/os/1999/08/divestiture.pdf>.

64. Michael G. Vita & Seth Sacher, *The Competitive Effects of Not-For-Profit Hospital Mergers: A Case Study*, 49 J. INDUS. ECON. 63 (2001).

the effect of mergers in the petroleum industry.⁶⁵ The FTC also has sought to improve the aggregate measures of performance mandated by Congress under the Government Performance and Results Act.⁶⁶

In the late 1980s, the DOJ, acting pursuant to the terms of a consent decree, funded research concerning the effects of the restructuring of AT&T.⁶⁷ DOJ researchers also reviewed the consequences of selected mergers of airlines in the 1980s⁶⁸ and examined the effects of mergers of accounting firms in the 1990s.⁶⁹ DOJ economists currently are performing research on the consequences of mergers in other industries.

Of agency sponsored evaluations that have been completed, one of the most interesting was conducted in the 1990s by the Peruvian competition agency, the National Institute for Defense of Competition and Protection of Intellectual Property (Indecopi).⁷⁰ During her tenure as Indecopi's chair, Beatriz Boza invited academics and practitioners to serve as researchers in residence in Indecopi and to study various aspects of Indecopi's substantive programs and administrative procedures. As such, the researchers had extensive access to internal Indecopi records and to the agency's officials.

The results of the researchers' work, which Boza called an "academic audit," were published following internal discussion and comment within Indecopi.⁷¹ The publication of the studies fulfilled Indecopi's commitment to the researchers at the outset of the project to publish the completed papers (redacted to omit nonpublic information), without regard to whether the researchers praised or criticized the agency. Boza and her colleagues regarded the preparation and publication of the studies as important vehicles for improving the quality of Indecopi's decisionmaking and increasing public awareness of the institution.

In addition to these completed evaluation projects, a number of competition agencies in recent years have initiated projects to evaluate the impact of individual cases. Among other developments, various bodies with competition policy duties in the UK have been conducting research to document the consequences of individual cases.⁷²

65. U.S. FED. TRADE COMM'N, *THE PETROLEUM INDUSTRY: MERGERS, STRUCTURAL CHANGE, AND ANTITRUST ENFORCEMENT* (2004); Christopher T. Taylor & Daniel S. Hosken, *The Economic Effects of the Marathon-Ashland Joint Venture: The Importance of Industry Supply Shocks and Vertical Market Structure* (U.S. Fed. Trade Comm'n Bureau of Econ. Working Paper No. 270, 2004), available at <http://ssrn.com/abstract=532427>.

66. Government Performance and Results Act of 1993, Pub. L. No. 103-62, 107 Stat. 285.

67. See PETER W. HUBER, *THE GEODESIC NETWORK: 1987 REPORT ON COMPETITION IN THE TELEPHONE INDUSTRY* (1987) (evaluating the effects of the AT&T divestiture).

68. Gregory J. Werden et al., *The Effects of Mergers on Price and Output: Two Case Studies from the Airline Industry*, 12 *MANAGERIAL & DECISION ECON.* 341 (1991).

69. Mary W. Sullivan, *The Effect of the Big Eight Accounting Firm Mergers on the Market for Audit Services*, 45 *J.L. & ECON.* 375 (2002).

70. Indecopi's creation and early operations are examined in *PERU'S EXPERIENCE IN MARKET REGULATORY REFORM 1993-1998* (Beatriz Boza ed., 1998).

71. *THE ROLE OF THE STATE IN COMPETITION AND INTELLECTUAL PROPERTY POLICY IN LATIN AMERICA: TOWARDS AN ACADEMIC AUDIT OF INDECOPI* (Beatriz Boza ed., 2000) (presenting the results of an academic audit of Indecopi's competition policy programs).

72. See Paul A. Geroski, Chairman, Competition Comm'n, Remarks at the Opening of the Centre for Competition Policy at the University of East Anglia: Is Competition Policy Worth It? (Sept. 14, 2004), available at http://www.competition-commission.org.uk/our_peop/members/chair_speeches/pdf/geroski_uea_140904.pdf (reporting on research performed by the UK Department of Trade and Industry and the

5. *The FTC Vertical Restraints and Abuse of Dominance Impact Evaluations*

As mentioned above, in the late 1970s and early 1980s the FTC carried out a program to evaluate the consequences of vertical restraints cases and one abuse of dominance case commenced in the 1970s.⁷³ The impetus for the studies was a letter transmitted in 1978 by a congressional committee to the FTC Chairman, Michael Pertschuk, asking that the FTC perform case studies of vertical restraints cases. The Senate Subcommittee on Antitrust and Monopoly suggested that the FTC examine its experience with vertical restraints enforcement to explore the possibility indicated the previous year in *Continental T.V., Inc. v. GTE Sylvania Inc.*⁷⁴ that there might be an empirical basis for adopting a per se rule of illegality for various nonprice vertical restraints. Thus, the original impetus for conducting the studies was external, by a congressional request, rather than internal to the FTC.

The congressional letter led the FTC to devise a project to assess the effects of various vertical restraints cases initiated in the 1970s. The FTC chose to focus the vertical restraints evaluations of its past enforcement activities in five industries: shoes, blue jeans, audio components, industrial gases, and hearing aids. The cases in question addressed price and nonprice distribution practices, with the majority of cases challenging minimum resale price maintenance arrangements. In addition to reviewing vertical restraint matters, the agency decided to use the project as an occasion to evaluate the consequences of other areas of enforcement. The tentative initial plan was to include at least one case involving abuse of dominance—a major area of enforcement priority for the FTC in the 1970s—and at least one case involving horizontal mergers. Responsibility for carrying out the project was assigned to the Assistant Director for Planning in the FTC's Bureau of Competition, John Kirkwood, who assembled a team of two FTC attorneys and two FTC economists to manage the project.

The agency's decision to proceed with an evaluation of past competition cases was not well received by the professional staff of the Bureau of Competition, the FTC's antitrust enforcement unit. A central concern was that any study that questioned the wisdom or efficacy of the FTC's past interventions would be used to undermine its position in pending competition cases or, more generally, would be used to discredit the agency's competition policy work. A further concern was that any findings that found the FTC had achieved beneficial results in the cases studied would be dismissed by outsiders as the product of efforts by the FTC to manipulate the design and execution of the study to generate positive assessments. In short, outcomes favorable to the FTC's interests would be disregarded, whereas outcomes critical of the FTC's work would be embraced as evidence of institutional failure. Whatever the result, argued the staff of the Bureau of Competition, the agency could not "win."

To inform the design of the study, the FTC impact evaluation team retained two prominent industrial organization scholars, Richard Caves and Ben Klein, to prepare

UK Competition Commission).

73. By way of disclosure, the author of this Article was a member of the FTC team that oversaw the design and implementation of the impact evaluation project described here. The author was the principal FTC representative for the evaluation of the Xerox abuse of dominance consent decree. As recounted here, details of the project are based on the account provided in FED. TRADE COMM'N, *supra* note 60.

74. *Cont'l T.V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36 (1977).

research protocols. The protocols suggested approaches that researchers might use to evaluate the outcomes of FTC vertical restraint cases and sketched the sequence of investigative steps that the FTC might follow to conduct the studies.⁷⁵ Based on the Caves and Klein protocols and discussions with the FTC's Bureaus of Competition and Economics, the impact evaluation team decided that the case studies would have essentially four elements: (1) a review of publicly available material relevant to the case, including academic literature and trade press accounts that provided information about the effects of the FTC's intervention; (2) an examination of the FTC decision in the case, the FTC's internal files concerning the case, and the record of any administrative proceedings; (3) the preparation of a report evaluating the impact of the case based on items (1) and (2); and (4) recommendations for collection of further data to test more fully the conclusions presented in the researcher's written report.

The impact evaluation team then confronted the question of which researchers would conduct the studies. One possibility was to perform the work in-house by assigning the studies to economists in the FTC's Bureau of Economics. The agency rejected this alternative on the ground that reliance on FTC insiders would raise questions about the credibility and reliability of the results. This concern led the FTC to seek outside researchers to perform the studies. The Caves and Klein research protocols suggested the strategy chosen to enlist outside researchers: hire well-regarded and relatively junior academicians to perform the studies. Because the FTC could pay no more than \$10,000 per study, it was highly unlikely that more senior and better known economists could be persuaded to do the work.

With the assistance of Caves, Klein, and other outside academics, the impact evaluation team sought to identify economists who, though junior in their academic careers, had shown promise in the field of industrial organization economics. The choice of researchers from this set of candidates posed its own difficulties. Researchers often approach a topic of inquiry—for example, the analysis of vertical restraints—with a point of view shaped by training, personal philosophical perspectives, and past research. With a close reading of the researcher's previous written work and some inquiry into the individual's policy preferences, one could determine whether the researcher would be inclined to take a favorable view of the FTC's theory of the case and its decision to prosecute.

Had it chosen to do so, the FTC could have skewed the results of the studies by giving research contracts to academics with a generally positive view of vertical restraints enforcement or the prosecution of other cases, such as abuse of dominance matters, that the agency meant to study. Aware that such an approach easily could discredit the entire enterprise, the agency consciously sought to assemble a collection of researchers with a balance of philosophical perspectives.⁷⁶ Though the researchers disagreed about the value of vertical restraints cases, they shared the characteristic of having highly regarded technical skills and training. For example, the selection process produced a number of researchers, among them, Victor Goldberg, Howard Marvel, and Sharon Oster, who have

75. The Caves and Klein research protocols are reproduced in FED. TRADE COMM'N, *supra* note 60.

76. FED. TRADE COMM'N, *supra* note 60, at 9 (reporting that the FTC impact evaluation team selected the vertical restraints researchers "on the basis of both their academic expertise, particularly in vertical restraints, and the diversity of perspectives they would bring to the project").

achieved considerable distinction in their subsequent professional careers. The impact evaluation team selected Timothy Bresnahan to perform the abuse of dominance study of the Xerox consent decree. Bresnahan was well on his way to becoming a leading figure in the field of industrial organization economics.

As the researchers assembled publicly available information to prepare their studies, the FTC's impact evaluation team gathered the FTC's records on the cases and supplied the relevant records to each researcher. The outside economists received the internal pre-complaint memoranda prepared by the Bureau of Competition and Economics, among the most important confidential agency records provided, concerning the decision to prosecute. These documents laid out and analyzed the results of the pre-complaint investigation and the proposed theory of the case.

To gain access to these and other nonpublic records, the researchers signed nondisclosure agreements that forbade the publication of nonpublic information. The FTC's contracts with the outside experts permitted the researchers to publish their work and committed the agency to use its best efforts to publish the studies. The contract also required the researchers to provide, before submission for publication in journals or in other works, the FTC with drafts of their papers to enable FTC attorneys to ensure that nonpublic material had been excised. Subject to this restriction, the researchers could freely publish their work, and several researchers did publish versions of their studies in professional journals.⁷⁷

Beyond the process of omitting nonpublic information, the researchers had to submit an initial draft for substantive review and comments by the FTC staff. To a considerable extent, the FTC's review of the drafts focused on technical issues—such as the accuracy of the researcher's characterization of the FTC's internal deliberations and the agency's theory of the case, the completeness of the researcher's literature review, and the thoroughness of exposition. At the same time, the preliminary conclusions included in some of the initial drafts rekindled the anxiety of some FTC attorneys who had opposed the initiation of the evaluation project. The researchers issued some favorable assessments of the FTC's cases, but some were decidedly unfavorable. The receipt of the initial drafts again raised the question of whether the agency should carry the project to its intended conclusion.

The impact evaluation team and other internal FTC reviewers pressed the researchers, as well as those sympathetic and hostile to the FTC's cases, to defend their methodology and conclusions. Some discussions with the academics who authored unflattering assessments of the FTC's work were contentious. Despite intense disagreements between the FTC and some researchers and internal FTC discomfort with the unfavorable evaluations, the FTC fulfilled its commitment to publish the results and permitted the researchers to seek publication of their papers.

From a number of perspectives, the impact evaluation project was a success. The published studies presented highly informative critiques of the FTC's economic theories in each case, offered a useful preliminary assessment of the likely effect of FTC intervention, and proposed a sensible methodology for conducting further empirical study. The studies yielded valuable insights into the cases examined, made strong

77. For two prominent and oft-cited examples see Bresnahan, *supra* note 61, and Howard P. Marvel, *Exclusive Dealing*, 25 J.L. & ECON. 1 (1982).

contributions to our understanding of price and nonprice vertical restraints, and the compulsory patent licensing remedy imposed to resolve the abuse of dominance prosecution against Xerox.

An intended element of the original impact evaluation program—the pursuit of follow-on work to conduct further empirical testing—never took place, as the FTC chose not to allocate additional resources for these undertakings. Thus, an inherent limitation of the studies is their relatively thin empirical foundation. The outside researchers were able to assess the coherence of the FTC’s theory of the cases and to evaluate the wisdom of its decision to prosecute in light of evidence available to the FTC at the time the complaints were issued. Beyond collecting publicly available information about industry developments, the researchers were unable to conduct empirical tests of their hypotheses about the likely economic consequences of each case.

From an institutional perspective, the FTC’s willingness to perform the studies and its actual execution of the project seems to have enhanced the agency’s reputation by demonstrating its commitment to subject enforcement choices to rigorous analysis by outsiders. At the same time, one must acknowledge that the performance of ex post assessments does come with internal costs for the agency that go beyond the allocation of resources to perform and manage the studies. The pursuit of the project did arouse internal opposition. In one instance, a respondent in a pending vertical restraints case introduced the critical assessment of an FTC case by one outside researcher to bolster its case for exculpation. The FTC ultimately decided to dismiss its complaint in the matter,⁷⁸ though it is difficult to tell what impact the researcher’s study had upon the agency’s ruling.

C. Process-Oriented Evaluations of Competition Agency Operations and Organization

The second category of modern evaluation experience takes the form of efforts to analyze key elements of the process and operations of competition agencies. Unlike the case-specific studies described above, the process-oriented reviews do not seek to analyze the competitive consequences of individual agency interventions. The focus of process-oriented studies has been to evaluate the quality of agency decisionmaking and to test the soundness of procedures used to carry out the agency’s responsibilities.

1. Process-Based Evaluations by Competition Agencies

The past quarter-century has featured extraordinary innovation and change in the design of competition systems and the management of competition agencies. The growth in the number of new systems, a phenomenon often depicted as a source of troublesome complexity, has spurred a remarkable process of experimentation with institutional design and unprecedented attention to the role that the choice of institutional arrangements plays in shaping substantive policy results. This period has witnessed dramatic upgrades in jurisdictions that were among the first nations to adopt competition laws, such as Australia and Canada, as well as in other jurisdictions that are more recent arrivals to the community of nations with competition laws, such as South Africa and South Korea. It is the rare jurisdiction that has not taken steps in recent years to retool

78. *In re Beltone Elec. Corp.*, 100 F.T.C. 68, 221 (1982).

important aspects of the structure, responsibilities, or remedies of its competition agency.

In recent years, the Competition Directorate of the European Commission (DG Comp) has conducted a number of formative studies of its own procedures and, more generally, of the operation of the EU's system of competition law. In many instances, these assessments of process and policymaking have inspired far-reaching reforms.⁷⁹ In the period since 2000 alone, the results flowing from DG Comp's process-oriented evaluations have included the creation of the position of chief economist within the top management tier of the institution, the establishment of "devil's advocate" panels to test the strength of theories and evidence on which cases might be based, an internal reorganization that, among other effects, redistributed authority for the review of mergers, and a modernization program that decentralizes key decisionmaking functions to the national competition authorities of the EU member states and vests the national courts with expanded adjudication responsibilities.

The EU also has actively engaged in broad programmatic reviews of important areas of DG Comp policymaking. One focal point of attention has been the retooling of enforcement protocols, including the issuance of new guidelines concerning the licensing of intellectual property.⁸⁰ Still more recently, DG Comp has completed a substantial inquiry into the implementation of its remedies in competition cases.⁸¹

Another noteworthy example of process-oriented reform is the Office of Fair Trade (OFT) in the UK. A basic reassessment of the adequacy of remedies for serious antitrust offenses led OFT to develop and gain acceptance for proposals to treat the creation of cartels as crimes. Reflecting on its experience in dealing with competition issues arising in formerly or currently regulated sectors, OFT has played a leading role in devising innovative approaches to improving relations with sectoral regulators and coordinating operations. OFT leadership also has been engaged in internal assessments about how best to organize the case investigation and litigation.

A final representative illustration comes from the FTC's experience in the 1990s of evaluating the effectiveness of its process for designing remedies for mergers.⁸² The study constituted the FTC's first effort to evaluate the quality of its merger remedies since the adoption of the Hart-Scott-Rodino Antitrust Improvements Act in 1976, which established the modern U.S. framework for advance notification of certain mergers.⁸³

79. The dramatic modern reforms of the European competition policy system since 2000 are detailed and analyzed in WOUTER P.J. WILS, *PRINCIPLES OF EUROPEAN ANTITRUST ENFORCEMENT* 1-59 (2005).

80. Philip Lowe & Luc Peepkorn, *Singing in Tune with Competition and Innovation: The New EU Competition Policy Towards Licensing*, in *INTERNATIONAL ANTITRUST LAW & POLICY* 265 (Barry E. Hawk ed., 2005).

81. See DG COMP, EUROPEAN COMM'N, *MERGER REMEDIES STUDY* (2005), available at http://www.ec.europa.eu/comm/competition/mergers/others/remedies_study.pdf.

82. See BUREAU OF COMPETITION, *supra* note 63. The FTC's 1999 merger remedies study is one element of a larger collection of merger control process reviews that the U.S. agencies have undertaken, beginning in the early 1990s and continuing until the present. See Timothy J. Muris, Chairman, Fed. Trade Comm'n, Remarks at the George Mason University Law Review's Winter Antitrust Symposium: Improving the Economic Foundations of Competition Policy 11-13 (Jan. 15, 2003), available at <http://www.ftc.gov/speeches/muris/improveconfoundatio.htm> (describing FTC studies of effects of mergers).

83. The congressional proponents of the merger reforms adopted in the 1976 legislation relied extensively on the work of academic researchers who found that divestitures imposed as remedies for consummated mergers seldom achieved their intended remedial goals. See Elzinga, *supra* note 32 (discussing the antimerger provision of the Clayton Act).

Among other consequences, the creation of compulsory notification requirements and mandatory waiting periods facilitated more extensive recourse to settlements that used divestitures and various conduct-related undertakings to resolve discrete competitive concerns.

The FTC study was an internally generated initiative that emerged from research being conducted by the Compliance Division of the FTC's Bureaus of Competition and Economics. In their experience with mergers, the professional staff of these operational units had identified a number of potential weaknesses in FTC remedy practice. Existing practice dictated that the FTC track compliance with divestiture obligations only to the point that the agency approved a specific divestiture proposal and the merging parties actually divested the assets in question. The FTC's process took no account, in any systematic way, of whether the purchaser of the divested assets eventually used the assets to make sales in the relevant market or whether the purchaser's participation in the market with the divested assets had any competitive effect. In addition to its own experience, the FTC staff was aware of a wealth of academic literature that doubted the effectiveness of many merger remedies, including remedies achieved after the implementation of the Hart-Scott-Rodino reforms in the late 1970s.⁸⁴

Reflecting on these experiences, the FTC's professional staff in 1995 recommended that the agency carry out a systematic assessment of its divestiture process. Compared to earlier FTC research on merger remedies, the novelty of the proposed study was its emphasis on questioning the buyers of divested assets about the ultimate disposition and use of those assets. The FTC carried out the study in two parts. It first undertook a pilot study to test its methodology, which consisted mainly of open-ended telephone interviews of purchasers of assets divested pursuant to settlements in merger cases. After accounting for the results of the pilot inquiry, the FTC carried out a more comprehensive assessment and examined 35 orders entered from October 1989 through September 1994 in which it had required the divestiture of assets. The chosen five-year period permitted the agency to examine divestitures whose effects would have had an impact in the market and that took place recently enough that the buyers would have relatively clear memories of the divestiture experience. The 35 cases also involved a broad range of industries, purchasers, and divested assets.

The FTC relied mainly on a team of its own attorneys and economists to conduct the study, although it obtained comments on drafts from a business school professor, David Ravenscraft, who had extensive experience with data collection and analysis of mergers and acquisitions. The FTC staff identified 50 buyers who had purchased assets divested in the 35 cases and interviewed 37 of them. The staff also interviewed eight of the merging party respondents and two third parties. The FTC published the results of its inquiry in 1999 and highlighted four findings: (1) three-quarters of the divestitures examined in the study "succeeded to some degree;"⁸⁵ (2) settlements that compelled the divestiture of ongoing businesses succeeded more frequently than divestitures of discrete assets; (3) the possibility of continuing entanglements and relationships between purchaser and seller required close attention by the competition agency, as such links

84. The relevant literature is collected and reviewed in Mikhail S. Kouliavtsev, *Some Empirical Evidence on the Effectiveness of Antimerger Relief in the United States*, 43 *ECON. INQUIRY* 370 (2005).

85. BUREAU OF COMPETITION, *supra* note 63, at 8.

sometimes posed unexpected problems for the purchasers, and sometimes may have been critical for the purchaser's success in using the divested assets; and (4) smaller firms succeeded at the same rate as larger firms in using the divested assets.

The 1999 divestiture study was highly informative for the FTC and other competition agencies. The study's results inspired significant adjustments in the agency's evaluation of settlement proposals involving divestitures, and altered the types of demands that the FTC subsequently made in settlement negotiations. The study led a number of competition agencies in other jurisdictions to reevaluate their own approaches to remedies in merger cases and, in some instances, to undertake their own variants of the FTC's inquiry.

In certain respects, the study's aim was relatively modest. It did not measure the actual competitive effects of FTC merger remedies only, but rather attempted "to draw conclusions about whether the buyer of the divested assets was able to enter the market and maintain operations."⁸⁶ Yet even so limited an inquiry is valuable. If a basic assumption of an agency's merger remedy policy is that divested assets will remain in the relevant market and assist a firm other than the merging parties to compete for sales, it is worth knowing whether the purchaser succeeded in using the assets to accomplish any sales. If the answer to this question is "no" in more than a trivial number of circumstances, the agency needs to rethink its assumptions about the efficacy of its merger control policy. If the answer is typically "yes," it is still helpful to know what makes for a successful divestiture and to use that knowledge to improve merger solutions in the future.

Perhaps the greatest limitation of the FTC study is its cursory description of the criterion used to measure the success of the purchaser in deploying the divested assets. The FTC reported that of the 37 divestitures it studied, "28 appear to have resulted in viable operations in the relevant market."⁸⁷ The FTC added that in each of the 28 successful cases, "the approved buyer acquired the assets, began operations, and was operating in the relevant market within a reasonable time."⁸⁸ The FTC study did not specify what practical test it used in deciding whether a firm had engaged in "viable operations"—a term that might encompass a single sale, or thousands of sales, in the relevant market.

The experience of DG Comp, OFT, FTC, and other agencies underscores several important elements of effective process-oriented review. Two characteristics stand out. The first is the importance of the competition agency in stepping back and undertaking a deliberate and comprehensive assessment of the efficacy of specific programs or procedures. A meaningful review of an agency's practices cannot be done in a half-hearted manner. In all instances, the competition agency in question was willing to devote some of its best human resources to examine selected areas of operation and to suggest improvements. The second, a closely related characteristic, is the role that comparative study has played in suggesting paths for reform. Benchmarking across jurisdictions served to illuminate operational or programmatic alternatives and supplied an experience base with which to assess the qualities of such alternatives.

86. *Id.* at 9.

87. *Id.* at 10

88. *Id.*

2. Process-Oriented Evaluations by Multinational Bodies

The peer review exercises, described in Part IV.A.4 above, have provided an important means for process-oriented evaluation. The positive reaction to peer review projects that international organizations have received from their members and affiliates suggests that competition agencies would welcome greater efforts by the larger and global multinational bodies, such as the International Competition Network, OECD, or UNCTAD, to focus more attention on issues of agency management and operations. Among many competition agencies, there appears to be a keen desire to meet with peer institutions to discuss, analyze, and benchmark each other with respect to basic operational matters. Important focal points for such engagements might include the process for developing an agency-wide strategic plan, the optimal approach for organizing the agency's staff, whether it is organized along functional lines, by sector, or with an independent unit for economists, the design of internal quality control measures, the relationship of the legal services department to case handling units, and the best way to coordinate activities with other government bodies, such as sectoral regulators, with substantive portfolios that overlap the portfolio of the competition agency.

In multinational gatherings, the extensive attention given to the question of *what* competition agencies should do, whether they are to give top priority to prosecuting cartels or to apply a dominance test or a substantial lessening of competition standard in merger control, has tended to overshadow the equally important question of *how* they should do it. In the formal meeting rooms, delegates often discuss whether one analytical concept is superior to another. In the conversations during breaks, meals, or social gatherings, the delegates frequently ask each other how their home agencies decide what to do and press for practical details about how common administrative tasks are carried out. There is considerable room for multinational bodies to serve their members' interests by putting questions of management and internal procedure on the agenda more frequently and more prominently.

3. The OECD Performance Management Initiative

One particularly noteworthy initiative by a multinational body to improve performance measurement is the OECD Competition Committee's recently established Institutional Assessment and Development Project. The OECD project seeks to use lessons from performance measurement experience in private and public institutions in recent decades to prepare an instrument for assessing the quality of internal management and procedures of competition authorities and for indicating areas for improvement. As tested in various public and private settings, the key steps of the performance management methodology are (1) the articulation of the "core principles" of the organization to be studied; (2) selecting types of behavior that indicate whether the organization is adhering to the core principles; (3) creating an internal assessment procedure that uses the behavioral criteria to measure the organization's operations; (4) conducting an assessment of the organization, focusing mainly on the organization's managers but also involving consultation in some instances with external experts; (5) using the results of the assessment to improve the structure or decisionmaking processes of the organization; and (6) repeating the assessment on a regular basis to evaluate the implementation of reforms and to identify further areas for improvement.

To adapt this framework to the assessment of competition agencies, OECD consulted various individuals—“key informants”—with expertise in the operations of competition agencies to identify the operational characteristics of successful competition agencies and prepared an assessment instrument, principally, a questionnaire, to apply to individual competition agencies. The aim of the project is to develop a methodology that competition agencies can apply systematically to evaluate existing procedures on a continuing basis over time and to incorporate the results of each inquiry into the formulation and execution of reforms. As tested and applied within a number of individual jurisdictions, the project seeks to develop a more scientific and empirically based foundation for competition agencies to raise the quality of their management and operational routines.

Two potential contributions of the project deserve emphasis. First, the very effort to identify the “core principles” and key behavioral characteristics could provide valuable insights into what separates a “good” or “superior” competition agency from a merely adequate institution.⁸⁹ Views about the quality of competition agencies often are based on idiosyncratic impressions about the volume and type of an agency’s outputs—mainly the prosecution of cases. The initiation and litigation of cases provide readily observable events for analysis by academics, practitioners, and journalists. To a large degree, enforcement actions, especially “big” cases involving easily recognized respondents, commonly are assumed to be the proper measure of what a government agency has done.⁹⁰

In critical respects, the case-centric vision of competition policy is unduly narrow. Taken on its own terms, the case-centric perspective gives excessive weight to spectacular matters involving easily recognized respondents and routinely undervalues the “small” case that can make “big” law.⁹¹ By lauding the enforcement official who files the enforcement action, the case-centric view of competition policy tends to overlook what the cases actually accomplished. This is the equivalent of measuring an airline’s quality by the number and size of departing flights without devoting equal attention to how, when, and where the airplanes descend to earth.

Equally serious weaknesses of case-centrism are its disregard for non-litigation policy instruments and its indifference to investments in institutional capability that are vital to successful policymaking. Case-centrism treats advocacy, for example, the filing of comments on proposed legislation or administrative regulations, as a feeble substitute for prosecuting cases. Nor does case-centrism respect an agency’s commitment to build the intellectual capital that informs the choice and execution of litigation and non-litigation interventions alike.

89. The 2003-2004 Annual Report of the Australian Competition and Consumer Commission illustrates how a competition authority can identify the “core principles” that motivate its efforts and can specify the behavioral traits by which it measures fulfillment of such principles and invites outsiders to assess its progress. AUSTRALIAN COMPETITION & CONSUMER COMM’N, ANNUAL REPORT 2003-04 (2004), available at http://www.accc.gov.au/content/item.phtml?itemId=607851&nodeId=file4240b7666fed2&fn=ACCC%20AR2003-04_whole%20report.pdf.

90. See generally Kovacic, *supra* note 12 (emphasizing the cumulative nature of competition policymaking).

91. Timothy J. Muris, Chairman, Fed. Trade Comm’n, Address Before the American Bar Ass’n Antitrust Section: How History Informs Practice (Nov. 19, 2003), available at <http://www.ftc.gov/speeches/muris/031120compolicy.pdf>.

The antidote to case-centrism is a better awareness of the determinants of effective competition policy. Recent developments in the academic literature and public policy display a growing recognition of the role that institutional design and capability play in shaping the quality of government competition programs.⁹² A virtue of the OECD assessment project is its recognition that the “outputs” of competition agencies are not generated spontaneously. Instead, outputs come into being only through a combination of managerial choices and operational procedures. Projects that encourage agencies to strengthen the quality of inputs have true promise to raise the quality of outputs. Over time, good technique tends to beget good results.

The second potential contribution of the project is the creation and assembly of a substantial base of knowledge about how competition agencies might best answer basic and recurring questions about how to structure, manage, and operate an institution. The creation and refinement of a commonly accepted standard for process-oriented evaluation would facilitate the analysis of individual jurisdictions and permit extensive comparisons across jurisdictions. If the project ultimately engages a large number of competition agencies, each participant will have a common frame of reference with which to discuss key operational and organization issues with its peers and a large base of experience with which to inform its own decisions.

In January 2005, the competition agency of Portugal agreed with OECD that Portugal would serve as the jurisdiction for prototyping the project. From January through April, representatives of the OECD collaborated with the Portuguese Competition Authority in applying the assessment tool. An initial summary of the results of the project was presented by Paul Maylon, an OECD consultant and a principal architect of the project, and officials from the Portuguese Competition Authority at the June 2005 Evaluation Roundtable of the OECD.

V. CONSTRUCTING A METHODOLOGY FOR CONDUCTING PERFORMANCE EVALUATIONS

This Part proposes methodologies that competition agencies might consider in using ex post assessments to derive insights about the substance and process of policymaking.⁹³ It is worth acknowledging at the outset the reasons why competition agencies might be disinclined to spend resources in evaluating the effects of completed interventions or operational procedures. An evaluation might not find that the agency has achieved favorable results. For example, the evaluation might reveal that specific enforcement initiatives either had no effect or yielded perverse results. A competition agency understandably might hesitate to undertake research that casts doubt upon the wisdom of its past actions and, perhaps, raises questions about current enforcement programs.

As suggested above, concern that the application of performance measurement techniques might unmask, and possibly make transparent to external observers, an agency’s miscalculations about the value of specific interventions is not a sound basis for

92. Kovacic, *supra* note 10, at 864-65 (recognizing that antitrust standards must remain attuned to the continually changing global economy); Timothy J. Muris, *Looking Forward: The Federal Trade Commission and the Future Development of U.S. Competition Policy*, 2003 COLUM. BUS. L. REV. 359, 399 (2003) (noting and addressing the challenges facing the FTC in making competition policy).

93. See KUSEK & RIST, *supra* note 3 (discussing how public institutions can design and implement evaluation methodologies).

dispensing with evaluation. There may be instances in which a government agency properly might withhold information about frailties or vulnerabilities in its operations out of concern that disclosure might assist outsiders in frustrating the attainment of important national goals.⁹⁴ Competition policy, by contrast, is more likely a field in which fuller disclosure and analysis of outcomes will stimulate public discussion that serves over time to improve the quality of competition policy. The social benefit of performance measurement in correcting enforcement or process weaknesses or channeling resources toward activities with the strongest contributions toward improving consumer welfare outweigh the immediate costs of reputational discomfort to the competition agency.

A second possible impediment to developing and applying performance measurement techniques is expense. To fund evaluations, a competition agency must spend resources that otherwise might be made available for the development of new cases and for the fulfillment of mandates imposed by law, such as the review of proposed mergers. As discussed above, expenditures for evaluation help insure that outlays for substantive programs are well spent. Evaluation helps inform the competition agency's decision about what to do next—just as navigation informs a captain's judgment about how to handle a ship. Without some dedication of resources to ex post evaluation, it is doubtful that a competition agency can have confidence that its enforcement programs are achieving appropriate ends, especially where the initiative in question, such as the proper approach to merger analysis in technologically dynamic sectors, can rely heavily on untested intuitions about the effects of intervention. Even where there is broad consensus about the social value of applying the substantive command, such as a prohibition against supplier cartels, ex post evaluation can identify ways in which a competition agency can improve its enforcement of the command. For these reasons, a routine allotment of funds for evaluation activities should be part of the competition agency's annual budget.

A third and more serious obstacle involves the design of the performance measurement process. Evaluating the effects of individual prosecutorial decisions can present difficult measurement and data collection issues. To decide that an agency ought to assess the quality of its organization and operational procedures does not indicate exactly how the assessment should be performed.

One response to this concern is that even simple analytical models supported by limited data promise to provide a more confident basis for policymaking than many competition authorities currently possess. The exercises undertaken by the EU and the United States to evaluate the implementation of merger remedies in one sense have involved relatively simple analytical techniques. The agencies have not sought in these exercises to measure the actual competitive impact of divestitures. Instead, DG Comp and the FTC have employed simple but informative methods to test their assumptions about the likelihood that a purchaser of divested assets would deploy those assets to compete in the relevant market. The effort to determine whether divested assets remained in the market and were used to compete against the merging parties gave the competition

94. In the context of discussing information security issues see Peter Swire, *A Model for When Disclosure Helps Security: What is Different About Computer and Network Security?*, 3 J. TELECOMM. & HIGH TECH. L. 163 (2004) (providing a highly informative model for analyzing when the disclosure of information about a system is likely to set in motion developments that increase or decrease the security of the system).

agencies considerably valuable information about the design of future merger remedies.

The balance of this Part provides a second response to concerns about methodology by proposing approaches that competition agencies might take to design new performance measurement systems or to enhance existing programs for ex post evaluation. This Part analyzes methodology issues that affect the design of case-specific evaluations and process-oriented evaluations alike, and highlights concerns that relate specifically to each form of performance measurement.

A. Constructing an Evaluation Methodology

A competition agency considering the establishment or enhancement of a performance measurement system faces at least three basic methodological issues. First, how should it go about framing methodological options and identifying superior evaluation techniques? Second, should the agency use its own employees to conduct the evaluation exercise, or should it enlist the participation of outsiders? Third, should the agency disclose the results of its evaluation exercises only to its own personnel, or should the outcome of the assessments be made available in some form to audiences outside the agency? Each issue is considered below.

1. Choosing a Process for Designing a Performance Measurement Methodology

The process for designing a case-specific or process-oriented performance methodology should involve contributions from competition policy insiders and outsiders. Agencies should envision the planning process as having four elements, described below.

a. Preliminary Discussions within the Competition Agency

The first element of the planning process should include internal discussions with the competition agency's management team to establish a consensus, in principle, for carrying out a performance measurement exercise and to identify a bureau within the agency that will be responsible for measurement projects. Many agencies have found that an appropriate location for this responsibility is an office dedicated largely to policy analysis functions. The initial intramural discussions should be informed by the agency's own initial reading of the literature on performance measurement for competition agencies. A policy office within the competition agency is a natural candidate to perform the literature survey.

b. Consultations with Expert Outsiders

The second key element in designing a methodology is the consultation of expert outsiders. This consultation can be done in a variety of ways. One approach tested in a number of jurisdictions is for the agency to consult a relatively small group of expert outsiders who are chosen for their familiarity with the issues posed by the form of evaluation the agency has in mind. In planning a case-specific evaluation project, the agency should consult industrial organization economists with experience doing empirical work. This was the methodology that the FTC followed in the late 1970s when it gave contracts to Richard Caves and Ben Klein to draft protocols to guide its

assessment of vertical restraints and abuse of dominance cases. In preparing a process-oriented review, the agency should seek advice from experts who are intimately familiar with the management and operation of a competition agency, as well as experts with extensive experience in conducting performance management reviews for public or private institutions.

A second repository of external expertise that ought to be tapped in this initial exploratory phase of planning is other competition authorities. As suggested in this background Article, competition agencies have accumulated valuable experience with performance measurement. Discussions with competition agency officials who have conducted such evaluations are a valuable and necessary element of initial preparation. The materials contained here and the country contributions to be presented at the OECD session provide a starting place for identifying the relevant competition agency personnel.

A third source of expertise external to the competition agency consists of other government institutions with experience in performance measurement. Increased concern in recent decades with improving the performance of public agencies has generated a number of innovations for assessing the quality of agency management systems and specific substantive outputs.⁹⁵ Within any single jurisdiction, a variety of government agencies other than government bodies traditionally associated with auditing functions may have valuable experience in designing evaluation programs. The foregoing approach to planning a performance measure exercise, which emphasizes consultations with a comparatively small number of advisors, can be carried out in a relatively quiet and nonpublic manner. A competition agency might prefer to begin with this low-key approach as a means of identifying possibilities best suited to its own circumstances and resources. This initial consultation phase by itself may provide the agency with a confident basis for launching the evaluation exercise without further steps to identify and vet options for performance measurement.⁹⁶

A competition agency could decide to engage in a more elaborate planning process that supplements the low-key and nonpublic consultation between competition agency staff and small numbers of expert outsiders. The more elaborate process could include issuing a preliminary project description, work plan, or simply posing methodology issues for public comment. A still more robust form of public involvement would be to host a public conference, seminar, or workshop to gather the views of academics, practitioners, business officials, consumer groups, and other government officials about the appropriate design and substantive focus of an evaluation exercise. Where the competition agency is uncertain about the resolution of technical issues, the public workshops also can be an occasion for eliciting commentary about the results of previous efforts by competition agencies to conduct performance assessments.

95. KUSEK & RIST, *supra* note 3.

96. For example, the World Bank has established a "Measuring Results" project that uses ex post assessments to evaluate the implementation of law reform initiatives and other development related measures that have received assistance from the Bank. A description of the Bank's Measuring Results project is available at <http://www.worldbank.org/wbsite/external/projects>.

c. Budgeting

The initial planning phase is an appropriate occasion to begin formulating a budget for the evaluation exercise. For the better funded competition agencies, good practice would dictate that some funds be earmarked annually for performance measurement. For the reasons described earlier in this paper, the agency should progress toward a norm that encourages the inclusion of some outlays for performance measurement in each budget cycle.

For poorly funded agencies that struggle unsuccessfully to satisfy even the most urgent operational needs, two possibilities come to mind. One is to rely on regional cooperation and regional associations to serve as focal points for the planning and, possibly, the execution of performance measurement exercises. This is one of a number of areas in which regional bodies can facilitate the pursuit of institution-building projects that might exceed the capabilities of any single competition agency. The second possibility is to approach donor agencies to seek subsidies for evaluation exercises. In recent years, national and multinational donors have given increased attention to performance measurement, and there is reason to believe that requests by competition agencies would not simply be brushed aside or deferred to infinity or beyond.

d. Identifying Restrictions on the Use and Dissemination of Confidential Records

The initial planning process should address restrictions on the use of confidential data in the performance management exercises, both the proprietary business records in the competition agency's possession and the confidential work product (such as memoranda prepared by professional staff and sensitive internal decisionmaking records) of the competition agency itself. Most, if not all, competition agencies will need to consider whether and with what conditions confidential records can be provided to agency outsiders who will participate in the evaluation exercises, and what restrictions must be observed in any public release of the results of an evaluation exercise. A further issue for some jurisdictions is whether the work product of the evaluation exercise itself must be revealed pursuant to the disclosure requirements of various public records statutes. For these and related questions, the initial process of planning an evaluation program ordinarily draws upon the advice of the agency's legal services department.

Experience in individual jurisdictions is likely to vary, but some general observations about the treatment of confidential records are possible. Most jurisdictions have mechanisms for permitting outside experts to gain access to confidential records upon signing nondisclosure agreements. The typical nondisclosure agreement forbids the outsider from revealing confidential information to specified categories of individuals, unless the competition agency expressly authorizes the disclosure. The nondisclosure agreement also permits academics and other commentators to write about the project, but requires that the author give the competition agency drafts of manuscripts to ensure that confidential information has not been disclosed. Competition agencies that have published their own reports or papers on evaluation exercises also have found effective mechanisms to supply informative accounts of their results without disclosing business or other information that must be kept secret.

The preservation of confidential data is another area in which competition agencies

have assembled considerable experience. An agency planning its first evaluation exercise can consult its peer organizations for advice about how to manage the flow of information used in and generated by an evaluation program. It is quite likely that agencies experienced in performing evaluation exercises can provide templates of nondisclosure agreements or supply descriptions of safeguards used to ensure that confidentiality obligations are observed.

2. Conducting Ex Post Assessments: Using Insiders, Outsiders, or Both

One evaluation methodology would rely solely on the competition agency's own personnel to conduct evaluations. The agency would establish a process for its own personnel to select completed enforcement initiatives (both litigated cases and consent agreements) or operational procedures and analyze their quality. For example, a case-specific evaluation would consist of having agency personnel review specific enforcement episodes in detail, studying the deliberative processes that led to the agency's intervention, interviewing those involved in the decision to prosecute, and collecting data on effects, such as by consulting customers or competitors of the respondents. Internal self-evaluation could be performed as a collaborative effort between case handlers, economic units, and policy bodies.

As suggested above, there are considerable advantages to engaging the talents of expert outsiders in performing assessments and not simply in the design of the measurement process. When they conduct or contribute to the actual assessment, an outsider can provide an expert perspective that the agency itself does not possess and can make the assessment more objective by bringing a presumably neutral point of view to the exercise. The outsiders could be nongovernment employees working under contract to the competition agencies, or could be employees of government institutions, such as national audit offices that occasionally examine competition agencies.

The level of participation by outsiders could range from more expansive forms of involvement, such as performing case studies, to more limited contributions, such as offering comments on studies prepared by agency insiders. Compared to relying solely on insiders, participation by outsiders also is likely to increase the exercise's credibility in the eyes of external groups. Even if an audit is not made public in any form, the agency probably will obtain a more meaningful perspective on its work if it engages outsiders on a confidential basis to participate in preparing the audit or commenting on reports that present the results.

One recent example that warrants attention by competition agencies is the International Development Research Centre (IDRC) project mentioned in Part III.F above. In this project, the Canadian government is funding research by scholars in seven developing countries⁹⁷ to identify and analyze factors that impede competition in the private sector of their economies. One important dimension of the IRDC study is the examination of how interventions by competition policy agencies in these countries have affected economic performance. To perform the country studies, the IRDC project director, Simon Evenett, has recruited a team of scholars in each developing country. Drafts of some country studies have been completed, and the project's aim is to make the

97. The nations are Argentina, Brazil, China, Egypt, India, Mexico, and South Africa.

studies and data sets assembled available to other scholars and analysts.

One could imagine a collaboration in which a competition agency canvassed existing case-specific evaluation research within its own borders and sought to form a partnership with the organization conducting the research. For existing or proposed projects of this type, the competition agency might offer its cooperation to the outside researchers, including access to agency personnel and discussions about publicly available information of which the outside researchers might not be aware. A still less drastic form of productive involvement would have the competition agency host or participate in conferences in which the results of such research are presented.

3. *Disclosure of Results*

One disclosure technique reveals the results of ex post assessments solely to enforcement agency personnel. For some competition policy agency insiders, this approach might seem to be the least threatening form of evaluation. If the findings are negative, only the agency will be aware of them.⁹⁸ An evaluation program by which the competition agency performs assessments but limits disclosure of the results to its own personnel is considerably better than no evaluation system at all. Even if an agency is unwilling to reveal evaluation results to outsiders, its chances of identifying and correcting flaws in case selection and procedure are likely to improve if it engages in a purely intramural exercise of candid self-assessment. As the agency uses and applies its internal evaluation methodology, it may become more comfortable over time in revealing at least some information about its assessments to outsiders.

The alternative to purely in-house distribution is to make evaluation results public in some form. The competition authority could issue public versions of the evaluations that delete references to sensitive information, such as confidential business data or material that discloses enforcement intentions regarding pending matters. A general presumption favoring public disclosure is warranted as a means of improving the substance and perceived legitimacy of a competition policy program, as well as the rigor of internal decisionmaking. Even limited forms of public disclosure can elicit useful suggestions from outsiders about improvements in the choice of interventions and procedure.

By its willingness to subject its operations and decisions to public discussion, the competition agency can increase public confidence in the competition policy system. The knowledge that the agency will engage in future public discussion about assessments of its work could supply additional motivation to agency managers and professional staff to carry out their responsibilities conscientiously.

Here it is useful to consider, again, the experience of the FTC with its study of merger remedies in the 1990s. The FTC divestiture study underscores the benefits of public disclosure.⁹⁹ As noted above in Part IV.C.1, the merger remedies study was conducted by FTC attorneys and economists, who consulted an outside academic about the design and execution of the inquiry. After deleting references to confidential and nonpublic information, the agency released the study to the public. The FTC remedies

98. Some jurisdictions may have public disclosure statutes that could require a competition agency to make such assessments, or redacted versions of them, available if requested. The design of a purely internal study would have to account for the types of information that might have to be revealed.

99. See BUREAU OF COMPETITION, *supra* note 63.

study was both informative and influential. Had the FTC chosen to conduct the study internally and make no public disclosure of its results, the exercise would have been valuable for what it told the FTC about needed improvements in the design of remedies. It is also clear that publication of the results stimulated a useful public debate about merger remedies and served to inform other competition agencies about how to improve merger control.

As a technique for disseminating the results of case studies and evaluations, competition policy agencies might commit themselves to engage in the detailed public analysis of specific cases or internal procedures that have been the subject of analysis by agency insiders or outsiders. Discussions of individual cases and of internal procedures that engage enforcement officials, private practitioners, business operators, and academics would increase understanding of the motivations for and consequences of individual enforcement decisions and agency procedures. Whether conducted in seminars, workshops, or conferences, the public give-and-take provides useful feedback to the competition agency and stimulates thinking about evaluation among external constituencies. Experience with discussion and debate between competition agencies and external groups concerning issues of doctrine and enforcement policy creates good reason to expect that a dialogue on matters relating to performance management would yield improvements in agency case selection and management.¹⁰⁰

B. Special Considerations for Process-Oriented Performance Measurement

The central task for applying a process-oriented evaluation methodology is to define what constitutes good internal agency procedures. Presented below is a suggested beginning for a list of good practice characteristics by which the quality of a competition agency's management techniques and procedures might be tested.¹⁰¹

1. Mechanism for Strategic Planning

To be effective, a competition agency, be it old or new, must have a conscious process for setting goals and creating steps to accomplish them. To do otherwise is to be the passive captive of external demands, whether in the form of complaints from consumers or business operators, or requests for action by public bodies. Even the most humbly funded competition agency must develop a strategic plan that defines what it will seek to achieve in the coming year or years.

100. See A. Douglas Melamed, *International Antitrust in an Age of International Deregulation*, 6 GEO. MASON L. REV. 437, 444 (1998).

It is important to keep in mind . . . that antitrust in [the United States] is based, not on a political or legislative code, but rather on a broad, general statute that has been sustained through what is really a common law process—by which I mean, not only a process of federal court litigation, but more broadly a dialog among academic and business communities, the enforcement agencies and the courts.

Id.

101. The discussion in Part V.B is derived in part from William E. Kovacic, *Achieving Better Practices in the Design of Competition Policy Institutions*, in *ON THE MERITS & CURRENT ISSUES IN COMPETITION LAW AND POLICY* 195 (Paul Lugard et al. eds., 2005).

2. Maintenance and Disclosure of Databases

Each competition agency should prepare and provide a full statistical profile of its enforcement activity. Good databases are indispensable to the tracking and analysis of an agency's activities over time. Despite their importance to performance measurement, the maintenance and public disclosure of comprehensive and informative databases on enforcement are relatively uncommon for competition policy agencies. Every authority should take the seemingly pedestrian but often neglected step of developing and making publicly available a database that (1) reports each case initiated; (2) provides the subsequent procedural and decisional history of the case; and (3) assembles aggregate statistics each year by type of case. Each agency should develop and apply a classification scheme that permits its own staff and external observers to see how many matters of a given type the agency has initiated and to know the identity of specific matters included in the category of enforcement activity.

Among other ends, a current and historically complete enforcement database would promote better understanding and analysis, inside and outside the agency, of trends in enforcement activity.¹⁰² For example, access to such databases would give competition agencies greater ability to benchmark their operations with their peers. For poorly funded institutions, this is another area in which regional or global organizations can make immediate and major contributions.

3. Explanation of Actions Taken and Not Taken

Competition agencies should take measures to progress toward a norm that favors explanations for all important decisions to prosecute or not to prosecute. One might define as "important" any matter in which a competition authority conducts an elaborate inquiry. The norm suggested here would dictate that the agency seek as often as possible to explain why it decided not to intervene following an extensive investigation.

4. Assessment of Human Capital

Continuous institutional improvement requires a competition agency to regularly evaluate its human capital. The capacity of an agency's staff deeply influences what it can accomplish. An agency routinely must examine the fit between its activities and the expertise of its professionals. Has the agency developed a systematic training regimen for upgrading the skills of agency professionals? If the agency is active in areas such as intellectual property that require special expertise, has it acquired the requisite specialized skills by hiring some patent attorneys? Do government statutes and regulations that control public sector employment permit the agency to recruit needed expertise in a timely manner?

102. See generally Richard A. Posner, *A Statistical Study of Antitrust Enforcement*, 13 J.L. & ECON. 365 (1970) (discussing a formative treatment of the value of good statistical records for the analysis of competition policy).

5. Investments in Competition Policy Research and Development

An essential element of continuous institutional improvement is the enhancement of the competition agency's knowledge base. In many activities, particularly in conducting advocacy, the effectiveness of competition agencies depends on establishing intellectual leadership. To generate good ideas and demonstrate the empirical soundness of specific policy recommendations, competition authorities must invest resources in "competition policy research and development."¹⁰³ Regular outlays for research and analysis serve to address the recurring criticism that competition policy lags unacceptably in understanding the commercial phenomena it seeks to address.

6. Recognition of Policymaking Interdependencies

Increasingly, efforts to formulate effective competition policy will require competition agencies to study more closely how other government institutions affect the competitive process. Many jurisdictions resemble a policymaking archipelago in which various government bodies other than the competition agency deeply influence the state of competition. Too often each policy island in the archipelago acts in relative isolation, with a terribly incomplete awareness of how its behavior affects the entire archipelago.

It is ever more apparent that competition agencies must use non-litigation policy instruments to build the intellectual and policy infrastructure that connects the islands and engenders a government-wide ethic that promotes competition. Building this infrastructure requires competition authorities to make efforts to identify and understand the relevant interdependencies and to build relationships with other public instrumentalities. In a number of instances, the study of collateral government policies will reveal how existing and proposed forms of public intervention impede competition. The capacity to study and document the effects of such impediments and to undertake advocacy initiatives to correct them will be important elements of an agency's program. On the report card we use to measure competition agencies, the suppression of harmful public intervention should weigh just as heavily as the prosecution of a case that forestalls a private restraint.

7. Benefits of Comparative Study

As suggested above, comparative study can play an informative role in improving the performance of individual competition agencies.¹⁰⁴ A competition agency should not consider adjustments in its own organization, procedures, or management techniques

103. See Muris, *supra* note 92 (analyzing the concept of "competition policy research and development" and its role in determining institutional capability); see also Akenori Uesugi, *Where Japanese Competition Policy is Going—Prospect and Reality of Japan*, in INTERNATIONAL ANTITRUST LAW & POLICY 57, 76 (Barry Hawk ed., 2005) (describing the Japanese Fair Trade Commission's 2003 creation of the Competition Policy Research Center to bolster its research capabilities).

104. See Wolfgang Kerber & Oliver Budzinski, *Competition of Competition Laws: Mission Impossible*, in COMPETITION LAWS IN CONFLICT 31, 36-40 (Richard A. Epstein et al. eds., 2004) (providing an excellent review of the use of international benchmarking as a way to improve the competition systems of individual jurisdictions, and for a less sanguine view than this Article holds of the potential net benefits of comparative benchmarking).

without examining experience abroad. Whatever the issue may be, whether analytical methodology, investigative techniques, personnel policy, or advocacy, foreign practice frequently has much to teach any competition authority.¹⁰⁵

One criterion by which to evaluate a competition agency's internal procedures is the degree of attentiveness to, and analysis of, relevant foreign experience. The ready availability in electronic form of foreign materials concerning many elements of competition agency enforcement and operations facilitates the comparative analysis contemplated here. Multinational organizations can spur this process by creating databases that describe and collect the results of evaluation projects carried out within individual jurisdictions.

C. Special Considerations in Case-Specific Evaluations

Past experience with the evaluation of specific competition agency interventions suggests considerations that agencies should take into account in carrying out the evaluation of cases or other actions.

1. Reviewing a Collection of Matters Rather than Analyzing Single Cases

An agency may learn more about its selection of cases or other interventions if it studies several related matters rather than focusing on a single intervention. For example, the FTC's study of vertical restraints cases in the late 1970s and early 1980s examined a range of matters within individual industries to develop a more general sense of the impact of its resale price maintenance and nonprice vertical restraints enforcement program.¹⁰⁶ It may not always be possible to study several matters of a specific type, and, as the FTC-sponsored study of the Xerox abuse of dominance settlement shows, the study of an individual case can be highly useful. Nonetheless, the design of a case-specific evaluation project should consider the possibility of studying two or more matters of a specific type.

2. Understanding Enforcement Choices in Context

Good case studies demand the skills of the historian and political scientist no less than the skills of an economist.¹⁰⁷ Researchers conducting case studies should be pressed to account for the context that motivated the decision to prosecute. In addition to evaluating consequences *ex post*, a good case study should seek to recreate as much as possible the *ex ante* assumptions that guided the agency's decision to intervene. This is particularly true for interventions that turn out to be ineffective or in some sense

105. See generally Kovacic, *supra* note 58 (describing what older competition authorities can learn from institutional innovations undertaken by newer competition authorities); Sheridan Scott, *Cartel Enforcement: International and Canadian Developments*, in INTERNATIONAL ANTITRUST LAW & POLICY 33, 41-48 (Bary Hawk ed., 2005) (discussing the value of comparative analysis and international benchmarking in informing Canada's consideration of reforms to its competition statutes).

106. See *supra* notes 73-78 and accompanying text (describing FTC vertical restraints evaluations).

107. See generally DAVID J. GERBER, LAW AND COMPETITION IN TWENTIETH CENTURY EUROPE: PROTECTING PROMETHEUS (1998) (discussing the importance of history and political science as sources of insight about the development of competition systems and as tools for understanding specific enforcement measures).

counterproductive. Agencies seldom engage in unsuccessful endeavors because they enjoy failure or take pleasure in retarding social progress.

Commentators sometimes attribute the failure of an agency's substantive interventions to the irrationality, primitiveness, or general perversity of its incumbent decisionmakers.¹⁰⁸ This explanation often is a mistaken, intellectually lazy answer to the question of why agencies fail. It ducks the harder and more perplexing question of why intelligent individuals made choices that prove to be seriously ill-conceived. As one scholar made the point: "Individual regulatory experiments and episodes must be judged against a standard true to the particular historical moment."¹⁰⁹

A case study may reveal that an agency made a faulty decision because it relied upon the wrong analytical model despite the availability of better models. One part of the researcher's task is to explain why the agency persisted in embracing the flawed model and neglected superior alternatives. The researcher should be pressed to suggest what types of institutional adjustments might serve to avoid being trapped on the wrong path in the future. This type of analysis is possible only if the researcher declines to build and demolish a strawman, and instead takes the ideas and institutional influences that guided the competition agency at the time of the decision to prosecute on their own terms.

3. Sensitivity to Institutional Capabilities and Long-Term Consequences

One factor in assessing a decision to intervene is the care with which a competition agency avoided serious mismatches between its institutional capability and the analytical demands posed by the case. In some sense, a public institution always will be forced to run at a level above its capacity. There is, however, a difference between trying to operate at 105% of capacity and 200% of capacity. The first condition is manageable. The second begets failure. Sensitivity to the longer term institutional demands associated with initiating a new matter and effort to achieve a rough match between commitments and capability ensure that incumbent managers do not impose significant negative externalities on successors and on their agencies as a whole.

D. Case-Specific Evaluations: Building Mandates into Settlement Decrees

Courts sometimes have played an important role in promoting efforts to assess the results of relief obtained in settlements. Since 1974, the federal district courts have been required to approve antitrust settlements in cases prosecuted by the DOJ.¹¹⁰ One provision of the settlement approved by Judge Harold Greene that resolved the government's monopolization claims in *United States v. AT&T Co.* required periodic evaluations of the efficacy of the relief negotiated between AT&T and the DOJ.¹¹¹ This requirement generated informative assessments of the effects of the AT&T divestiture and analyses of the need for continuing restrictions on the business operations of the successors to the Bell Telephone system.¹¹²

108. See Kovacic, *supra* note 12, at 383-84 (collecting authorities).

109. THOMAS K. MCCRAW, *PROPHETS OF REGULATION* 308 (1984).

110. See 15 U.S.C. § 16 (2000).

111. *United States v. AT&T Co.*, 552 F. Supp. 131, 194-95 (D.D.C. 1982), *aff'd sub nom.*, *Maryland v. United States*, 460 U.S. 1001 (1983).

112. For an example of the evaluations produced to satisfy the requirements of the AT&T settlement, see

The evaluation mandated by Judge Greene in *AT&T* in some ways resembled a command that Judge Charles Wyzanski incorporated into the decree issued at the close of the DOJ's monopolization case in *United States v. United Shoe Machinery Corp.*¹¹³ After finding that United Shoe had engaged in illegal monopolization, Judge Wyzanski imposed a number of controls on the defendant's future conduct. Paragraph 18 of the decree provided that ten years after the decree's effective date "both parties shall report to this Court the effect of this decree, and may then petition for its modification, in view of its effect in establishing workable competition."¹¹⁴ In the mid-1960s, Judge Wyzanski conducted the review contemplated in the 1953 decree.¹¹⁵

The settlement in *AT&T* and the decree in *United Shoe* provide models that enforcement agencies and courts might adopt for other antitrust settlements. To ensure that the consequences of settlements, especially provisions of uncertain effect, receive subsequent scrutiny, an enforcement agency or a court could mandate periodic review of their effects. Where settlements are not subject to judicial approval, the competition agency might use its administrative discretion to incorporate mandatory assessment provisions in at least a sample of its orders. The certainty of occasional ex post review would add discipline to the negotiation of settlement terms and increase the empirical foundation for designing antitrust remedies in the future.

VI. CONCLUSION

In 1940 Walton Hamilton and Irene Till authored a gloomy assessment of the first 50 years of experience with antitrust enforcement in the United States. The two scholars observed that "[t]he Sherman Act has been called a 'charter of freedom' for American industry. Why has it not been a success?"¹¹⁶ Among other findings, they concluded that future improvements would require the correction of various institutional flaws of the U.S. competition agencies. In particular, Hamilton and Till criticized what they believed to be the competition agencies' preoccupation with bringing the next case without the benefit of an internal management system that selected matters on the basis of continuing industry analysis, used such analysis to frame effective remedies, and monitored results.¹¹⁷

Since Hamilton and Till reviewed U.S. experience 65 years ago, competition agencies have evolved considerably and have accomplished a great deal. Despite the passage of time, their emphasis on the importance of sound institutional design and operations as keys to the substantive success of competition agencies rings true today. In

HUBER, *supra* note 67.

113. *United States v. United Shoe Mach. Corp.*, 110 F. Supp. 295 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954).

114. *Id.* at 354.

115. Emphasizing that United Shoe's market share fell from 85% in 1953 to 62% in 1964, Judge Wyzanski rejected the government's request that the company be split into at least two successor firms. *United States v. United Shoe Mach. Corp.*, 266 F. Supp. 328 (C.D. Mass. 1967). The Supreme Court later reversed Judge Wyzanski's ruling, stating that he lacked discretion to impose a divestiture in this instance. *United States v. United Shoe Mach. Corp.*, 391 U.S. 244 (1968).

116. Staff of S. Temporary Nat'l Econ. Comm., 76th Cong., Investigation of Concentration and Economic Power: Antitrust in Action 4 (Comm. Print 1940) (written by Walton Hamilton and Irene Till).

117. *Id.* at 30-35.

discussions of competition policy there is a natural tendency to focus on current issues of doctrine and enforcement policy. These issues, after all, often have an immediate and powerful intellectual appeal. Amid discussions about the pressing substantive issues of the moment, it is easy to lose sight of the practical questions that beset efforts to formulate and execute competition policy programs. How should an agency set priorities? What is the best way to organize the investigation and litigation of cases? Which internal quality control mechanisms work best?

The exercises of preparing performance measures and conducting evaluations provide valuable tools for answering these critical questions about the administration of competition policy. The assessment of outcomes of substantive interventions can generate useful information about such basic matters, such as the choice of cases and the design of remedies. The routine evaluation of internal procedures supplies regular opportunities to determine whether the agency's organizational infrastructure and management techniques put the competition authority in the best possible position to select promising substantive initiatives and bring them to a successful close. A competition policy system is only as good as the institutions entrusted with its implementation.¹¹⁸ By emphasizing internal review and improvement, a competition agency puts proper emphasis on the expansion of its knowledge base and institution building as indispensable predicates to success.

Evaluations may indicate needed adjustments in the competition agency's statutory authority. In recent decades, many competition authorities have sought and obtained important enhancements in the framework of laws, and there is every reason to believe that a key to effectiveness over time will be the installation of periodic upgrades to account for past experience and new conditions.¹¹⁹ A program of performance measurement and evaluation can supply a better empirical foundation for designing and justifying needed changes.

Evaluation promises to play a larger role that extends well beyond a competition authority's intramural decisions about case selection and management. The legitimacy of government efforts to enforce competition policy commands depends substantially on the ability of competition policy authorities to demonstrate to a variety of external observers (e.g., legislators, business managers, and the public at large) that chosen forms of intervention to curb restraints on business rivalry improve economic performance. Neither the intuitions of public enforcement officials nor the hypotheses of academic theorists are likely to provide a confident basis for outsiders to conclude that the

118. See Jean-Jacques Laffont, *Competition, Information, and Development*, in ANNUAL WORLD BANK CONFERENCE ON DEVELOPMENT ECONOMIES 237 (Boris Pleskovic et al. eds., 1999) (describing how the quality of implementing institutions determines the success of a competition policy system); Timothy J. Muris, *Principles for a Successful Competition Agency*, 72 U. CHI. L. REV. 165, 176 (2005) (discussing institutional characteristics, including investments in "competition policy research and development," that improve a competition agency's prospects for success).

119. See, e.g., PHILIP CLARKE & STEPHEN CORONES, *COMPETITION LAW AND POLICY: CASES AND MATERIALS* 8-16 (1999) (describing modern reform of Australia's competition system); Paul Collins & Jeffrey D. Brown, *National Antitrust Laws in a Continental Economy: A Comparison of Canadian and American Antitrust Laws*, 65 ANTITRUST L.J. 495 (1997) (reviewing the enhancement of Canada's competition laws in the 1980s and their subsequent implementation); Kris Dekeyser & Celine Gauer, *The New Enforcement System for Articles 81 and 82 and the Rights of Defence*, in INTERNATIONAL ANTITRUST LAW & POLICY (Barry Hawk ed., 2005) (describing the new EU framework for enforcement of Articles 81 and 82 of the EC Treaty).

government agency has selected the correct mix of policies or chosen the ideal procedures to execute programs. Competition policy institutions face a continuing need to justify their value, and one cannot expect critical observers to be satisfied with only unsubstantiated assurances of efficacy.¹²⁰

Larger multinational bodies, such as ICN, OECD, and UNCTAD, and smaller regional associations have taken promising steps to promote the formulation and application of performance measures. Case seminars and peer review exercises have been important sources of evaluation, and the regular meetings of these bodies have supplied useful venues for individual agencies to share know-how about the results of their own evaluation efforts. Considerable work remains to encourage competition agencies to accept a norm that treats evaluation as a vital tool for ensuring that government competition programs achieve desired ends. In the past, the evaluation initiatives of the major international competition policy institutions have focused on performing evaluations for their members. The focus for the future has to shift more toward projects that assist and prod competition agencies to devise and apply their own performance measurement methodology.

Ex post evaluation of results is important for all institutions, but it is especially critical for competition agencies. Competition policy relies heavily on experimentation and evolutionary adjustment to ensure that doctrine and enforcement policy make defensible distinctions between benign and dangerous commercial practices. Even for the most experienced and self-assured competition agency, competition policy is a work in progress, and performance measurement provides a vital element of the process by which agencies test new techniques and adapt proven methods to new circumstances. The aphorism often applied to the growth of individuals rings true in this field of public policy: It is what you learn after you know it all that really counts.

120. See KATZMANN, *supra* note 37, at 205 (“[W]ithout studies indicating whether antitrust policy is technologically capable of achieving various economic goals, government is vulnerable to the charge that antitrust is a charade or a lightning rod that absorbs the frustrations of those who might otherwise push for greater state intervention in the economy.”).