RATING THE COMPETITION AGENCIES: WHAT CONSTITUTES GOOD PERFORMANCE?

William E. Kovacic[*]

INTRODUCTION

What is a good competition agency? Among competition policy specialists, this topic often emerges in casual conversation and scholarly debate. For all the attention the subject receives, discussions about agency

tury."[2] One Obama appointee to the federal antitrust agencies later claimed that "inadequate antitrust oversight" contributed to the economic crisis in the United States.[3] Candidate Obama pledged his administration "to reinvigorate antitrust enforcement,"[4] a phrase that his supporters repeated frequently in subsequent commentary.

DOJ provided the chief target for opprobrium. Many commentators said, at least in private conversations, that they meant to aim their criticism

From these results, one might conclude that the U.S. federal agencies in this decade have performed in a satisfactory manner, if not well above average. This is not the perspective of commentators noted above, including the current president of the United States. The Obama campaign statement contained no suggestion that the Bush administration antitrust agencies might be deemed "elite." Rather, the presidential candidate said the Bush administration antitrust enforcement program "may be the weakest" of any presidential administration since the late 1950s.[5] As indicated above, many others offered similarly unflattering assessments.

Could all of these observers have been talking about the same federal agencies? Has the FTC been an "elite five-star" agency, or is it part of the weakest antitrust program of the past half-century? Has the Commission performed at the top of its game during the Bush administration, or has it been dozing under the guidance of leaders whose ideological rigidity induces them to disregard facts constantly?

The modern commentary about the quality of U.S. competition policy begs for answers to two basic questions. First, by what criteria should the performance of competition agencies be judged? The lack of widely-accepted, consistently applied standards for assessing the quality of agency performance has afflicted the field of competition policy throughout its history, and the absence of such standards is a major impediment today to achieving consensus on what competition authorities ought to do.

Second, once the criteria for the agency report card have been set, how should they be applied in order to determine the grades? It is impossible to have a constructive conversation about agency performance without a common view about how to answer these questions. Without a common framework, there is no meaningful way to score agency performance in any one period or across time. U.S. competition policy is in serious trouble today if we cannot agree, as then candidate Obama suggested we cannot, that federal competition policy in this decade is at least as good as it was, for example, in the 1960s. Do commentators sincerely advance the view that the FTC of this decade does not surpass the agency of the 1950s and 1960s—an agency damned by many observers as gravely deficient?

Why care about the establishment and application of meaningful standards? Assessments of agency performance are important for several reasons. Agency reputations can be likened to brands, and having a well-respected brand is an extremely valuable asset. Current perceptions of agency quality influence legislative decisions about budgets and additions to the agency's statutory authority, judicial decisions about whether to defer to an agency's positions, judgments by company officials about whether to comply with mandates subject to the agency's supervision, the level of mo-

new staff. A broadly held view that an agency is fulfilling its duties capably also contributes to citizen confidence in public governance and thereby strengthens the legitimacy of public administration.

A great deal of good policy is the result of cumulative, incremental improvements over time.[6] These improvements progressively enhance the agency's brand. The development of a strong brand is a slow growth that requires sustained contributions by agency leadership over time. The characterization of an agency's work as severely deficient in any one period not only can diminish hard-earned reputational capital, but also can induce incumbent leadership to disregard positive developments in an era that incorrectly is said to be deficient. If a period of public administration is said to be mediocre, new leadership might be inclined to indiscriminately write off initiatives pursued in that period and to devise new programs from scratch.

This presentation discusses the assessment of agency performance in two parts. It first discusses what the criteria for evaluating a competition agency should be. It then considers how the report card should be applied in practice. In setting out the design and application of evaluative criteria, the presentation emphasizes investments in achieving superior institutional design and enhancing agency capability. These are long-term capital investments that provide the foundation for the identification and execution of successful programs. The returns to such capital investments tend not to be appropriable in the one period of any single leader's tenure, and the U.S. system of public administration provides relatively weak incentives for incumbent leaders to make them.

A central theme of this presentation is that the standards for evaluating competition agencies should press incumbent leaders to invest substantially in activities that improve the capacity of their agencies over the long term. Fred Hilmer, whose report in the 1990s led to the reformulation of Australia's competition system, makes this point when teaching executive education courses for business officials. He tells his students that the good things happening in their companies today probably result from investments their predecessors made five to ten years ago. He tells his students to ask themselves the following: What are you doing today to make sure that your successors will prosper five or ten years hence?[7] That is the norm that the criteria for evaluation should establish and promote inside a competition agency.

---

[6] On the cumulative, evolutionary nature of policy development in the field of competition law, see William E. Kovacic, *The Modern Evolution of U.S. Competition Policy Enforcement Norms*, 71 ANTITRUST L.J. 377 (2003) [hereinafter "Kovacic, *Antitrust Norms*"].

[7] I am grateful to Professor Hilmer for sharing this point with me.

theory and empirical knowledge.[9] What seems to be wisely conceived policy in one era might be proven to be unwise in a later period. This suggests that competition agencies should be given two grades. The first grade depends upon whether the agency's policies were consistent with the state of

The second weakness with case-driven activity measures is that they ignore non-litigation activities. For example, case-related activity measures overlook the significance of the FTC study,

It is not evident why the FTC's cases involving standard setting and settlements between producers of branded and generic pharmaceutical products would fall into an "except for" category and would not be relevant to an examination of the FTC's nonmerger antitrust enforcement program. These matters have considerable economic significance and raise important issues of competition policy. No sensible scoring principle would fail to treat these enforcement initiatives as noteworthy and important. To exclude

tracted little attention at the time of their commencement, yet these cases yielded doctrinal results that have influenced the future course of competition policy.

There must be standards of significance that are better—much better—than media attention. Two suitable criteria would be economic impact and doctrinal significance. How did the case improve the economic well-being of consumers? Did the case address issues important to the development of antitrust law?

We can examine the FTC's program during the Bush administration in light of both criteria. How do the Commission's abuse of dominance cases since 2001 compare to earlier matters in terms of economic effects?[36] The settlement achieved by the FTC in 2005 in *Unocal* has been worth approximately $500 million per year to consumers of gasoline in California. That is the most substantial measurable pay-off from an FTC abuse of dominance case since Congress established the agency in 1914. Since the settlement that resolved the matter in 2003, the *Bristol-Myers* case has yielded at least $3 billion to $5 billion in benefits, and the amount is growing. Gauged by observable economic effects, *Unocal* and *Bristol-Myers* belong on the list of the five most important abuse of dominance cases in the Commission's history.

Doctrinal significance is a separate test of the FTC's program since 2001. Three of the FTC's abuse cases (*Negotiated Data Solutions*, *Rambus*, *Unocal*) dealt with the operation of standard setting bodies. Three (*Bristol-Myers*, *Biovail*, *Unocal*) dealt with misuse of government regulatory processes. One (*Valassis*) concerned an invitation to collude, and another dealt with the use of patent settlements as instruments of improper exclusion. In what sense could these be said to be unimportant matters of antitrust policy or doctrine?

The significance of the doctrinal stakes is perhaps most evident in *Rambus*. In *Rambus*, the Commission failed to sustain its finding of liability, and the Supreme Court declined to take certiorari. In historical context, the doctrinal importance of *Rambus* stands out. Before *Rambus*, the FTC's most recent appearance before the appellate courts in an abuse case was *Borden, Inc. v. FTC* (*ReaLemon*)[37] in 1982. None of the agency's abuse of dominance cases in the 1990s was litigated to a decision on the merits. All settled. The Commission's remedy in *Rambus* included compulsory licensing of a patent, and the most recent appellate endorsement of compulsory

the Supreme Court. *Rambus* was not a litigation success for the FTC. But it presented doctrinal issues of the highest order.

The depiction of the FTC's abuse of dominance program since 2001 as either nonexistent (the Goldschmid critique) or substantively insignificant (the Pitofsky critique) is more than cheap talk spoken amid an election campaign or a presidential transition. It has the capacity to set expectations for the FTC leaders who will guide the agency during the Obama administration. If prominent advisors to the Obama campaign belittle the FTC's program during the Bush administration as null in activity or insignificant in substance, what type of program must new management pursue to be seen as sufficiently active and influential? What sorts of matters must agency managers initiate to surpass the economic results of *Unocal* or *Bristol-Myers*, or to exceed the doctrinal content of *Polygram* or *Rambus*? If the

reduce the already small number of industry participants. In the hardest cases, the agency confronts arguments—for example, involving diminished competitive capability, new entry, or technological change— that weigh against reliance on presumptions that ordinarily would shape the disposition of say, a three-to-two merger. What the agencies do, and why they do it, in these problematic matters is an interesting point of comparison over time.

In a recent paper that criticizes federal merger policy (especially DOJ enforcement) during the Bush administration, Jonathan Baker and Carl

ics, McDonnell Douglas's share of current sales overstated its competitive significance. The merger was not a three-to-two combination; it was two-to-two, with Airbus and Boeing remaining as the only credible supply sources in the eyes of the major customers—the commercial airlines and aircraft leasing companies.

One can spin out an intriguing, plausible counterfactual. If the FTC had said no and put the back of McDonnell Douglas to the wall, might the company have found a way to survive and ultimately prosper in the commercial aircraft segment of the aerospace market? Perhaps the company could have repositioned itself to succeed in what would become the burgeoning regional jet market and use this experience base gradually to migrate into larger aircraft designs. Perhaps the Air Force would have delivered on preliminary suggestions that it would support the development of a commercial freighter variant of the C-17 transport—giving McDonnell Douglas an important means to continue design and production operations for the largest types of commercial airframes.

Nor was the FTC's examination of defense-related features of the merger free from doubt. For all of its problems in commercial aircraft production, McDonnell Douglas remained a formidable designer and producer of major weapon systems. In a number of defense-related market segments, the merger with Boeing removed McDonnell Douglas as an independent center of design, developme

The petroleum mergers of the 1990s provide a second useful view of risk-taking during the Clinton administration. In 2004, the Government Accountability Office ("GAO") issued a report that studied seven petroleum sector mergers reviewed by the FTC from 1995

Exxon-Mobil, Shell-Texaco, BP-Amoco, and BP-Amoco/Arco, albeit with substantial divestitures.

Recent debates about merger enforcement policy ordinarily do not perform this type of comparison of cases across time. Commentators typically provide case studies in isolation without considering whether the agency in one period took risks that were outside a zone of acceptable practice estab-

ade dwarfs, in number of cases, the Antitrust Division's program in the 1950s, 1960s, and 1970s. The combined DOJ/FTC output of abuse of dominance cases in this decade exceeds the output of the Reagan Administration. The total DOJ/FTC output of civil horizontal restraints matters in this decade surpasses the government's output of such cases in the 1950s, 1960s, and 1970s. Immunity and exemption cases filed in this decade match or exceed levels from the 1950s, 1960s, and 1990s. A more detailed examination would identify other examples of how a cases-count-for-everything measure places the Bush administration higher on the ladder of activity than its predecessors in key areas of activity.

More fundamentally, there are good reasons to reject the Obama statement's suggestion that case counts are the best way to measure the strength or weakness of an antitrust program. One of the most important reasons to distrust measures of effectiveness that rest solely on case counts is that views of what constitutes good policy change over time. Owing to its grounding in industrial organization economics, competition law is inherently evolutionary. That is why a properly designed report card marks performance with two grades. One grade measures the agency's work by contemporary standards. The second grade assesses the agency's contribution, in any one period, of policies, doctrinal developments, or analytical concepts that prove to be durable and respected over the long term. The second grade inevitably typically gets filled in only after extensive experience with a contribution provided during a specific period.

Because the development of a competition system is cumulative and evolutionary, "good policy" in any one period may consist of taking different measures in light of trends in the state of current knowledge concerning theory and empirical study. The design of law enforcement programs illustrates the point. Good policy sometimes consists of backing away from existing enforcement frontiers, sometimes pushing enforcement outward, and sometimes sustaining the status quo. A characteristic of good practice is

This program usefully can undergo adjustments, refinements, and improvements. The same can be said of any program in any period. It does not