VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

Edited by Richard Schmalensee

Volume 2 | Number 2 | Autumn 2006

Competition Policy International

From the Editor

Richard Schmalensee

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

From the Editor

Whose welfare should competition policy protect? That is the subject of the first two articles in our Autumn 2006 edition. Is it society at large, including businesses whose profits, after all, ultimately inure to people? Or is it just those people who consume products? The fact that we are even having a debate over whether consumer or total (consumer plus producer) welfare is the right standard for competition policy is remarkable. The U.S. consensus that the antitrust laws should be about competition, not redistribution or protection of small business, is only about four decades old. And only in the last few years did the European Commission start focusing on consumer welfare as its guiding principle. Professors Michael Katz and Joseph Farrell, and Dr. Ken Heyer consider the debate over using consumer versus total welfare as the guiding principle for merger analysis. Both papers generally favor total welfare as the right ultimate objective. However, Katz and Farrell find some of the arguments for having agencies focus on consumer welfare more persuasive than Heyer, who favors a strict focus on total welfare. Moreover, since economists do not generally set policy, this debate is not over.

The debate about the goal of competition policy is followed by a six-paper symposium on state efforts to assist competitors. The EC Treaty prohibits Member States from granting aid to competitors that, roughly speaking, would distort competition in the European Community. This principle has resulted in significant political tension between the Commission—through the Directorate General for Competition (DG COMP), which enforces this aspect of EC law—and Member States. "State aid" was the subject of the May 2006 Antitrust Forum sponsored by the Jevons Institute for Competition Law and Economics at University College London. We are pleased to have papers based on the remarks made by Philip Lowe, the Director General of DG COMP, along with comments on his remarks by Professors Mathias Dewatripont and Frédéric Jenny. Alex Nourry and Nelson Jung then examine what they consider to be the new wave of protectionism sweeping the European Union, as well as the Commission's efforts to address this issue.

While the United States does not have an equivalent prohibition, there have been attempts to ban certain forms of state assistance on the grounds that

Richard Schmalensee is Professor of Economics and Management at the Massachusetts Institute of Technology and John C Head III Dean of MIT Sloan School of Management.

it violates the Commerce Clause of the U.S. Constitution. Professor Peter Enrich, our next contributor, has led this effort and recently argued a case along these lines before the U.S. Supreme Court. (The court rejected the case on procedural grounds without reaching the merits). Maureen Ohlhausen looks at a different, though very important, aspect of state interventions in the competitive process. She focuses on U.S. state legislation that restricts competition, such as laws that limit the interstate shipment of wine (a burden on those of us in Massachusetts who would like to buy wine from California wineries).

The U.S. Supreme Court's decision in *Illinois Tool Works v. Independent Ink* is our featured case for this issue.[1] Reversing long-standing precedent, the Court concluded, after a cogent analysis of tying jurisprudence, that, for the purposes of a tying case, one cannot just presume that a patent confers market power. But, as Richard Taranto argues that the Court also made substantive and methodological contributes to antitrust to which litigators should pay particular attention.

We introduce a new feature in this issue that we will repeat whenever there are worthy subjects: reviews of books on antitrust law. Professor Randal Picker reviews Herbert Hovenkamp's *The Antitrust Enterprise*,[2] and Professor Richard Whish examines Robert O'Donoghue and Jorge Padilla's *The Law and Economics of Article 82 EC*.[3]

We end where we started: the purpose of the antitrust laws. Our classic for this edition is drawn from the writings of Walter Eucken, which laid the foundation for the Ordoliberal, or Freiburg, School of competition policy in Europe. Christian Ahlborn and Carsten Grave have translated the work and written an introductory essay that examines this school of thought, which has profoundly influenced EC competition law, from the standpoint of consumer welfare.

On behalf of the journal's readers and its editorial team, I am delighted to extend my thanks to all the contributors to this issue.

1 Illinois Tool Works, Inc. v. Independent Ink, Inc., 126 S. Ct. 1281 (2006).

2 Herbert Hovenkamp, The Antitrust Enterprise: Principle and Execution (2006).

3 Robert O'Donoghue and A. Jorge Padilla, The Law and Economics of Article 82 EC (2006).

Volume 2 | Number 2 | Autumn 2006

Competition Policy International

The Economics of Welfare Standards in Antitrust

Joseph Farrell and Michael L. Katz

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

The Economics of Welfare Standards in Antitrust

Joseph Farrell and Michael L. Katz

There has been considerable debate concerning whether consumer surplus or total surplus should be the welfare standard for antitrust. This debate misses two critical issues. First, antitrust is not straightforwardly welfarist—it does not maximize but protects, and it does not forbid all actions that seem likely to lower some welfare measure. Rather, antitrust enforcement has both process and consequence components: anticompetitive actions that harm consumers are illegal but other actions that harm consumers are not. Second, the enforcement process involves multiple steps and multiple decision makers. Mergers, for instance, are proposed by the merging parties, reviewed and perhaps challenged by antitrust agencies, and reviewed by courts. Hence, a full discussion of what standard is or should be applied must specify by whom and how it fits in the overall process. We conclude that, while some popular arguments for a consumer surplus standard are weak, other arguments have some merit.

The authors are professors at the Department of Economics and Haas School of Business, respectively, at the University of California, Berkeley. The authors thank Dennis Carlton, Charles Clarke, Ken Heyer, Bruce Lyons, and Carl Shapiro for helpful comments and suggestions.

I. Introduction

What standard should antitrust analysis use to evaluate alternative outcomes? Economists often state that total surplus—the sum of producer surplus and consumer surplus—is the most sensible objective and that consumer surplus is used only because lawyers so interpret the relevant statutes.[1] In this paper, we conclude that the economic issues are much more subtle and less resolved than is generally understood. In large part, these issues are poorly understood because most of the debate has addressed the wrong question. Asking what welfare standard should be applied in antitrust enforcement conflates two separate questions. First, what should be antitrust policy's ultimate goal? Second, what objectives should specific agents (notably the antitrust agencies and the courts) within the antitrust enforcement system apply in their enforcement decisions?

We will argue that total surplus is an appropriate ultimate goal for antitrust enforcement, but that the case for basing enforcement decisions on analysis of total surplus is much less clear. We believe that total surplus is an appropriate ultimate objective because, as others have argued, there is a natural division of labor between efficiency-oriented policies and policies aimed at improving the distribution of income, and antitrust policy fits much better into the first category. Thus, we conclude that a sensible final goal of antitrust policy is to maximize total surplus without regard to distributional considerations.

It does not follow that antitrust agencies or courts should adopt a decision rule of the form: challenge or block behavior if and only if that behavior looks likely to lower total surplus. The antitrust enforcement process involves multiple steps and multiple decision makers. Mergers, for instance, are proposed by the merging parties, reviewed and perhaps challenged by antitrust agencies, and reviewed by courts. Hence, a full discussion of what standard is or should be applied must specify by whom and how it fits in the overall process. For several reasons, which we discuss below, it may be optimal to have specific agents within the broader system act to maximize a different objective (e.g., consumer surplus) even when the ultimate goal of antitrust policy is to maximize total surplus.

1 Consumer surplus is, in turn, defined as the difference between what a consumer is willing to pay for a good or service and what he or she actually pays. Producer surplus is defined as the amount of income a producer receives in excess of what it would require in order to supply a given number of units of a good or service. Intuitively, producer surplus can be thought of as economic profits. Another way of thinking about total surplus is that it is consumption benefits measured in dollars minus the costs of production.

For a discussion of some technical issues concerning the use of the measures in the presence of income effects and multiple commodities, *see* Andreu Mas-Colell, Michael D. Whinston, & Jerry R. Green, Microeconomic Theory §10.C (1995) and Jean Tirole, The Theory of Industrial Organization 7-12 (1993). *See also* R.D. Willig, *Consumer Surplus without Apology*, 66 Am. Econ. Rev. 589 (1976) and Jerry A. Hausman, *Exact Consumer Surplus and Deadweight Loss*, 71 Am. Econ. Rev. 662 (1981).

II. What Did Congress Intend and What Do Enforcers Do?

We begin by summarizing the debate regarding whether the total surplus effects or consumer surplus effects are the basis for determining the legality of firm conduct under U.S. antitrust policy. We argue that the standard currently applied is neither, because whether antitrust law allows particular conduct depends not just on the consequences of that conduct but also on characteristics of the conduct itself.

The major antitrust statutes are remarkably brief and vague, spawning widespread disagreement regarding antitrust goals and standards. Although one might imagine a wide variety of goals, almost all the debate features two or three contending criteria: consumer surplus, total surplus, and (unfashionably) the welfare of competitors.[2] These goals all are welfarist objectives in that each is a function only of economic agents' utility levels, not of the process by which those utilities are obtained or of other aspects of the outcome (e.g., whether consumers' behavior was legal or whether they consume cigars or tofu).

Robert Bork argued that the U.S. Congress intended a total surplus standard, which he confusingly called a "consumer welfare" standard.[3] Others, including Robert Lande, have argued that the U.S. Congress intended a true consumer welfare standard under which the Sherman Act would facilitate wealth transfers from producers to consumers.[4] Steven Salop argues that the current standard is a consumer surplus standard, basing his argument, in part, on the claim that efficiencies play little role in the actual practice of merger policy.[5]

---

2 We focus on the Sherman Act and Section 7 of the Clayton Act. There appears to be consensus that the Robinson-Patman Act sought to protect competitors in a way that today is widely discredited. Arguably the recent *Volvo* decision seeks to move Robinson-Patman away from that standard. Volvo Trucks North America, Inc. v. Reeder-Simco GMC, Inc., No. 04-905 (U.S. Jan. 10, 2006).

3 *See, e.g.*, Robert H. Bork, *Legislative Intent and the Policy of the Sherman Act,* 9 J.L. & Econ. 7, 7-48 (1966).

4 Robert H. Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust: The Efficiency Interpretation Challenged,* 34 Hastings L.J. 65, 65-151, (1982). *See also* Robert H. Lande, *Proving the Obvious: The Antitrust Laws Were Passed to Protect Consumers (not Just to Increase Efficiency)*, 50 Hastings L.J. 959, 963-66 (1999).

5 Steven C. Salop, *Question: What is the Real and Proper Antitrust Welfare Standard? Answer: The True Consumer Welfare Standard*, Statement before the Antitrust Modernization Commission §II.A (Nov. 4, 2005), *available at* http://www.amc.gov/public_studies_fr28902/exclus_conduct_pdf/051104_Salop_Mergers.pdf. *But see* William Kolasky & Andrew Dick, *The Merger Guidelines and the Integration of Efficiencies into Antitrust Review of Horizontal Mergers*, 71 Antitrust L.J. 207, 207-251 (2003) (arguing that efficiencies are important in the process).

Although those contributions contrast the consumer surplus and total surplus welfare standards,[6] others argue that they are nearly equivalent in a long-run perspective because short-run profits spur firms to serve consumers' long-run interests.[7] Indeed, as noted above, Bork thought it proper to dub a total surplus standard a "consumer surplus" standard.

This attempt to defuse the debate fails, however, because even if changes in consumer and total surplus approximately coincide in the very long run, antitrust probably cannot—and surely does not—conduct a very long-run analysis to evaluate a specific case. An analysis with a shorter time horizon (in practice, often two years) may well predict that consumer and total surplus will move in opposing directions. For instance, in the Canadian *Propane* case, the court apparently believed that the merger should be approved under a total surplus standard and blocked under a consumer surplus standard.[8]

Christopher Grandy departs from this consumer surplus-total surplus debate in two ways. First, he argues that Congress meant the Sherman Act to protect competitors rather than consumers. Second, he argues that this protection was meant only against acts that could naturally be called anticompetitive.[9]

This second departure is important. Claims that U.S. antitrust policy imposes a consumer or total welfare (or any welfarist) standard omit a crucial element of antitrust: that antitrust policy examines not only consequences (the change in consumer or total welfare), but also the process (the nature of the acts) that generates the consequences. Specifically, while antitrust may prohibit firms from harming consumers and/or efficiency, it does so only to the extent that firms do so through actions that are deemed anticompetitive.

For example, the models of medium-term effects that antitrust economists tend to use predict that entry into an oligopolistic market by an inefficient producer or

6 In a classic paper, Williamson argued that the use of a total surplus standard could make a very big difference in evaluating mergers that give rise to production efficiencies. Oliver E. Williamson, *Economies as an Antitrust Defense: the Welfare Trade-offs,* 58 AM. ECON. REV. 18 (1968), *reprinted in* 1 COMPETITION POL'Y INT'L 217 (2005).

7 *See, e.g.*, Steven C. Salop, *Efficiencies in Dynamic Merger Analysis*, Statement before the U.S. Federal Trade Commission (FTC) Hearings on Global and Innovation-Based Competition (Nov. 2, 1995), *available at* http://www.ftc.gov/opp/global/saloptst.htm, who argues that taking a dynamic perspective reduces the tension between total surplus and consumer surplus standards. For a discussion of why the dynamic perspective does not fully eliminate the tension, *see* Salop, *supra* note 5, §III.A.1.

8 Commissioner of Competition v. Superior Propane, Inc. (2003) 3 F.C. 529. *See* Thomas Ross & Ralph Winter, *The Efficiency Defense in Merger Law: Economic Foundations and Recent Canadian Developments*, 72 ANTITRUST L.J. 471, 471-503 (2005). Ross and Winter argue that the court may have misapplied the total surplus standard as a matter of economics: they conclude that the incremental deadweight loss due to the predicted price increase was drastically underestimated by being calculated as if the pre-merger price were at marginal cost.

9 Christopher Grandy, *Original Intent and the Sherman Antitrust Act: A Re-examination of the Consumer-Welfare Hypothesis,* 53 J. ECON. HIST. 359, 359-379 (1993).

in an industry with large economies of scale may well reduce total surplus. But we would be surprised if any court ruled that stand-alone entry harmed competition.[10] Similarly, a claim of excessive competition is unlikely to be a winning defense of price fixing.[11] Evidently, either we don't trust those models, or we don't believe in a purely welfarist total surplus standard.

One response is that entry plainly increases competition in a layman's sense of the term. To a lawyer, that might be the end of the story—an end that proves that antitrust is not purely welfarist. A sympathetic economist might be more apt to say that the models are informative but not conclusive concerning the effect of entry on surplus, and must be weighed against the well-established view that competition generally promotes efficiency. In other words, even in relatively simple problems such as stand-alone entry into an oligopolistic industry, our specific analyses inevitably omit much, and their conclusions must be taken with a certain amount of judgment. Here, a not unreasonable judgment might be that entry typically promotes total welfare in the long run more than the models capture.[12] This sophisticated view is compatible with subtle versions of the welfarist position that antitrust seeks to promote total surplus in the end, but it is incompatible with the strong form of the welfarist position that antitrust enforcement decisions should be based on an industry-specific, fact-intensive, detailed prediction of the effects that the conduct under examination has on total surplus.[13]

Claims that U.S. antitrust policy imposes a consumer or total welfare (or any welfarist) standard omit a crucial element of antitrust: that antitrust policy examines not only consequences (the change in consumer or total welfare), but also the process (the nature of the acts) that generates the consequences.

10 That said, the antitrust treatment of exclusive dealing does allow for the possibility that monopoly is preferable to competition in some circumstances within a vertical relationship.

11 Indeed, in *Professional Engineers*, the Supreme Court stated that "the Rule of Reason does not support a defense based on the assumption that competition itself is unreasonable." *See* National Society of Professional Engineers v. United States, 435 U.S. 679, 696 (1978).

12 Elsewhere, one of us has applied somewhat analogous reasoning to the economics of payment-card fee structures and interchange. Joseph Farrell, *Efficiency and Competition Among Payment Instruments*, 5 Rev. Network Econ. 26, 26-44 (2006) (suggesting that competition policy might wisely promote privately optimal—rather than socially optimal as estimated in models—choice by customers).

13 One can argue that merger policy also reflects a view that enforcement decisions should not be based solely on detailed, case-specific predictions of welfare effects. Specifically, horizontal mergers are typically allowed if it can be shown that there would be small competitive effects—without any formal assessment of efficiencies. This process reflects the existence of what has been called a "standard deduction" for merger efficiencies. *See* Michael Salinger, Director of the FTC Bureau of Economics, Four Questions About Horizontal Merger Enforcement, Presentation to American Bar Association Antitrust Section Economics Committee Brown Bag (Sept. 14, 2005), *available at* http://www.ftc.gov/speeches/salinger.htm. Kolasky & Dick, *supra* note 5, describe what they term the Chicago School view that agencies and courts are unlikely to be good at evaluating claims of efficiencies, which might imply advantages of a standard deduction over requiring or even allowing firms to itemize.

Alternatively, one might observe that, even if it reduces total surplus, entry into oligopoly (both in theory and practice) generally benefits consumers; thus consensus approval of such entry might reveal that the implicit welfare standard is consumer surplus rather than total surplus.[14] But this argument, too, is weak. Antitrust proudly eschews plenty of opportunities to promote consumer surplus, at least in the short or medium run. In particular, monopoly pricing is not itself illegal in the United States. Indeed, in its recent *Trinko* decision, the U.S. Supreme Court opined that "the mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system," and that "to safeguard the incentive to innovate, the possession of monopoly power will not be found unlawful unless it is accompanied by an element of anticompetitive *conduct*."[15] The Court is apparently reasoning that this rule promotes total or consumer surplus in the long run, so the policy is consistent with having a welfarist standard in the top-level sense. But the rule is not one that would emerge if agencies were to pursue total surplus or consumer surplus as estimated by available facts and economic models in particular cases.

Thus, in antitrust as it is practiced, both consequences and process count: it never answers only the question "does this practice reduce some measure of surplus?" It is incomplete and potentially misleading to say that antitrust protects consumer surplus, total surplus, or rivals' profits. Rather, conduct can violate the antitrust laws only if it is held to harm competition. As many have noted, the concept of harming competition is often hard to interpret, and too naïve an interpretation would prohibit many beneficial agreements. Thus, the law has evolved toward prohibiting only acts that both (a) hurt competition in an ordinary (if sometimes vague) sense and (b) hurt efficiency and/or consumer surplus. The debate over the so-called "standard" is the debate over the standard applied in prong (b). We think that the debate is clarified by keeping this two-pronged criterion explicit, and not seeking to have the second prong redefine the word "competition" or claiming that antitrust is straightforwardly welfarist.[16]

14 Steven Salop, *supra* note 5 at 10 & 11, appears to hold this view. Although we could imagine someone coming to a general judgment that total surplus is in the long run best promoted by putting zero weight on the profits of disappointed competitors while otherwise relying on an antitrust-style medium-run analysis, it is certainly not what would naturally be meant by "applying a total-surplus standard."

15 Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398, 407 (2004)(emphasis in original); *see also* R. Hewitt Pate, Assistant Attorney General, Competition and Intellectual Property in the U.S., Speech to EU Competition Workshop (June 3, 2005) *available at* http://www.usdoj.gov/atr/public/speeches/209359.pdf ("if a monopoly is lawfully obtained…we do not even object to setting a monopoly price.")

16 *See* Joseph Farrell, *Complexity, Diversity and Antitrust*, 51 Antitrust Bulletin 156 (Spring 2006) (arguing that it is a mistake to try to collapse these two components into a redefinition of the word competition to (almost) mean a surplus standard).

Some economists promote welfarism as a sine qua non of reasoned policy analysis. For instance, Ross & Winter, *supra* note 8 at 474, write that "welfarism…may appear so obvious that it must be

Merger policy—the bulk of agency antitrust practice—might appear to contradict our claim that there are two prongs, because merger enforcement focuses purely on consequences (i.e., competitive effects and efficiencies). But we would argue that this focus is consistent with our interpretation because almost all horizontal mergers satisfy prong (a); that is, they reduce competition in a natural sense. Hence, merger analysis can focus on whether a transaction satisfies prong (b). And even here, the process is not truly welfarist. In particular, with minor exceptions, even merger policy does not seek to maximize a welfare measure, but only tries to ensure that such a measure does not fall as a result of a merger.[17]

Having established that there are two prongs to the analysis, we will spend the rest of this essay considering the relative merits of consumer surplus and total surplus as welfare standards in prong (b).

III. Do Distributional Concerns Justify Use of a Consumer Surplus Standard?

Perhaps the leading philosophical claim made in favor of a consumer surplus standard is that it better reflects society's judgments about the appropriate distribution of economic welfare than does a total surplus standard. The use of total surplus implicitly assumes that the distribution of income is socially optimal, so that taking a dollar away from one member of society and giving it to another member would not affect social welfare. As one textbook put it,

footnote 16 cont'd

satisfied in any serious discussion of merger analysis. [...The view] that competition policy is about \`protecting competition' [...] is without economic foundation." Although we agree that total surplus is the appropriate ultimate objective of antitrust enforcement, this view does not imply that day-to-day antitrust enforcement should be based on seeking in the instance to evaluate a welfarist measure. Because it is impossible to predict long-run effects with certainty, it could easily be consistent with a long-run welfarist view to adhere to well-chosen non-welfarist principles (e.g., protect competition). Kolasky & Dick, *supra* note 5 at 207, write that "it is efficiency, not competition, that is the ultimate goal of antitrust...\`efficiency is the goal, competition is the process'."

17 With fairly limited exceptions, antitrust does not ask whether an alternative merger would yield higher welfare: the U.S. Dep't of Justice & Federal Trade Commission, Horizontal Merger Guidelines §5.1 (1992) (as amended Apr. 8, 1997) *reprinted in* 4 Trade Reg. Rep. (CCH) ¶13,104 (hereinafter "Merger Guidelines") contain provisions for checking whether a failing firm is being sold to the best acquirer from a welfare perspective, and in evaluating efficiencies, the Merger Guidelines §4 discuss examining whether an alternative deal would achieve the efficiencies without the adverse competitive effects. However, it generally is not the case that a merger can be successfully challenged on the grounds that a different merger would yield higher consumer or total surplus. Similarly, when an exclusive dealing contract is challenged, there is not a full-blown investigation to determine the best possible vertical contract from the perspective of social welfare.

"Implicit in our use of total surplus is the claim that society is best off when the total surplus is maximized. But you might be worried that there is some kind of value judgment behind that claim. If you are, you are correct; there is. The value judgment is that a dollar to each person is given the same weight, whether that person is a consumer or a producer, rich or poor."[18]

It is, however, a widely held view that an additional dollar is worth more to society in the hands of a poor person than those of a rich one. This view underlies the support for a variety of redistributive policies, including progressive income taxation and the provision of government-subsidized health insurance for low-income families.

There are at least three rationales for antitrust enforcement's use of total surplus as a measure of social welfare even in the presence of such distributional concerns. The first is to view the use of total surplus as a response to uncertainty about distributional effects. For instance, the 1991 Canadian *Merger Enforcement Guidelines* stated: "[w]hen a dollar is transferred from a buyer to a seller, it cannot be determined *a priori* who is more deserving, or in whose hands it has a greater value."[19] If enforcers do not, or cannot, undertake a case-by-case determination of relative deservingness, then it may be best simply to assume that all affected parties are equally deserving.

A second rationale is the following. If outcome A yields greater total surplus than outcome B, then in principle it is possible to design a system of wealth transfers, starting from A, such that at least one person ends better off than in B and no one is worse off. This idea is known as the Kaldor-Hicks criterion.[20] If the redistribution is actually done, then A is preferred by all parties to B. In practice, however, such compensation often is impossible given the limited available information about effects of A versus B on individual consumers and producers. And, when compensation is not paid, some parties typically will prefer B to A. Hence, the use of a total surplus standard imposes particular value judgments. The rationale in this case can loosely be stated as adopting a principle of maximizing total surplus and then counting on the process to balance out gains and

18 Michael L. Katz & Harvey S. Rosen. Microeconomics 362 (3rd ed. 1998).

19 Part 5.5, footnote 57, *available at* http://www.competitionbureau.gc.ca/internet/index.cfm?itemID=1682&lg=e , with the notation that (in view of Superior Propane, cited supra) "This Part no longer applies. Readers should consult the decision of the Federal Court of Appeal in the Commissioner of Competition v. Superior Propane Inc. and ICG Propane Inc 2001 FCA 104."

20 *See* J. de V. Graaf, Theoretical Welfare Economics 82-90 (1963). There are also technical conditions regarding the size of income effects that may come into play.

losses over time to ensure a fair distribution of economic benefits. This rationale builds on the uncertainty idea above. Suppose that enforcement decisions are always made to maximize total surplus. Then, on average, everyone will be as well off as possible. Nonetheless, any particular individual or firm may be better or worse off than if different decisions were made.

The lack of a guarantee leads to the third rationale for use of total surplus as antitrust's measure of social welfare even in the presence of distributional concerns. This rationale is a division of labor among public policies: if antitrust enforcement and some other public policies focus on total surplus, other public policies can redistribute that surplus in accord with notions of fairness.[21] A number of reasons suggest that antitrust policy is poorly suited as a redistribution vehicle in comparison with various tax and subsidy schemes.[22] Its principal shortcoming is that antitrust enforcement does not, and—without a fundamental change in the nature of analysis—cannot, take a comprehensive view of distribution. It would become necessary to examine the relative income distributions among consumers, workers, and firm owners. In many instances, data would be lacking.

To illustrate this shortcoming, consider how a consumer surplus standard handles distributional issues. Consumer surplus can provide a very a poor approximation to a welfare measure that weights impacts using ordinary notions of distributional preferences. One reason is that rich and poor consumers may be differentially affected by an antitrust decision; distributional concerns would suggest weighting the impact on the poor more heavily, but a consumer surplus standard insists that they count equally. If a central goal of antitrust enforcement is to redistribute income, then why treat rich and poor consumers alike? Another problem with using consumer surplus to embody a preference for wealth redistribution from rich to poor is that the owners and workers of firms are people too. Use of a consumer surplus standard entails treating all consumers as equally deserving at the margin, yet treating the same people unequally in their roles as workers and capital owners. The merger of makers of expensive fountain pens illustrates how a consumer surplus standard can go wrong in this regard.[23] Lastly, when the market is not a final-goods market, the direct buyers are themselves firms, so a natu-

21 Steven Salop, *supra* note 5 at 17, recently argued against this rationale on the grounds that the tax authorities do not compensate members of society for the wealth transfers induced by specific mergers or anticompetitive firm conduct on a case-by-case basis and that, if authorities did so, the transactions costs would be enormous. This criticism misses the point that, in the face of transactions costs, it is desirable to implement policies that work well on average (rather than exactly case by case) even when one has strong distributional preferences. Instead, taxes and various social subsidy programs are intended to equalize the marginal social value of income across consumers, subject to informational constraints and the need to take transactions costs into account.

22 Not all commentators might agree. Lande, *supra* note 4, for instance, has argued that Congress intended antitrust largely as a strategy for wealth redistribution.

23 United States v. Gillette Co. & Parker Pen Holdings, Ltd., 1993-1 Trade Cas. (CCH) ¶ 70,210 (D.D.C. May 5, 1993).

ral interpretation of a consumer surplus standard favors buying firms over selling firms.[24] We are aware of no evidence that the wealth distribution of shareholders varies systematically according to a firm's place in the value chain.

A different argument for use of a consumer surplus standard is based on imperfections in corporate governance. There is evidence that part of free cash flow coming into widely held corporations is dissipated by managers serving their own interests rather the owners'. Although such expenditures promote managers' welfare, they are likely to be inefficient because managers are constrained in how they can spend these funds without running afoul of corporate governance. This observation might justify underweighting increases in profits (before dissipation) relative to changes in consumer surplus. But as the quotation from *Trinko* cited earlier in this paper notes, profits also can induce efficient investment, so this argument does not provide strong support for use of a consumer surplus standard.

In summary, we believe that antitrust is not a good policy tool for redistributing income, and even if it were, we doubt that distributional concerns provide a sound basis for preferring a consumer surplus standard over a total surplus standard.

IV. A System-Level Perspective: Decision Rules versus Objectives

Most antitrust economics literature assumes policy optimization by a single decision maker. In fact, antitrust enforcement involves multiple layers of decision makers. In a multi-layered decision process, one should not presume that each participant is or should be tasked with maximizing the overall objective.

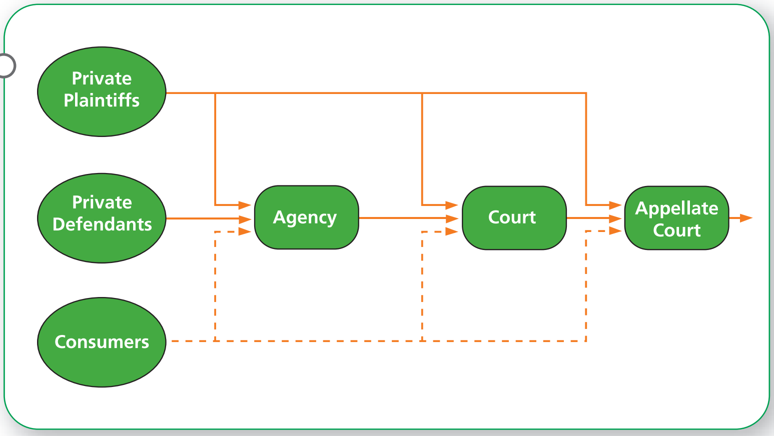
Two important examples illustrate this general point. First, in the U.S. advocacy legal system, although parties' lawyers are officers of the court, legal ethics charges them primarily to be their clients' advocates, even though the final goal is justice. Second, suppliers in competitive markets pursue profits, yet act to maximize total surplus; as Adam Smith noted, "it is not from the baker's benevolence that we expect our bread." As these examples make clear, commentators should not simply jump from a belief that welfare measure *W* is the appropriate final goal to a belief that the agencies ought to base their challenges to firm conduct on estimates of that conduct's effects on *W*. The rules that particular decision makers within the overall system should use could well differ from the ultimate goal of antitrust policy.

24 Here we assume that the analysis focuses on the immediate impact on direct buyers. When direct buyers are not final consumers, subtle economic issues arise regarding pass-through. For a discussion of how this was debated in the context of the proposed merger between Heinz and BeechNut, *see, e.g.,* Jonathan Baker, *Heinz Proposes to Acquire BeechNut*, *in* The Antitrust Revolution (John Kwoka & Lawrence White eds., 4th ed. 2004).

Both internal and external considerations might affect what standard the agencies and agency staff should use. Internal considerations concern the motivation and compensation of agency staff and management. External issues include: accounting for self-selection by the firms triggering investigations (e.g., choosing to merge in the light of their predictions of how the proposed transaction will be treated); the generally weak participation in the process by final consumers; and the passive or reactive role of the courts as adjudicating agency challenges but not themselves initiating challenges. Figure 1 provides one schematic and simplified overview of the process. The lines indicate points at which various parties may first enter the system. We use dashed lines to represent the fact that—because they are typically numerous, unorganized, and have small individual stakes in the outcome—consumers often play a more limited role in the proceedings than do other parties. Among the diagram's simplifications, it does not illustrate the various components of decision making within the agency, and there may be one or more additional rounds of appeal.

Figure 1

An agency or court is one component of a complex system



Internal issues would arise even if the agency were a dictator; external considerations arise because it is not. That said, the issues overlap. Just as the overall enforcement system comprises several decision makers playing different roles, so does a single agency's decision-making structure. For instance, staff members typically investigate firm conduct and then make recommendations to management personnel who serve as gatekeepers. We proceed by very briefly discussing internal considerations and then discussing several external considerations in turn. We examine a series of models, which establish that, even when the overall objective of antitrust policy is to maximize total surplus, it may be optimal to instruct specific antitrust enforcers to pursue decision rules based on alternative welfare measures.

A. INTERNAL CONSIDERATIONS

Antitrust agencies are not unitary decision makers. Instead, they are organizations (with important elements of hierarchy) in which many different people participate in decisions. To illustrate some of the issues, consider a stylized simple agency with only two members. The staff member collects information, analyzes it, and makes a recommendation to the agency manager. In our experience, such recommendations recite some, but not all, of the underlying information collected. Based on the information forwarded to him or her, the manager decides whether to proceed to litigation. If there is litigation, the staff member argues the case in court.

The decisions made by the staff member and manager will depend on their personal preferences and the nature of their compensation. Presumably success or failure in litigation will affect the staff member's compensation, at least in the long run. Thus, a rational staff member will take into account the probability of winning the case when making a recommendation to the manager. There may also be selection issues: economists and lawyers who choose to work in the government are unlikely to be a random sample of all economists and lawyers. This self-selection can matter because antitrust enforcement clearly entails important elements of judgment. In the presence of these internal considerations, it is not self-evident what the optimal standards to ask the staff and management to pursue are.

B. SELECTION BY THE PARTIES

Antitrust enforcement arises in response to actions taken by firms. If two firms do not wish to merge, antitrust never requires them to do so. Similarly, if a manufacturer enters into exclusive contracts with its distributors, the agencies may investigate and challenge that practice, but they do not proactively force the firm to adopt a specific contractual regime. The fact that antitrust enforcement is reactive gives firms important influence over antitrust outcomes.

Two recent papers on merger analysis investigate implications of the fact that firms choose which mergers are proposed and, thus, receive antitrust scrutiny. These models treat the antitrust enforcement agency and the court as a single entity. Firms predict that entity's enforcement behavior, and that prediction affects what mergers are proposed. The private parties' choices of which mergers to propose and the enforcement entity's choices of which proposed mergers to allow interact to determine which mergers are consummated. In each of these models, total surplus may be better served if antitrust agencies protect consumer surplus than if they protect total surplus.

Let M, R, and S denote the merger-induced changes in the profits of the merging parties, the profits of other suppliers, and level of consumer surplus, respectively. Although we use the mnemonic R for "rival," the other suppliers could also be suppliers of complementary goods and services. The associated change in total surplus is $W = M + R + S$.

Bruce Lyons argues that antitrust enforcement should account for self-selection by firms. Specifically, firms choose the most profitable of permissible mergers, knowing that some profitable mergers would be blocked by antitrust enforcement.[25] Figure 2a illustrates the logic of this model. For simplicity, assume that $R \equiv 0$ for all possible mergers. The four black dots in figure 2a represent mutually exclusive possible mergers or merger strategies (for the moment ignoring antitrust constraints), where each merger is characterized by its effects on the merging parties' profits, M, and consumer welfare, S. Given the assumption that $R = 0$, a merger's effect on total surplus is equal to $M + S$. Because rational firms will never propose mergers for which $M < 0$, the figure displays only profitable mergers. If antitrust enforces a consumer surplus standard, only mergers in green shaded region I would be allowed; if antitrust enforces a total surplus standard, then only mergers falling in either the green shaded regions I or orange shaded region II would be allowed.

Assume that all involved can perfectly predict the profit and consumer welfare consequences of any proposed merger, which also implies that the parties can perfectly predict which mergers would be allowed under any given antitrust standard. Profit-maximizing firms will choose the merger with the highest value of M (the most profitable merger) that will not be blocked by antitrust.[26] Hence, in figure 2a, under a consumer surplus standard the firms would propose merger *a*, while under a total surplus standard they would propose merger *b*, which is more profitable but harms consumers. All points on the line with a slope of -1 running through point *a* involve the same total surplus. As shown, a consumer surplus standard induces a higher level of total surplus than does a total surplus standard!

This logic illustrates that the standard adopted by antitrust enforcers is not the full story about what happens: even if in the end we want to maximize total surplus, in some circumstances antitrust authorities should challenge a different set of mergers than the set of all mergers that lower total surplus.

Examples, however, can tell us little about whether such circumstances hold in practice, or about whether the allowed set should be related to consumer surplus specifically. Figure 2b illustrates the case in which a consumer surplus standard would induce merger *c*, yielding lower total surplus than merger *d*, which would be induced by a total surplus standard. Which case, figure 2a or figure 2b, is more likely? Lyons considers conditions under which figure 2a is more likely than figure 2b, but at this stage we view that discussion as exploratory. Several factors affect which standard is preferable.

25 Bruce R. Lyons, Could Politicians be more Right than Economists? A Theory of Merger Standards (Univ. of East Anglia Centre for Competition & Regulation Working Paper CCR 02-1, revised May 2002). For a similar analysis, *see* Sven-Olof Fridolfsson, Essays on Endogenous Merger Theory: A Consumer Surplus Defense in Merger Control (2001) (Ph.D. dissertation, Stockholm University).

26 A number of issues arise regarding the order in which various firms choose to propose what would be incompatible mergers. For a fully specified model that addresses these issues, *see* Lyons, *supra* note 25.

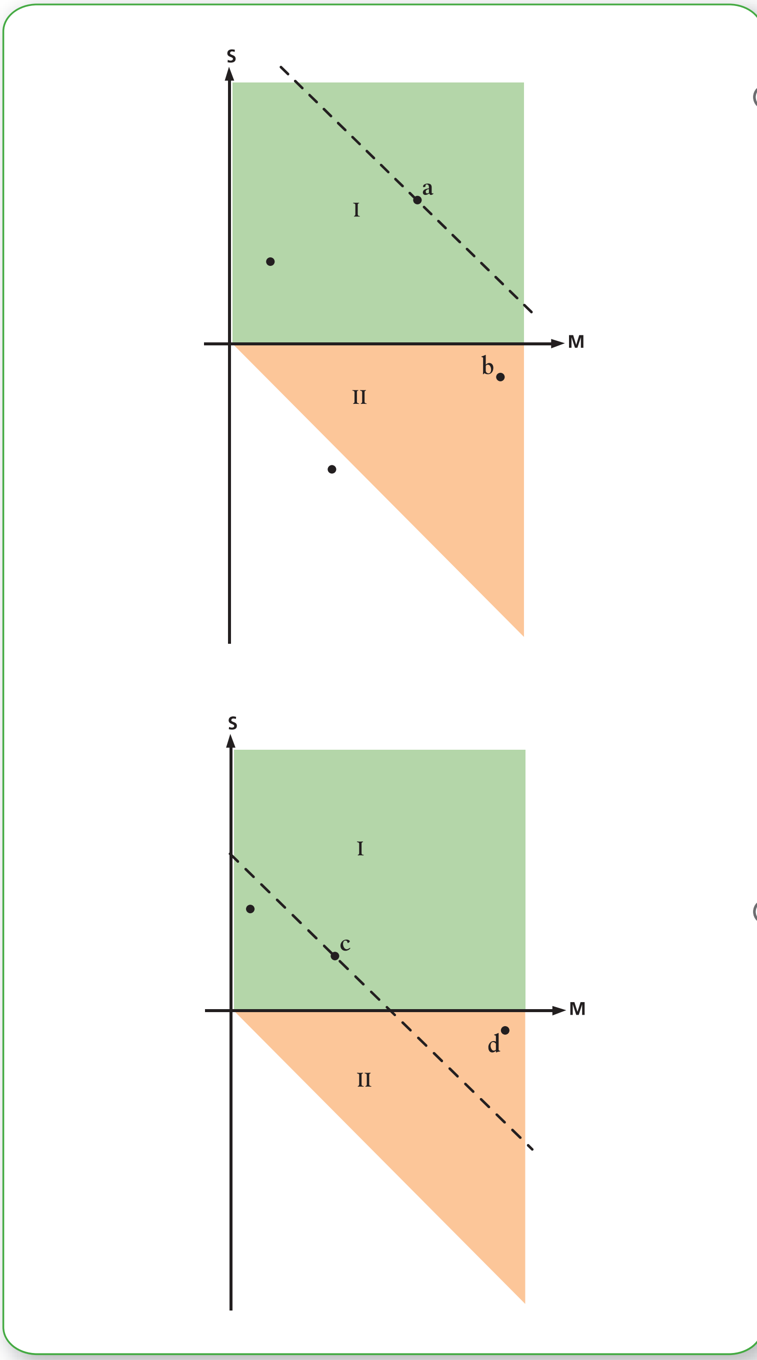


Figure 2a

A consumer surplus standard promotes total surplus

Figure 2b

A consumer surplus standard reduces total surplus

For example, one factor is whether efficiencies are inextricably linked to adverse competitive effects. In the extreme case, suppose there is no link and that all efficiencies can be achieved in a range of different ways, some of which have significant adverse competitive effects and others (e.g., fixed-fee licensing) not. In this setting, a consumer surplus standard would tend to push firms toward achieving desired efficiencies in ways that do not have significant adverse effects. There are, however, at least two reasons to expect efficiencies and competitive effects to be linked. First, if the merging firms do not significantly compete against one another, then they have a joint incentive to cooperate on achieving efficiencies even without a merger. Second, firms that are closer competitors might also have more similar operations and, thus, the potential for greater efficiencies. When efficiencies and competitive effects are linked, there may be a tradeoff between the realization of efficiencies and avoiding adverse competitive effects. The implications of this tradeoff for the choice of welfare standard remain to be explored.

A second factor is the nature of the efficiencies. For example, if all efficiencies take the form of fixed-cost savings and every merger has some adverse competitive effects, then a consumer surplus standard would block all mergers, while a total surplus standard would allow those that increase total surplus. Hence, in this setting, a total surplus standard would give rise to greater total surplus than would a consumer surplus standard.

Even if the model cannot show whether consumer surplus is likely to be the better standard, we take two lessons from it. First, as we have noted, it is important to consider the whole process. Second, Lyons' model suggests the intuition that (a) the outcome reflects both what firms push for and what antitrust pushes for, and (b) if we want to maximize gains in total surplus (northeasterly movement as shown in figure 2) and firms always push eastwards, there is something to be said for someone adding a northerly force.

We also observe that, although the Lyons model is described in terms of merger enforcement, the logic applies to other areas of antitrust as well. Firms often choose among alternative courses of conduct (e.g., the types of contracts they sign with distributors or the aggressiveness of their pricing policies) that affect profits and consumer welfare. Economically rational firms will choose profit-maximizing actions subject to the constraints imposed by antitrust enforcement.

David Besanko and Daniel Spulber offer a different model in which selection by the potentially merging parties affects the optimal welfare screen to apply in approving or blocking mergers.[27] In their model, time-consistency concerns can make it optimal for the legislature to impose something like a consumer surplus standard on the agencies and courts.

27 D. Besanko & D.F. Spulber, *Contested Mergers and Equilibrium Antitrust Policy,* 9 J.L. Econ. & Org. 1 (1993).

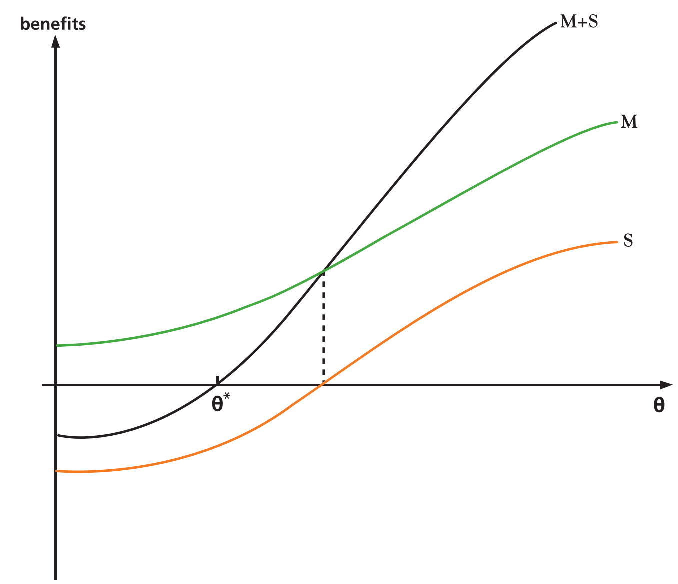


Figure 3

A simple merger selection process

In Besanko and Spulber's model, unlike Lyons', private parties do not choose among mutually exclusive mergers to propose. Instead, private parties consider each member of a set of mergers and choose whether to propose it; there is no linkage across mergers. Thus, the two models, in different ways, simplify the complex reality that if firms A and B merge it may affect whether A/B and C are allowed to do so, and whether D and E are allowed to. For each possible merger in the Besanko and Spulber model, the parties have private information about a parameter, θ, that affects both the change in the merging parties' profits, $M(\theta)$, and the change in consumer surplus, $S(\theta)$. Again for simplicity, assume that $R \equiv 0$. Besanko and Spulber assume that both profits and consumer surplus are increasing in θ. This pattern would arise, for example, if the merger's principal effect were a reduction, measured by θ, in variable costs of production. Figure 3 illustrates this process.

Although the merging parties know the value of θ, antitrust enforcers know only the population distribution and have no other relevant merger-specific information. Thus, enforcers can pick only a single probability, ρ, with which to reject any proposed merger. The model realistically assumes that it is costly to propose a merger that is blocked. These costs include legal and administrative costs, as well as costs that arise from the disruption an enterprise suffers when its future structure is uncertain. Formally, if a proposed merger is rejected, the would-be merging firms are worse off by T than they would have been had they not proposed it. Therefore, given policy ρ, firms will propose a merger if and only if $\rho M(\theta) - (1 - \rho)T \geq 0$, or

$$M(\theta) \geq \frac{1-\rho}{\rho} T.$$

Hence, for any policy ρ, precisely those mergers with large enough values of θ will be proposed.

The optimal policy in this setting is to set ρ at the value that solves

$$M(\theta^*) = \frac{1-\rho}{\rho} T,$$

where θ^* is the value, illustrated in figure 3, such that $M(\theta) + S(\theta) \geq 0$ if and only if $\theta \geq \theta^*$. Label the solution as ρ^*.[28]

At this point, self-selection and time-consistency issues arise. If firms believed that any proposed merger would be blocked with probability ρ^*, then firms would propose exactly those mergers that improve welfare. Indeed, this is how ρ^* was calculated. If the agency recognizes this fact and seeks to maximize total surplus, then it should allow all proposed mergers. But if firms foresee what the agency will do, then they will propose inefficient mergers as well as efficient ones. In short, ρ^* is inconsistent with equilibrium behavior when the enforcer acts to maximize total surplus.

Now suppose that a legislative body directed the enforcer to approve mergers based on a consumer surplus standard. Observe that—because only mergers that yield positive profits to the merging parties are proposed—the level of consumer surplus from a merger is always lower than the level of total surplus. Moreover, mergers with low values of θ harm consumers. If enough mergers that increase total surplus are bad for consumers, enforcers might then reject all mergers conditional on knowing only that $\theta \geq \theta^*$. Although that outcome would generally be neither an equilibrium nor efficient, it does open the way to one that would be. In particular, if the expected value of consumer surplus is negative conditional on $\theta \geq \theta^*$, then there exists a weight, ω, such that the expected value of the weighted sum of total surplus and consumer surplus, $\omega(M + S) + (1 - \omega)S$ is equal to zero conditional on $\theta \geq \theta^*$. Hence, an agency with the objective of maximizing this particular weighted sum of total and consumer surplus will find it optimal to block ρ^* of those merger that it reviews even when it knows that only mergers for which $\theta \geq \theta^*$ are proposed. In other words, the threat to block ρ^* of the proposed mergers will be credible, and the optimal challenge probability is consistent with the private incentives and information of the active participants (in technical terms, this outcome is a perfect Bayesian equilibrium).

28 Simple algebra yields $\rho^* = \frac{T}{T + M(\theta^*)}$.

This model thus confirms the idea that rules for challenging mergers should be evaluated in the context of the system as a whole. We are reluctant to take more than that from the model, however, for several reasons. First, the model assumes that the legislature can commit to a rule but the antitrust agency cannot—even though private parties have frequent observations of agency decisions and, thus, one would expect the agency to form a reputation. Second, the model does not ring true in terms of institutional behavior. The agencies do not view themselves as making merging more costly in order to induce firms to propose only the most profitable ones. Third, if the mechanism of enforcement policy is to raise the cost of merging, then allowing all proposed mergers subject to payment of a well-calibrated tax would be a better policy; the cost to firms would be the same, but the government would collect the revenues rather than simply have economic value dissipated through unproductive activities. Fourth, and finally, the model relies on the strong assumption that, across a set of potential mergers, the most profitable mergers also generate the greatest increase in consumer surplus. This pattern may hold for variable cost reductions, but one would expect the opposite pattern to hold for competitive effects: increased market power would raise profits and—due to deadweight loss—lower consumer surplus by more, thus reducing total surplus. In this setting, any form of a merger tax (including random rejection of merger proposals) would result in mergers less favorable to consumer and total surplus unless it deterred all mergers. Hence, optimal enforcement policy would either block all mergers or allow all mergers, depending on the average effects of a merger on total surplus.

THIS MODEL THUS CONFIRMS THE IDEA THAT RULES FOR CHALLENGING MERGERS SHOULD BE EVALUATED IN THE CONTEXT OF THE SYSTEM AS A WHOLE.

We close by noting that, like the Lyons model, the Besanko and Spulber model and its broad lesson can be applied to antitrust enforcement generally. However, the concerns with the model that we have just expressed also extend to the broader setting.

C. THE AGENCIES AS AGENTS

Another strand of the literature examines the implications of lobbying. Suppose that exposure to the parties tends to tip the agencies toward a relatively sympathetic view of the parties' position. Consumers do not usually engage in lobbying or in other ways participate in the process.[29] Hence, building a pro-consumer bias (relative to a total surplus standard) into the agency's objective function may counteract the bias that can arise from asymmetric lobbying.

29 If consumers are end-users (individuals), this assumption is natural because each consumer has relatively little at stake and may not be well informed. (One might ask whether consumer groups such as Consumers Union help to resolve this problem.) When direct buyers are not final consumers, intervention by direct buyers is more likely, but intervention by final consumers may be even less likely.

Neven and Röller offer a model that makes this point.[30] In their model, the merging parties and other business enterprises affected by the merger engage in lobbying, but consumers do not. The agency is influenced by lobbying but faces a threat of punishment if it fails to apply the statutorily mandated welfare standard. Specifically, the agency chooses the enforcement action (e.g., approving or challenging a merger) that maximizes $I + \alpha(B_M + B_R) \geq 0$, where I is the welfare standard the agency has been instructed to apply and B_M and B_R are lobbying expenditures by the merging parties and rivals, respectively. The fact that monitoring of the agency is imperfect tends to raise α, while limitations on lobbying (what Neven and Röller call transparency) reduce the effectiveness of bribes, tending to lower α.

The lobbying expenditures, B_M and B_R, are equal to the difference between what the relevant firms are willing to spend to have the merger approved in comparison with what they are willing to spend to have it blocked. The merging parties are willing to spend up to M to get the merger approved, while the rivals are willing to pay up to R.[31] Observe that B_R is negative if rivals are harmed by the proposed merger. Given these bidding (bribing) rules, a merger will be approved under a consumer surplus standard if and only if

$$S + \alpha(M + R) \geq 0,$$

and it will be approved under a total surplus standard if and only if

$$S + (1 + \alpha)(M + R) \geq 0.$$

Neven and Röller compare the resulting levels of total surplus when the agency is instructed to apply a consumer surplus standard and a total surplus standard. They find that neither standard dominates the other. Intuitively, instructing the agency to apply a consumer surplus standard compensates for the lack of consumer lobbying. Suppose, for example, that oversight of agency decision-making and private-lobbying activities leads to $\alpha = 1$, so that the agency maximizes the sum of consumer surplus and the bribes. Then, because firms are willing to bid up to the value of the merger, the agency will approve the merger if and only if $S + M + R \geq 0$, which maximizes total surplus. Thus, a consumer surplus standard leads to the first-best outcome in this setting.[32]

In other cases, however, a total surplus standard may yield superior performance. Clearly, if legislators or some other oversight body can perfectly monitor

30 D.J. Neven & L.-H. Röller, *Consumer surplus vs. welfare standard in a political economy model of merger control,* 23 Int'l. J. Ind. Org. 829 (2005).

31 We are presenting a greatly simplified summary of the analysis. *See* Neven & Röller, *id.*, for details of the lobbying game and the equilibrium expenditure levels.

32 One may have to treat $\alpha = 1$ as a limiting case in their model: at several points, the paper presents results that implicitly or explicitly assume $\alpha < 1$.

and control the agency (i.e., if $\alpha = 0$), then a total surplus welfare standard will lead to the first-best outcome, while a consumer surplus standard would reject some efficiency-enhancing mergers for which $S < 0 < S + M + R$. For values of α between 0 and 1, both standards give rise to biases, and for specific parameters either can yield a superior decision.

As in Lyons' analysis, a central issue is whether one can say more than "anything is possible." A sympathetic view is that it is a different form of the same intuition: the merging parties and affected rivals push the outcome in their preferred directions, and if consumers pushed equally hard then the outcome would tend to be efficient. Because consumers seldom do so, the gap can be filled by weighting the agency's objective function more towards consumer interests and less towards those interests that are otherwise well represented in the forces that combine towards the overall outcome. More succinctly, non-consumer interests are represented and, hence, consumer interests should be too.

D. THE ROLE OF THE COURTS

Antitrust agencies do not make decisions in isolation: they are subject to judicial review. For instance, the Antitrust Division of the U.S. Department of Justice cannot block a merger or a business practice, but can only challenge it in federal district court.[33]

This distinction would not matter if courts gave the agency extreme deference, so that the decision to challenge would be tantamount to blocking. Perhaps at one time it was a reasonable approximation in some areas of antitrust that "the only consistency . . . is that the government always wins."[34] If that was ever the case, it clearly no longer is, either in the United States or in the European Community, as demonstrates by the *SunGard*, *Oracle*, *TetraLaval*, and *AirTours* decisions.[35] Thus the distinction does matter for how the agency should decide what to challenge. As a general principle, the agency and the court each should take into account that its decision only matters if the other condemns the prac-

33 Judicial oversight is not complete. Litigation is costly for the parties, and these costs can be a source of agency bargaining power (although the converse also holds). Thus, given the high costs of delay faced by the partners in an unconsummated merger, an agency decision to challenge a merger often leads the parties to abandon the merger rather than defend it in court. There are, however, parties that prevail in litigation and are allowed to merge despite the agency's objections, so judicial oversight is meaningful.

34 United States v. Von's Grocery, 384 US 270, 301 (1966) (Stewart, J., dissenting).

35 United States v. SunGard Data Systems, Inc. and Comdisco, Inc., 172 F. Supp. 2d 172 (D.D.C. 2001); United States v. Oracle Corp., 331 F. Supp. 2d 1098 (N.D. Cal. 2004); Commission v. Tertra Laval BV, 2005 E.C.R. I-01113 (Grand Chamber); Airtours, PLC v. Commission, 2002 E.C.R. II-02585 (Ct. First Instance).

tice.[36] Consider, for instance, the extreme case in which the court always reaches the correct finding, the agency occasionally errs in its own assessment, and there are no litigation costs. In this case, the agency should challenge everything, ignoring its own estimates of the effects on consumer surplus, total surplus, or any other welfare measure—a challenge is costless and leads to the optimal decision. In the other direction, suppose the courts have less information than do the agencies, and the courts trust the agencies to pursue the right objective.[37] In this case, the courts should give the agencies extreme deference and, anticipating deference, the agencies should challenge cases if and only if they believe that the conduct under investigation would lower the relevant social welfare measure—a wholly deferential court plays no screening role. Once again, the decision calculus of an antitrust enforcer should account for that enforcer's role in the overall system.[38]

E. DOES A SYSTEM-LEVEL PERSPECTIVE SUPPORT A CONSUMER SURPLUS STANDARD?

The analyses discussed above show that, even if the overall objective of antitrust policy is to maximize total surplus, it may be optimal for enforcement agencies to use decision rules that apply a different standard. A central shortcoming of these analyses as a basis of policymaking is that each identifies possibilities but offers little guidance as to how often a consumer surplus standard is likely to lead to a higher level of total surplus than would a total surplus standard, or whether some third standard might be best. This is not a criticism of earlier authors; it simply means that much work remains to be done in this area. Clearly, the foundations for a total surplus rule, in the practical sense in which it would be actually used, are a good deal shakier than most economists have understood, but it is not yet time to abandon the edifice.

V. Bargaining and Remedies

The agencies often negotiate settlements with private parties and courts may impose remedies. What objective should agencies and courts pursue in negotiating and designing these conditions? A sensible candidate might be to turn a profitable, yet welfare-reducing, merger into a somewhat less profitable but welfare-

36 For a discussion of multi-party decision making when each party has veto power, *see* Raaj Sah & Joseph Stiglitz, *The Architecture of Economic Systems: Hierarchies and Polyarchies*, 76 AM. ECON. REV. 716, 716-727 (1986).

37 By less information, we mean that the agency's information in any given case is a sufficient statistic for the court's.

38 These issues could be explored further in the hierarchical decision framework we sketched above in our discussion of internal decision-making structure within an agency. Here, the court would be the relatively passive final decision maker, while the agency would play the role of collecting information and proposing a decision.

enhancing or at least welfare-neutral one.[39] Here, too, the interaction of different parties affects the optimal welfare standard.

In the context of merger remedies, Farrell considers a model that also expresses the general idea of countervailing influence.[40] He argues that merger remedies are best modeled not as imposed by the agency but as negotiated between the agency and the parties. Without explicitly modeling the negotiation process, he suggests that one can expect its outcome to reflect a degree of compromise between the parties' goal (maximize M) and the agency's goal, which might be set by high-level policy to involve maximizing some weighted sum $kM + S$. Whenever the parties have any bargaining power, it is optimal to set $k < 1$ as a counterweight; if $k = 1$, then M would be over-weighted relative to S in the resultant. Indeed, if the parties have enough bargaining power, it is entirely possible that total welfare is best served by making $k \leq 0$. Here $k = 0$ would correspond to making the agency pursue consumer welfare and ignore the parties' profits, while $k < 0$ would correspond to a consumer focus with an actual hostility to profits.

Discussions of objectives often assume that participants can accurately evaluate the effects of mergers and of potential remedies. A complementary perspective on merger remedies is informational. To illustrate, consider a proposed merger that will affect only the profits of the merging parties, M, and consumer surplus S. Suppose for simplicity that the court can perfectly gauge S. Although the court may have a good estimate of M, the merging firms are likely to have a better one. This informational asymmetry matters if, for instance, a court is applying a total-surplus standard, finds that $S < 0$, and is uncertain whether $M + S \geq 0$. Under a total surplus standard, the firms have an incentive to claim that M is large.[41]

One resolution takes advantage of the same market mechanism that, in the market for a competitively supplied good, ensures that consumers only consume the good if their consumption value exceeds the marginal cost of production. Namely, make the parties pay a price for their conduct that is equal to the social cost of that conduct. If those who gain must compensate those who lose, this compensation provides a market-like test of what must otherwise be imperfectly judged: that the merger's gain in efficiency outweighs the pre-remedy harm to consumers.

In this view, the point is not that there are benefits when compensation is actually paid. (That is, we are not suddenly concluding that distribution is

39 Analysis of agency negotiations highlights the issue of whether the agency's objective is to maximize some welfare measure or see that it does not fall.

40 Joseph Farrell, *Negotiation and Merger Remedies: Some Problems, in* Merger Remedies in American and European Competition Law 95-105 (Francois Leveque & Howard Shelanski eds., 2003).

41 Recall that we are assuming the court knows the value of *S* and, thus, does not revise its projection of *S* in response to claims about the size of *M*.

important after all.) Rather, requiring actual payment might be the strongest available proof that compensation could be paid. Requiring actual compensation might also motivate firms ex ante to seek out socially desirable mergers, rather than less-efficient but more-profitable ones. Overall, then, compensation might promote efficiency even if, ex post, it is just a transfer or is even inefficient (that is, it costs the firms more to provide than it benefits consumers).

Another possibility is that an agency or court cannot perfectly predict the merger's effect on consumer welfare but can obtain a commitment from the merging parties (e.g., on price) that guarantees that S is positive. With such a commitment, the court can be sure that $W = M + S$ is positive if rational firms want to proceed with their merger.

Intuitively, such a requirement induces the active players—the merging firms—to take broad welfare effects into account. Graphically, the intuition is that requiring compensation makes the merging firms' indifference curves over M and S more like the social indifference curves. Without such a requirement, the firms' indifference curves are simply vertical lines, some of which are illustrated in orange in figure 4a. In contrast, the social indifference curves (drawn in green) are straight lines with slope minus one. Figure 4b illustrates how the firms' indifference curves become more like the social indifference curves when compensation is required—when consumers are harmed, the firms' profits net of compensation vary with the level of total surplus. Simple algebra demonstrates that, when compensation can be paid without transactions or agency costs, requiring compensation may raise total surplus and never lowers it.[42] The importance of this finding, however, is tempered by the reality that transfers are often costly, particularly if targeted at affected consumers, who may number many millions.

Intriguingly, it is not obvious why the compensation must be paid in a coin at all related to the competitive effects. For example, the merging firms might simply pay off buyers if there is no efficient remedy available to undo an increase in market power that is outweighed (in its effect on total surplus) by fixed-cost efficiencies. Although this is not conventional antitrust thinking, and (for instance) the U.S. Federal Communications Commission has been strongly criticized for allegedly seeking merger conditions that are not clearly aligned with competitive harms from the merger, compensation in a different coin is the heart of a market economy. A consumer can legally remove a DVD player from an electronics retailer only if he or she compensates the retailer. It would not be enough for the

42 To see why, consider two possible mergers, 1 and 2, for which $M_1 + S_1 > M_2 + S_2$ and $M_1 > M_2$. Absent a compensation requirement, the merging parties would choose merger 1, which maximizes both profits and total surplus. Suppose, counterfactually, that the compensation system induced the merging parties to choose the less socially desirable merger: $M_1 + \min\{0, S_1\} < M_2 + \min\{0, S_2\}$. This condition can hold only if $S_1 < 0$. There are now two cases to consider. If $S_2 \geq 0$, then $M_2 + S_2 \geq M_2$, and $M_1 + S_1 > M_2 + S_2$ implies $M_1 + \min\{0, S_1\} = M_1 + S_1 > M_2 = M_2 + \min\{0, S_2\}$, a contradiction. If $S_2 < 0$, then $M_1 + \min\{0, S_1\} = M_1 + S_1$ and $M_2 + \min\{0, S_2\} = M_2 + S_2$. Hence it cannot be the case that $M_1 + S_1 > M_2 + S_2$ and $M_1 + \min\{0, S_1\} < M_2 + \min\{0, S_2\}$.

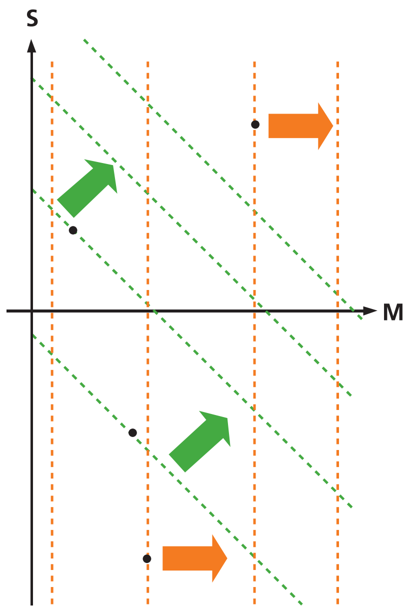


Figure 4a

Social and private indifference curves absent compensation

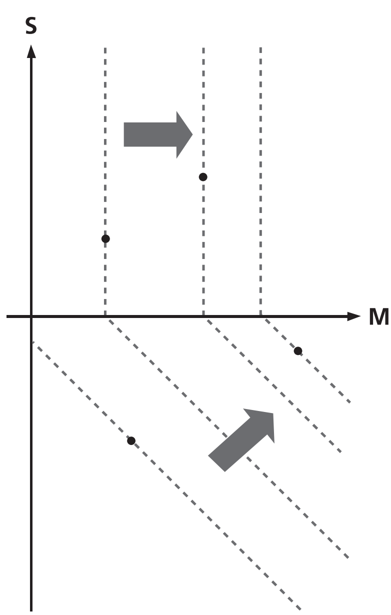


Figure 4b

Private indifference curves under the compensation scheme consumer to prove by expert testimony that he or she valued the player more than the retailer does. Nor is it regarded as suspect that he or she would compensate the retailer in a currency that bears no resemblance to the DVD player. This analogy raises interesting questions regarding why public policy treats a firm's obtaining the right to reduce competition so differently than a consumer's obtaining a good or service. One possibility is that it is harder for the agency to evaluate reliably—and harder for courts to judge whether the agency has evaluated responsibly—whether a distant remedy properly compensates consumers than whether a closely-tailored one does so. An important, related consideration is that allowing an agency to cut deals that involve unrelated conduct can permit the agency to engage in wide-ranging policy making without judicial review and contrary to the wishes of the legislature. For example, an agency might approve a merger conditional on the parties' agreeing to cease certain marketing practices that the agency finds distasteful but believes could not be successfully challenged in court.

VI. Conclusion

We have distinguished three layers of policy objective. At the highest level is the broad objective of governmental intervention in the economy and society. In the middle lies antitrust policy's objective within that overall policy framework. Lastly, there are the objectives of specific decision makers within the antitrust enforcement system.

We have argued that distributional concerns, however legitimate (or established) at the highest level of policy concern, should not be pursued through antitrust policy. In particular, arguments based on distributional concerns do not make a good case for the use of a consumer surplus standard in antitrust. However, analysis of the overall antitrust decision-making system suggests that, in some circumstances, a consumer surplus standard (or consumer surplus standard with a process component) can perform better than a total surplus standard, even if the ultimate goal is to maximize total surplus. Some of those arguments, unsatisfyingly, prove only possibilities. But several economic analyses have explored how outcomes may generally come closer to maximizing total surplus if someone, such as antitrust agencies, contributes a pro-consumer counterweight to firms' representation of their interests by choice of conduct and during lobbying, litigation, and bargaining. That argument, however, has not yet been thoroughly explored.

Analysis of the overall antitrust decision-making system suggests that, in some circumstances, a consumer surplus standard (or consumer surplus standard with a process component) can perform better than a total surplus standard, even if the ultimate goal is to maximize total surplus.

Where does this leave us? We believe that there is a strong case for using total surplus, together with appropriate non-welfarist process criteria, as the overall objective of antitrust policy—and arguably even the process element earns its place through the view that competition promotes total surplus. The case for instructing the agencies and courts to use total surplus (with or without process elements) as their standard is weaker. But we are a long way from being able to conclude that a consumer surplus standard, presumably alongside an anticompetitive behavior prong, is better. At this point, we believe one should not too confidently advocate either a total surplus or a consumer surplus prosecutorial and judicial standard. One of us would nevertheless recommend the use of a total surplus standard at this stage of our knowledge; the other believes that the somewhat murky status quo should muddle along until we understand more.

VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

Welfare Standards and Merger Analysis: Why Not the Best?

Ken Heyer

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Welfare Standards and Merger Analysis: Why Not the Best?

Ken Heyer

Over the past several decades, there has emerged a rough consensus among professional antitrust practitioners, and within the law and economics community generally, that the "competition" referred to in our antitrust statutes is not to be interpreted simply as pre-merger rivalry among entities. Rather, it is best viewed as a process, the outcome of which is welfare, with welfare—not rivalry—being the object of interest. Consistent with this interpretation, scholars, competition authorities, and the courts have come to treat antitrust law as condemning only those mergers whose effect may be substantially to reduce welfare.

In this paper, I argue for using the total welfare standard, rather than the more commonly employed consumer welfare standard. In doing so, I respond to three broad objections that have been raised. One is that use of a total welfare standard conflicts with antitrust law, or at least with legal precedent. A second is that employing a total welfare standard would clearly be more costly for antitrust agencies than employing one or another flavor of a consumer welfare standard. A third is that the total welfare standard ignores important distributional considerations—considerations that are better treated under some form of consumer welfare standard. Each of these objections is evaluated, and ultimately found unpersuasive.

The author is the Economics Director of the Antitrust Division, U.S. Department of Justice. The views expressed in this article are his alone, and do not represent the views of the U.S. Department of Justice.
He would like to thank Doug Melamed, Sheldon Kimmel, Dan O'Brien, Dennis Carlton, Carl Shapiro, Ralph Winter, Mark Frankena, Tim Brennan, Craig Conrath, and Nien-Huei Jiang for their comments on an earlier draft of this article. He would like also to acknowledge the excellent research assistance provided by Gillian Oak. An earlier version of this paper appeared as Economic Analysis Group Discussion Paper EAG 06-8, March 2006.

I. Introduction

Antitrust agencies in the United States, and increasingly around the world, have adopted what has been termed the consumer welfare standard for analyzing proposed mergers. For example, in a recent article, the then-Deputy Assistant Attorney General, and currently Assistant Attorney General of the U.S. Department of Justice's Antitrust Division stated, "Today, most would agree that proper enforcement of the antitrust laws focuses on consumer welfare." He added that "the enforcement authorities in the United States look most frequently at the question of what is best for consumers."[1] And, in a speech given shortly after succeeding Mario Monti as EC Competition Commissioner, Neelie Kroes observed that:

> "Consumer welfare is now well established as the standard the Commission applies when assessing mergers and infringements of the Treaty rules on cartels and monopolies. Our aim is simple: to protect competition in the market as a means of enhancing consumer welfare and ensuring an efficient allocation of resources."[2]

Under the consumer welfare standard, if a merger appears likely to harm consumers as a result of a reduction in competition—some would add significantly, or substantially—in any relevant market, the merger is illegal.[3] This article considers the basis for applying a consumer welfare standard, and examines the arguments for instead employing a total welfare standard (i.e., a standard that considers a merger's likely effect on all members of society, not simply the consumers of products produced by the merging firms).[4]

1 Thomas O. Barnett, *Substantial Lessening of Competition-The Section 7 Standard*, 2005 COLUM. BUS. L. REV. 293, 295-298. Barnett states explicitly that "The views and opinions expressed herein are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice."

2 Neelie Kroes, European Competition Policy – Delivering Better Markets and Better Choices, Presentation at European Consumer and Competition Day (Sep. 5, 2005).

3 As discussed later in this paper, U.S. competition agencies will at times consider efficiencies "not strictly in the relevant market, but so inextricably linked with it that a partial divestiture or other remedy could not feasibly eliminate the anticompetitive effect in the relevant market without sacrificing the efficiencies in the other market(s)."

4 This article does not address the proper welfare standard to apply in the case of civil non-merger investigations, such as those implicated by Section 2 of the Sherman Act or Article 82 of the EC's competition law. While, in principle, the economic case for a total welfare standard would seem to be equally applicable outside the narrow setting of merger policy, issues such as ease of application,

As an initial matter, it is quite clear that the relevant sections of U.S. antitrust law say nothing about welfare—consumer or otherwise. Rather, they state that mergers are illegal when their effect "may substantially reduce competition in any line of commerce."

Over the past several decades, there has emerged a rough consensus among professional antitrust practitioners, and within the law and economics community generally, that the competition referred to in our antitrust statutes is not to be interpreted simply as pre-merger rivalry among entities. Rather it is best viewed as a process, the outcome of which is welfare, with welfare—not rivalry—being the object of interest. Consistent with this interpretation, scholars, competition authorities, and the courts have come to treat antitrust law as condemning only those mergers whose effect may be substantially to reduce welfare.

That having been said, there remains a question of which welfare standard to use, and exactly whose welfare to consider. Several candidates suggest themselves. One is the welfare of consumers in each of the markets potentially impacted by the merger. Under this standard, a merger is permissible if (and only if) consumers in each and every one of the markets at issue are likely to be at least as well off after the merger as they were before it. One might call this an actual Pareto consumer welfare standard, though for reasons explained later in this paper, applying even this standard does not necessarily ensure that each and every consumer will be made better off.

A second approach would be to permit mergers whose net effect on consumers across all the (possibly multiple) markets served by the merging parties is positive. Using this standard, a merger would be permitted even if consumers are harmed in market A, so long as the benefits received by consumers in other markets (B, C, . . . , Z) served by the merging firms are in aggregate greater. One might refer to this as a potential Pareto consumer welfare standard.

A third approach, one that has not, to my knowledge, been adopted explicitly by any major competition authority, is to permit mergers whose predicted effect on the total welfare of members of society as a whole is positive.[5]

footnote 4 cont'd

among others, distinguish the two situations. For discussions of the appropriate standard to apply outside the merger setting, *see* Gregory Werden, *Identifying Exclusionary Conduct under Section 2: The 'No Economic Sense' Test*, 73 Antitrust L.J. 413 (2006), Steven C. Salop, *Exclusionary Conduct, Effect on Consumers, and the Flawed Profit-Sacrifice Standard*, 73 Antitrust L.J. 311 (2006) and A. Douglas Melamed, *Exclusive Dealing Agreements and other Exclusionary Conduct – Are There Unifying Principles*, 73 Antitrust L.J. 375 (2006). Neither, I would note, proposes that total welfare maximization be the test.

5 The welfare standard employed in Canada lies somewhere between a consumer and a total welfare approach. Section 96 (1) of the 1986 Competition Act of Canada explicitly provides for an "efficiencies defense" for mergers that might result in higher prices for consumers. As interpreted by Canadian

footnote 5 cont'd on next page

Application of this standard requires that weight be given not only to the welfare of those who consume the merging firm's products, but also to those doing the producing.[6]

In one very real sense, an economy's producers are consumers as well, albeit consumers of many items other than the ones they happen to produce. There is, of course, a difference between the act of consuming and the act of producing, and most individuals in our highly specialized economy do not restrict themselves to consuming only what they themselves produce; they enhance one another's welfare through trade. This distinction, however, hardly suggests a meaningful basis for weighting the welfare of individuals occupying these two roles unequally; much less, weighting the welfare of a market's producers at zero when determining whether a merger is, on balance, beneficial to society. In any event, it seems reasonable to place the burden of proof on those who would defend the use of a narrower, consumer welfare standard, rather than a total welfare standard that accounts for the wellbeing of all the members of an economy. As discussed later in this paper, it is far from clear that this burden has been carried.

This paper makes a case for employing the total welfare standard. In the course of doing so, it responds to a number of possible defenses for antitrust's current exclusive focus on the welfare of consumers in the relevant markets impacted directly by proposed mergers.

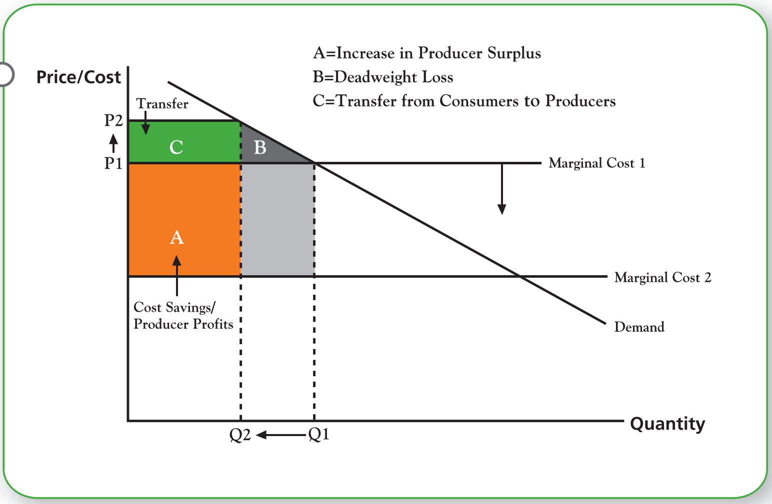
The issue is not a new one. In 1968, Oliver Williamson famously described, using what he termed the "naive tradeoff model," the tradeoff that arises when a

footnote 5 cont'd

courts in the recent *Superior Propane* litigation, an efficiencies defense for mergers where consumers are likely to be harmed must employ a so-called "balancing weights" approach—i.e., an approach in which harm to low-income consumers is afforded disproportionately greater weight. Elaborating on the logic employed in this decision, Ross and Winter (2005) show that use of a balancing weights approach, where the weights are those revealed in government tax or redistribution policy generally, would only in the most extraordinary circumstances produce a ruling different from that generated by the total surplus rule. For an excellent discussion of the *Superior Propane* case, other recent Canadian court decisions, and an economic analysis of Canada's efficiency defense, *see* Thomas W. Ross & Ralph A. Winter, *The Efficiency Defense in Merger Law: Economic Foundations and Recent Canadian Developments*, 72 ANTITRUST L.J. 471, 471-505 (2005). *See also* Discussion *in* Report of the Advisory Panel on Efficiencies to Scott Sheridan, Commissioner of Competition (Aug. 2005).

6 Although those most directly affected by a merger, and those for whom the merger's likely effects may be easiest to identify and calculate, are the merging firms and their customers, total welfare technically includes also the welfare of any and all who may be affected by the merger. This includes, in principle, the welfare of those who compete against the merging firms and, to the extent higher profits are shared with them, the workers at firms whose profits are affected. I do not, in this article, consider the effect of mergers on consumers, producers, or workers in countries outside of the competition authority's jurisdiction.

Figure 1



merger simultaneously produces cost savings—from realization of efficiencies and higher prices—from greater market power.[7]

Williamson went on to present what he termed "illustrative" results, showing that it may take very small percentage cost savings to completely offset the negative welfare effects of even a significant increase in market power. Indeed, much of the subsequent commentary on Williamson's influential article dealt more with the implicit assumptions that generated this contentious result, than with the proposal that merger policy employ a total, rather than a consumer, welfare standard.[8]

7 Merging firms may be made better off in yet another way that can leave consumers worse off. Consider a situation where two duopolists had been colluding—tacitly, perhaps. If a merger makes it easier for them to price discriminate and increase output, this may leave consumers worse off (though it may also leave them better off).

8 *See* Oliver E. Williamson, *Economies as an Antitrust Defense: The Welfare Tradeoffs*, 58 Am. Econ. Rev. 18 (1968), reprinted in 1 Competition Pol'y Int'l 217 (2005). In responding to critics and qualifying his admittedly naive model, Williamson recognized that accurately measuring a merger's net effect on total welfare properly incorporates any losses in producer surplus—that occur when, as is commonly the case, the merging firms were setting price above marginal cost even before exercising greater market power post-merger. Williamson concluded that adjusting for this did not materially affect his conclusions. For criticisms of Williamson's original article, *see* Michael E. DePrano & Jeffrey B. Nugent, *Economies as an Antitrust Defense: Comment*, 59 Am. Econ. Rev. 947 (1969); *see also* Alan A. Fisher & Robert H. Lande, *Efficiency Considerations in Merger Enforcement*, 71 Cal. L. Rev. 1580 (1983). For responses by Williamson to his critics, *see* Oliver Williamson, *Economies as an Antitrust Defense: Reply*, 59 Am. Econ. Rev. 954, 954-959 (1969) and Oliver Williamson, *Economies as an Antitrust Defense Revisited*, 125 U. Pa. L. Rev. 669 (1977). For a more recent treatment of similar issues, *see* L.-H. Röller, J. Stennek, and F. Verboven, Efficiency Gains from Mergers (The Research Institute of Industrial Economics, IUI Working Paper Series, No. 543, 2000), *available at* http://swopec.hhs.se/iuiwop/papers/iuiwop0543.pdf.

II. Welfare and Efficiency

The concepts of economic welfare and economic efficiency are closely related to one another. Economists say that an economy is operating at maximum efficiency when society is squeezing the greatest value—the highest level of welfare—out of its scarce resources. The independent actions of profit and utility-maximizing economic agents work towards producing this desirable outcome in ways that are familiar to students of elementary microeconomics. Competition among firms to obtain the patronage of consumers spurs them to produce those goods and services that are most highly valued by consumers, to do so at the lowest possible cost (for example, by finding ways of producing the same quantity of output with fewer inputs), and to drive prices down towards the marginal cost of production (thereby resulting in output up to the point at which additional value to consumers no longer exceeds the additional cost to society). In this way, competition works—"as if by an invisible hand" as Adam Smith famously observed—to squeeze the greatest possible value out of society's scarce resources.

One of the ways in which production costs are minimized is by efficiently combining the inputs that produce the goods and services we ultimately consume. The entities that typically accumulate and process inputs into final products are called firms, though at times inputs may be combined by independent agents through arms-length contracts with one another. Firms will at times seek to grow through merger, though an alternative may be to grow internally, or perhaps to expand via arms-length contracts with other, still independent, firms.

III. Mergers and Efficiency

Getting a product to market involves a number of discrete, yet critical, steps. These steps may include some or all of the following: basic research, applied research, product design, product manufacture, marketing, distribution, or service. Not uncommonly, firms that produce final products in close competition with one another have different strengths and weaknesses in these various steps. Auto firm A, for example, may be better than auto firm B when it comes to coming up with innovative ideas and quality control, while auto firm B may be better when it comes to marketing and post-sale servicing. Combining the best of both can produce synergies that in principle permit lower-cost production of an even better product.[9]

It is useful to discuss briefly why contracts that maintain the independence—especially the pricing independence—of two competitors with relatively different strengths will not always be a feasible or equally efficient method of obtaining the economic benefits expected from the merger. Industrial organization economists have been, at least since publication of Ronald Coase's landmark article on the

9 Many of these essentially complementary efficiencies may be available only by merging the operations of competitors.

theory of the firm, intrigued with the question of when and why the costs of organizing transactions within firms will tend to be lower than the costs of organizing transactions across entities via contract.[10] Several decades after the appearance of Coase's article, work by others, particularly Oliver Williamson, put more theoretical flesh on the bare bones that Coase had first exposed.[11] Over time, a number of industry studies provided a degree of empirical support for today's commonly accepted notion that transactions organized within a firm can, in many cases, economize greatly on the transactions costs associated with writing, monitoring, and enforcing contracts.[12] While important theoretical questions about the efficiency of operations even within a firm continue to be studied, there is little disagreement that it will frequently be less costly to conduct transactions within a firm than to do so across firms. And unless it is equally costly for a firm to grow internally rather than through acquisition, this implies that cost savings may be merger specific. In other words, merger will at times be the most efficient means through which firms satisfy the demands of consumers.

A merger can be thought of as a special sort of contract, an all-encompassing one, if you will, whereby the decisions of two formerly independent firms will be subject to the authority of a single entity. Or, put differently, it can be thought of as when two formerly independent firms contract to become a single firm. Firms may merge to obtain greater market power. They may also merge to achieve efficiencies—i.e., to reduce costs.[13] The efficiencies potentially realizable through merger are numerous, as are the means through which these benefits can be achieved.[14]

Broadly speaking, efficiencies will tend often to take either of two forms: ones that lower marginal production costs, and ones that generate savings in fixed

10 *See* R.H. Coase, *The Nature of the Firm*, 4 Economica 386 (1937).

11 *See* Oliver E. Williamson, Markets and Hierarchies: Analysis and Antitrust Implications (1975).

12 *See* Benjamin Klein, Robert G. Crawford, & Armen Alchian, *Vertical Integration, Appropriable Rents, and the Competitive Contracting Process*, 21 J.L. & Econ. 297 (1978). *See also,* Russell Pittman, *Specific Investments, Contracts, and Opportunism: The Evolution of Railroad Sidetrack Agreements*, 34 J.L. & Econ. 565 (1992); and Paul L. Joskow, *Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets*, 77 Am. Econ. Rev. 168 (1987).

13 I am referring here to a lowering of the cost function, not to a reduction of costs that occurs purely as a consequence of reducing quality or output.

14 This raises a question of how to treat reductions in marginal cost that arise because of procurement cost savings. The answer is that procurement cost savings arising from resource cost savings—e.g., fewer resources required when there are longer production runs—are welfare-enhancing. Indeed, they are efficiencies that likely result in greater output as well. Procurement cost savings that arise from merger-generated monoposony power, however, are less likely to generate increases in welfare. Unless the exercise in monopsony power is offsetting preexisting market power on the selling side, these benefits to the merged firm will likely result in lower output, and will in any event result in inefficient production of pre-merger levels of output.

costs. Efficiencies can lower the cost of producing existing products. They can also promote the development of entirely new or better products. One way in which this latter type of benefit—dynamic efficiencies—can in theory be enhanced is for merging firms to eliminate redundant R&D activities and instead allocate the firms' limited assets towards multiple, alternative projects. Dynamic efficiencies may themselves be realized in a variety of ways, and one may ask whether efficiencies that make innovative activity more likely to occur, or likely to occur at lower cost, are more properly viewed as fixed cost savings or marginal cost savings. The important point is that whatever label one applies, and regardless of how the benefits from dynamic efficiencies are split between lowering prices and developing entirely new products or processes, dynamic efficiency generates an increase in total welfare.[15]

DISTINCTIONS BETWEEN FIXED AND MARGINAL COST TEND TO BE PARTICULARLY IMPORTANT WHEN COMPETITION AUTHORITIES EMPLOY A CONSUMER, RATHER THAN A TOTAL, WELFARE STANDARD.

Distinctions between fixed and marginal cost tend to be particularly important when competition authorities employ a consumer, rather than a total, welfare standard. The reason is as follows: unlike changes in marginal cost, changes in fixed cost generally do not alter the firm's profit-maximizing price, or the level of output at which the firm maximizes its profits, unless they affect the firm's very viability.[16] As a result, pure fixed cost changes, no matter how large, may have no effect at all on the welfare of consumers in the relevant market.[17] In terms of their effect on a firm's profit-maximizing price, higher fixed costs can be compared with someone breaking into the company's headquarters and stealing a large sum of money from the firm's

15 Another oft-cited category of possible efficiencies from merger is realization of scale economies. For reasons given in Farrell and Shapiro, however, achieving through merger pure scale economies will often not be merger specific. Or if it is, the fact that the merging firms had not been achieving these efficiencies without merging may suggest strongly that the market is not performing competitively. Joseph Farrell & Carl Shapiro, *Scale Economies and Synergies in Horizontal Merger Analysis*, 68 ANTITRUST L.J. 685 (2001). Röller et al. categorize potential efficiencies from merger as either rationalization, economies of scale, technological progress, purchasing economies, and reduction of slack (managerial and X-efficiency) (*See* Roller et al. *supra* n. 8). *See also,* William J. Kolasky & Andrew R. Dick, *The Merger Guidelines and the Integration of Efficiencies into Antitrust Review of Horizontal Mergers*, 71 ANTITRUST L.J. 207 (2003).

16 Although it sometimes surprises attorneys to hear it, reducing the marginal cost of a firm with market power makes it profitable for the firm to reduce price and increase output. Because firms maximize profit by setting marginal revenue equal to marginal cost, fixed cost savings do not have this effect.

17 Under a consumer welfare standard, even fixed cost savings would properly be given some weight if they were ultimately passed along to consumers in the form of lower prices. Arguing that all costs must be recovered in the long run, some would contend that fixed cost savings would, eventually, be reflected in lower prices. This intuition no doubt provides part of the rationale for the willingness of the U.S. antitrust agencies, as reflected in the Horizontal Merger Guidelines, to "consider the effects of cognizable efficiencies with no short-term, direct effect on prices in the relevant market." The extent to which fixed cost savings actually will be passed through, and how quickly that might occur, will depend at least in part on the strength of competition post merger.

safe. Conversely, lower fixed costs are akin to some anonymous benefactor depositing a large sum of money into the firm's bank account. In the first case the firm is worse off, and in the second it is better off, yet in neither case is there reason to expect a change in the price the firm finds it most profitable to set, or the level of output the firm finds it most profitable to produce.

Importantly, however, unlike in the case of pure money transfers, fixed cost savings have significant efficiency implications for the economy as a whole. As discussed below, by freeing up resources for use elsewhere in the economy, fixed cost savings enhance an economy's total welfare. These potential benefits from merger are given zero weight when applying a narrow consumer welfare standard.[18]

IV. The Actual Pareto Consumer Welfare Standard

As an initial matter, when a merger has no effect other than to lower the (quality-adjusted) price for final goods sold in a market, some consumers in that market will benefit and no consumers will be harmed. Those who had been consuming the product before the merger will be able to purchase their original quantities at a lower price, and additional surplus will be obtained by consumers who, at the now-lower price, consume even greater quantities than before. In addition, individuals who had in the past maximized their utility by consuming zero quantities of the product may be better off by making some purchases at the now-lower price. Thus, all consumers of the product appear to be better off.

Even in the case of price-lowering mergers, however, it will not necessarily be true that all consumers everywhere will be better off. An efficient merger may drive one or more rivals out of business, and consumers who preferred the version offered by exiting rivals may now find themselves worse off. Related to this point, efficiencies sometimes arise from combining complementary assets and standardizing on a single platform or a single standard. Where the two merging

18 The Horizontal Merger Guidelines issued by the U.S. Department of Justice and the U.S. Federal Trade Commission outline the approach to efficiencies taken by the U.S. federal competition authorities. They state, "The Agency will not challenge a merger if cognizable efficiencies [i.e., efficiencies that are merger-specific, that have been verified and do not arise from anticompetitive reductions in output or service] are of a character and magnitude such that the merger is not likely to be anticompetitive in any relevant market." In a footnote, the Guidelines qualify this statement, noting that although the Agency generally focuses on whether cognizable efficiencies likely would be sufficient to prevent even short-term price increases in the relevant market, "The Agency also will consider the effects of cognizable efficiencies with no short-term, direct effect on prices in the relevant market," though the Guidelines state that benefits from such efficiencies "will be given less weight because they are less proximate and more difficult to predict." This qualification permits U.S. competition authorities, in some circumstances, to depart from what I am referring to as the consumer welfare standard. Nevertheless, neither the U.S. competition authorities, nor competition authorities in most other economies, appear willing to adopt explicitly and unambiguously a total welfare standard for merger enforcement. U.S. Dep't of Justice & Federal Trade Commission, Horizontal Merger Guidelines §4 n.36 (1992) (revised Apr. 8, 1997) [hereinafter *Merger Guidelines*].

firms had previously been offering competing and incompatible methods of satisfying consumers, efficient standardization will typically strand the investments of consumers who had invested in the to-be-jettisoned standard. They will be left worse off, even though consumers in the market as a whole are better off from having a better, cheaper, more ubiquitous standard as a result of the merger.

Producers in the relevant markets may be either better off, or worse off. Certainly the merging firms believe that they will be better off, as evidenced by the fact that they have chosen to merge, presumably, voluntarily. Rivals of the merging firms are likely to be harmed, since a price-lowering merger may well force them to compete harder, perhaps by lowering their own prices, and they may well lose business to the more efficient merged firm. In addition, firms not even in the relevant markets may be worse off to the extent that demand for their product falls when consumer patronage shifts to products whose price has fallen as a result of the merger. While these latter categories of producers are worse off—indeed, producers may collectively be worse off—an actual Pareto consumer welfare standard would bless the merger.

As emphasized recently in work by Steve Salop**,** merger-generated efficiencies can in theory actually lower total welfare—as a consequence of shifting sales towards the merging parties and away from their rivals.[19] Salop presents an example where two relatively high-cost firms with relatively small shares achieve marginal cost savings through merger. As a consequence of lowering their marginal cost, they reduce price somewhat. This, in turn, results in greater sales for them and lower sales for what may be a (still) more efficient rival.[20] Although the reduction in the merged firm's marginal cost will likely lead to at least somewhat greater sales in the relevant product market, the pre-merger level of output will be produced at higher total cost. In such a circumstance, the net effect would be gains for consumers, but quite possibly lower total welfare—after one adjusts for the net negative effect on producers as a whole.[21]

19 Steven C. Salop, *Question: What is the Real and Proper Antitrust Welfare Standard? Answer: The True Consumer Welfare Standard*, Presentation to Antitrust Modernization Commission (Nov. 4, 2005).

20 In answer to the question "How do higher-cost firms manage to survive in the market?" Salop suggests that it may be because they provide a differentiated product. To the extent that consumers value the particular variety produced by the "high cost" firms, Salop may not be making the point as strongly as he might. Given that many oligopoly models—Cournot among them—generate equilibria where the lowest cost firm finds it profitable to price high enough to keep its rivals viable, the general point can be made even more strongly by simply assuming that all firms in the market produce and sell a homogenous product.

21 Yet another somewhat counterintuitive result can occur in circumstances where a merger produces no cost savings at all, but where higher prices set by the merging firms induce customers to shift some of their patronage toward substitutes. If the products to which customers turn have relatively high price-cost margins and demand for the good whose price has risen is sufficiently inelastic, total welfare can increase precisely because price has gone up. I thank Ralph Winter for this observation.

Scenarios such as this may or may not be rare, yet they represent another category of cases—ones where prices are actually lowered by a merger—in which application of a pure consumer welfare standard would be costly to the economy as a whole. Taking fully into account such possibilities—akin in certain respects to what are termed "second best" considerations[22]—may be very difficult in practice. However, the example alerts us to the possibility that looking only at a merger's effect on the welfare of consumers and the merging firms can be too narrow a focus if it ignores inefficient shifts in production across firms and leads one falsely to conclude that a merger has raised total welfare, when in fact it has not. The original Williamson diagram, therefore, suggests incorrectly that when a merger results in lower prices total welfare inevitably rises. This is not a general result.

Though concern for the welfare of parties affected only indirectly by the effects of a merger ought not be completely irrelevant to enforcement decisions, difficulties associated with estimating such effects (ones that may well be second order), and requiring that a general equilibrium analysis be conducted of every merger, would seem to argue for imposing a fairly high burden of proof on those asking that competition authorities base enforcement decisions on such arguments.

V. Fixed Cost Savings and Total Welfare

There are a number of situations in which merging parties anticipate capturing efficiencies that, under traditional consumer welfare criteria, would not help them avoid an adverse enforcement decision (or court ruling). Thus, the choice of standard may be of more than simply theoretical interest. Mergers contemplated because they will likely produce significant fixed cost savings tend to be of this form. Consider, for example, the situation where firm A and firm B compete with differentiated products, and where firm A happens to have a good deal of unutilized capacity in its factory. The reasons for this disequilibrium may be several, but let's assume that the available capacity is temporarily excessive because there has been a significant and unanticipated drop in demand for firm A's product.[23]

The excess capacity in the hands of firm A can, let us assume, be used to produce the entire projected output of competing firm B. In such circumstances—that I suspect are hardly unique—consolidating all the production of both firms in the partially vacant plant of firm A would clearly lower the total cost of producing this output. This is because part of the cost to society of producing firm B's output in firm B's plant is that resources are being used for this purpose that

22 R. G. Lipsey & Kelvin Lancaster, *The General Theory of Second Best*, 24 Rev. Econ. Stud. 11 (1957).

23 Unused capacity need not, of course, be excess in an economic sense. It may instead be serving a valuable function in the event that demand for the firm's product increases in the next period. By assumption, I am ruling out this explanation. In addition, the fact that factory capacity is durable implies that the firm cannot readily and immediately adjust its capacity to fit the now-smaller projected demand for the output of its product. Over time, of course, firms will adjust their investment decisions to reach a new, long-run equilibrium. Nevertheless, this will not necessarily happen quickly.

could be reallocated elsewhere in the economy, producing a net gain in total welfare. In particular, assuming the fixed cost savings are merger-specific, a cost of prohibiting a merger between firms A and B would be the opportunity cost of continuing to run the plant of firm B.[24] If firm B's plant would be closed by a merged firm, and particularly if the merged firm would continue supplying the market with all or nearly all of both firm A and firm B's pre-merger output, economic benefits could be substantial.

Benefits might consist of the value—net of their production costs—of alternative products produced out of the now empty plant. Or, if the highest alternative use of the plant is to tear it down and sell it as scrap, then the value of that scrap (net of demolition costs)—plus the value in its highest alternative use of the land on which the plant currently sits—would represent economic benefits from the merger. These would all be net benefits to the economy—an increase in total welfare. The fact that they do not involve a reduction in the merged firm's marginal cost—and thus do not result in any pass-through to the merged firm's consumers—does not change the fact that the merger is welfare enhancing.[25] Under a consumer welfare standard, the merger would be blocked if there is a small increase in market power.

VI. Marginal Cost Savings Not Fully Passed Through

As discussed above, fixed cost savings tend not to be passed on to consumers in the relevant product markets at all, while marginal cost savings in markets potentially raising competitive concerns generally result in lower prices. That having been said, the degree of pass-through from mergers that lower marginal costs will differ from case to case, as it is a function of many factors—including both demand conditions and the particular oligopolistic game being played by firms in the market.[26]

Much like fixed cost savings, those marginal cost savings that do not result in lower prices are both benefits to society as a whole and, under a consumer welfare standard, not an acceptable defense to a transaction that is likely to raise price.

24 As discussed in Farrell & Shapiro (2001), such efficiencies are more likely to be merger-specific when firm A can't simply produce and sell the total output of the two merging firms at constant cost through its own plant. This may be because the products of the merging firms are branded and contracting costs are substantial. *See* Farrell & Shapiro, *supra* n. 15, at 12.

25 Of course, if the marginal cost of production at firm A's plant is lower than the marginal cost of production at firm B's plant, efficiencies would be even greater—though in this case we would expect the lower marginal cost of production to translate into some pass-through to final consumers and a concomitant increase in consumer surplus (i.e., welfare).

26 A pure price taker—an infra marginal producer—will not find it profitable to pass on in the form of lower prices even marginal cost savings. Rather, it will keep those savings as rents.

VII. The Benchmark: Total Welfare

Let us begin with the standard definition of total welfare. In any single market, total welfare is conventionally defined as total surplus—the difference between the value consumers place on output, minus what it costs society to produce that output. Across all markets, total welfare is simply the sum of all surplus, irrespective of how it happens to be divided between consumers and producers. In a hypothetical world populated by only a single individual, that individual would do best by organizing his or her affairs so as to maximize the total value obtained from the scarce resources he or she has to work with.[27]

Adding to the population does not obviously negate this core principle; however, it does introduce issues of distributional equity, which I discuss in some detail below. In any event, the difference between the value to consumers and the cost of production is exactly what economists mean by total welfare. From the standpoint of society as a whole, maximizing it would seem, at least at a first approximation, to be a desirable objective. Anything short of this is akin to asking society to make do with less, rather than with more.[28]

What reasons might there be for departing from this standard when developing merger policy? We can consider at least three categories of objections. One is that use of a total welfare standard conflicts with antitrust law, or at least with legal precedent. A second is that employing a total welfare standard would be more costly for antitrust agencies than employing one or another flavor of a consumer welfare standard. A third possible objection, one neatly abstracted away from in our example of a single individual populating the economy, is that the total welfare standard ignores important distributional considerations—ones that are better treated under some form of consumer welfare standard. Each of these objections is evaluated, and ultimately rejected, in the analysis below.[29]

27 Of course, in such a world it is hard to imagine any need for a merger policy. Still, the point holds.

28 It is worth noting that literal application of a pure consumer welfare standard, as that term is being defined here, would appear to immunize consumer buyer groups that exert efficiency-reducing monopsony power over sellers. I suspect that many supporters of a consumer welfare standard for sellers would be uncomfortable applying its logic equally to the buyer side of the market. Moreover, economic cost-benefit analyses of proposed government activities and regulations tend to employ a distribution-neutral framework, though these studies may attempt to estimate, or even propose ways of ameliorating, associated distributional consequences.

29 A somewhat different line of argument for employing a consumer welfare standard has been presented by Lyons. Lyons observes that because firms, rather than enforcement officials, get to select which mergers are proposed, they will select those mergers that are most profitable, not necessarily those that maximize welfare, from among the set that competition authorities permit. Lyons then shows that because of this self-selection effect, enforcement by competition authorities of a total welfare standard instead of a consumer welfare standard can, under some circumstances, actually lower total welfare. While Lyons' theoretical results are intriguing, his analysis fails to show that use of a total welfare standard is actually likely to produce such perverse outcomes, much less that use of a total, rather than a consumer welfare standard will systematically bias outcomes and lead to lower total welfare. *See* Bruce R. Lyons, Could Politicians Be More Right Than Economists? A Theory of Merger Standards (University of East Anglia Centre for Competition & Regulation Working Paper CCR 02-01, revised, May 2002).

VIII. Legal Impediments to Use of a Total Welfare Standard?

Section 7 of the Clayton Act prohibits mergers, the effect of which "may substantially reduce competition *in any line of commerce*." (emphasis added) It does not say anything about consumers.[30] An argument could certainly be made that the "in any line of commerce" language implies that mergers are illegal whenever the result is net harm in any relevant market (irrespective of whether the net benefits outside of that relevant market would be even greater). While under this argument one might condemn welfare-enhancing mergers whose primary benefit is to consumers in some other markets, it does not, by itself, support condemning mergers whose benefits to some in the relevant market (namely, those producing the goods being consumed), exceed the harms to others in the same relevant market (those doing the consuming).

Only a seemingly arbitrary decision to weigh more heavily the welfare of some individuals in society than others would do that. If, in particular, a merger causes harm to consumers of product A and yet the fixed cost savings from no longer producing and selling product A would exceed this harm, then treating the welfare of consumers and producers of product A equally would seem to imply that the merger enhances (total) welfare "in any line of commerce."

The literal language of Section 7 would seem, if anything, more likely to rule out use of a potential Pareto consumer welfare standard than to trump a total welfare standard. In the former case, at least the beneficiaries whose gains outweigh the harm to be suffered by individuals within a specific line of commerce (or relevant market) are by definition outside that line of commerce. Conceivably, therefore, consideration of these benefits might run afoul of the law's prohibition against mergers likely to reduce competition substantially "in any line of commerce."

Nevertheless, the federal antitrust agencies, if not yet the courts, have stated explicitly that under certain circumstances they will employ their prosecutorial discretion to not challenge such mergers. In particular, the most recent edition of the U.S. Department of Justice and U.S. Federal Trade Commission's Horizontal Merger Guidelines observes that:

30 I recognize that there is a literature debating just what objective function those who legislated the Sherman Act were really seeking to maximize, and that there are those who feel strongly that the act was passed to protect the merging firms' consumers. This article takes no position on the original intent of U.S. merger policy's founding fathers. To the extent that legislative history truly presents a bar to use of a total welfare standard, an implication of this article is that new legislation to correct this error would be desirable. In any event, for those countries where there is less clearly a legal bar to use of a total welfare standard, such use would be in their economies' best interest.

> "In some cases . . . the Agency in its prosecutorial discretion will consider efficiencies not strictly in the relevant market, but so inextricably linked with it that a partial divestiture or other remedy could not feasibly eliminate the anticompetitive effect in the relevant market without sacrificing the efficiencies in the other market(s)."[31]

Incorporating into one's decision-making out-of-market consumer benefits that are inextricably linked to in-market consumer harms makes eminent economic sense. Less clear is why these benefits should, from either a legal or an economic perspective, receive greater weight than the benefits to producers—whether the latter are achieved in-market or out-of-market.

All of this having been said, the language of the Clayton Act explicitly concerns itself with maintaining competition, not welfare. Doesn't the concern with competition imply that the Clayton Act aims to protect the beneficiaries of competition—consumers?[32] Perhaps. And yet, even here the case for a consumer welfare standard is less than clear cut. Merger to monopoly, for example, reduces competition by definition.[33] Nevertheless, if the merger-specific marginal cost reductions are large enough, even in these cases the affected consumers are better off, and welfare (however defined) rises. Does the Clayton Act condemn such mergers? Should it do so? If we agree that the ultimate determinant of whether a merger under the law (a law that is conspicuously silent as to the welfare measure it endorses) depends (or should depend) on its effect on welfare rather than on competition per se, then it seems fair to consider whether the appropriate measure of welfare should be consumer welfare or total welfare.

IX. Costs of Change?

Once a precedent, or a policy, has been around for a sufficiently long period of time, individuals are likely to have come to rely on it. More specifically and more significantly from an economics perspective, the reliance that individuals place

31 Merger Guidelines, *supra* n. 18, at 13.

32 One might fairly observe that because the law talks about competition, not about benefits, mergers that reduce competition—for example, merger to monopoly—are at the heart of the statute's concerns. In this respect, use of a consumer welfare standard could perhaps be deemed a sensible way of making operational the statute's notion of reduced competition. On the other hand, to get from competition to consumer welfare requires the introduction of a benefit standard. Otherwise, the law is only protecting the act of rivalry itself, which as discussed earlier, the law has (appropriately, in the view of many) moved beyond.

33 And, it certainly satisfies the statute's concern with mergers that "tend to create a monopoly."

on a longstanding policy may have led them to sink investments in anticipation of the precedent not being overturned. Both from the standpoint of equity and efficiency, change—even change to a better policy (or standard)—can impose significant costs. Is this consideration likely to present a serious objection to shifting from a consumer to a total welfare standard in merger analysis? It is not likely that it would, and the application to merger policy of the stare decisis doctrine seems weak.

In particular, there seems little reason to believe that a change in standard would cause either inefficiency or an inequitable effect on those who have taken past actions in reliance on the current standard. A change in standard would not be applied retroactively to mergers that have already been consummated, and on a going forward basis it is hard to imagine significant costs of shifting to a total welfare standard for mergers that have not yet even been proposed. We hardly have a situation where market participants, relying on the consumer welfare precedent, have made significant sunk investments based on the assumption that a consumer welfare standard would continue to be used on into the future. Shifting towards a total welfare standard for review of future mergers would seem to provide guidance that is no less clear to potential merger parties. It would also have the added benefit of encouraging an even larger number of efficient mergers than have taken place in the past.

SHIFTING TOWARDS A TOTAL WELFARE STANDARD FOR REVIEW OF FUTURE MERGERS WOULD SEEM TO PROVIDE GUIDANCE THAT IS NO LESS CLEAR TO POTENTIAL MERGER PARTIES.

One cost of changing standards would be a need for the antitrust bar, consultants, and courts who have become educated in just what does or doesn't satisfy a consumer welfare test to become retrained in what constitutes an increase in total welfare. My sense is that these costs are likely to be relatively small. Certainly, they will be far smaller than the costs that were incurred in the course of moving over the past three decades towards a more economics-based approach to merger analysis generally.

X. Relative Costs of Administering the Alternative Standards

If the costs—to competition agencies, firms, consultants, courts—of employing a total welfare standard were likely to be significantly higher than the costs of employing a consumer welfare standard, this would be an argument for sticking with what we have—warts and all. It is not obvious, however, that this is true. And even if it were, in determining how much weight to be given administrative ease when deciding on a welfare standard, one should be mindful of the old saying about looking under a street lamp for one's lost keys simply because the light there is better. Indeed, it is plausible that in many investigations it would be easier, rather than more difficult, to employ a total welfare standard.

Supporters of a consumer welfare standard might contend that applying it is fairly simple; all one needs to do is determine whether price will rise or fall. Unfortunately, even this is far more difficult to determine in practice than it is to state in principle—even if one were to assume away the potential for merger-specific improvements in product quality.

In order to gauge the actual effect on a product's consumers, economists require reasonably accurate information about the shape of the demand curve for that product within the relevant range. The costs of obtaining reliable information of this sort may be considerable. Absent this information, it is difficult to estimate confidently the extent to which marginal cost savings of any given amount will be passed on to consumers in the form of lower prices.[34]

In addition, estimating the price effect from a merger-generated reduction in marginal cost of any given amount requires information about the competitive game being played by market participants. Cournot, Bertrand, and other specific types of oligopolistic competition have different implications for the extent of pass-through from a given cost reduction, just as they have different implications for the extent to which a merger will enhance market power. And, all of this is complicated even more by the fact that the type of competitive game being played may itself change as a result of the merger, implying either greater or lesser pass-through than if the game remained unchanged.[35]

Beyond these difficulties, and importantly, in order to apply the consumer welfare standard the analyst needs not simply an estimate of merger-specific efficiencies, but also an estimate of the anticipated merger-induced reduction in marginal cost. The practical difficulties of distinguishing between those cost savings that impact incremental sales, and thus will to some extent factor into future pricing, and those that are fixed, and hence are unlikely directly to affect the profit-maximizing price, can be substantial.[36] In my experience, considerable resources tend to be spent (and wasted) by merging parties, by their consultants, by the compe-

---

34 Werden, however, shows that in the case of differentiated products where merger leads to neither enhancements in quality nor changes in differentiation (such as, through repositioning), the marginal cost reduction necessary to offset the anticompetitive effect is independent of the shape of demand; it depends only on margins and diversion ratios. And, accurate information about these variables may be easier to obtain than information about demand. Werden further argues that, in these circumstances, a very large reduction in marginal cost is generally required to offset the price-increasing effects of an otherwise anticompetitive merger. *See* Gregory J. Werden, *A Robust Test for Consumer Welfare Enhancing Mergers Among Sellers of Differentiated Products*, 44 J. Indus. Econ. 409 (1996).

35 To be sure, this may complicate considerably the analysis of a merger's likely effects under any welfare standard.

36 In addition, marginal cost savings will generally not be achieved immediately, implying a need properly to discount for the more distant effects on price. While even fixed cost savings may take time to materialize, these at least tend to be one-time cost savings, implying less need to discount a stream of future benefits.

tition authorities themselves, and by the courts in attempting to draw what would, under a total welfare standard, be a far less important distinction.

It might be argued that even if the costs of implementing a consumer welfare test are nontrivial, they are necessarily less than the cost of calculating the total welfare effect of a merger. Why? Because calculating total welfare requires the analyst to do all of the above *plus* estimate and add in any fixed or marginal cost savings that will not be passed on to consumers.

While this is correct if one's goal is to calculate precisely a merger's total effect on welfare, it is not correct if one's purpose is instead to determine whether total welfare is likely to increase. There will surely be many situations where the analyst would be able to conclude from the likely magnitude of merger-specific cost savings—whether marginal or fixed—that these benefits to society would exceed any plausible deadweight welfare loss.[37] In such cases, a total welfare standard would likely be far easier than a consumer welfare standard to apply.[38] Moreover, if one is unable to estimate with a reasonably high degree of confidence claimed efficiencies that benefit producers rather than consumers (such as, merger-specific fixed cost savings), it would not be inconsistent with the use of a total welfare standard to evaluate the merger largely on the basis of what one believes likely to be the effect on final consumers alone. A total welfare standard may, therefore, not only be more desirable conceptually, but also less costly to implement.

XI. Welfare Standards and the First Theorem of Antitrust

Finally, one might argue that employing a consumer welfare standard is less costly because it lends itself to ready application of what some might consider the "first theorem of antitrust"—i.e., if consumers like it, the merger is pro-competitive and should be permitted. If they dislike it, the merger is anticompetitive and should be blocked. Because pure fixed cost savings do not translate into lower prices for consumers in the relevant markets impacted by a merger, and

37 Treating as welfare neutral the pure transfer of surplus from consumers to producers from even a modest post-merger price increase—which is what a total welfare standard would do—the deadweight loss from many mergers would often be quite small relative to any significant cost savings. This, even after controlling for the factors that may have biased upwards the estimates contained in Williamson's naïve model. Moreover, as discussed in Roberts and Salop, firm-specific efficiencies generated by merger may, over time, spill over to the market as a whole as other firms in the economy gradually appropriate these benefits for themselves. Gary L. Roberts & Steven C. Salop, *Efficiencies in Dynamic Merger Analysis*, 19 World Competition L. & Econ. Rev. 5 (1996).

38 It is worth recalling also from the discussion in Section III, *supra*, that even mergers likely to satisfy the consumer welfare standard, as currently applied, may result in harm to at least some consumers. Such may be the case, for example, where a merger efficiently leads the combined firm to focus on a single standard or platform (thereby stranding some groups of customers), or where merger-specific efficiencies in some markets are inextricably linked to competitive harm in others.

because a total welfare standard explicitly permits a tradeoff between harms to consumers and (potentially larger) gains to producers, determining from the views of customers whether to challenge a merger provides at most only limited guidance under the total welfare standard.

The case for relying on the views of customers as a simple, shorthand way of determining whether a merger is likely to enhance consumer welfare is not, however, anywhere as strong as some suggest.[39] Indeed, difficulties in applying this proxy are an important reason why competition authorities do not simply poll customers for their bottom-line views on a merger, but dig more deeply into the rationale behind customer views, where possible; examine natural experiments; and rely increasingly on sophisticated empirical techniques—particularly econometric analysis. These efforts tend to belie the claim of some that the first theorem is a surefire and low-cost method of answering our ultimate questions.

Why can one not readily determine from the views of customers whether a merger is likely to satisfy the consumer welfare standard? The reasons are several. For one thing, typically there are many consumers with different demands and tastes. Polling a segment of them, particularly if price discrimination is not feasible, will not necessarily determine whether a merger is likely to produce a reduction in consumer welfare. Second, where the merger threatens potentially to raise price to each consumer by only a relatively small dollar amount, consumers are unlikely to have given serious thought to the question; they will rationally have found it unprofitable to invest in obtaining the information relevant to developing a strong and informed opinion.

Also very important is the fact that many, perhaps even most, mergers that come before competition authorities involve inputs, not final products. The immediate consumers of these inputs, and the ones most frequently quizzed about the merger's likely impact, are not themselves final consumers. Unfortunately, the effects on some, or even all, of these customers can be quite different from the effects on the ultimate, final, consumers to whom they sell. Purchasers of intermediate goods frequently employ different production techniques in turning out competing final goods. To the extent that some of these producers rely less heavily on a particular input than do others, the impact on the former group may be positive even if a merger threatens to raise the incremental costs for that firm and its rivals. In effect, firms that face relatively small cost increases may benefit on net from the fact that consumers shift towards them, and away from competitors whose costs have increased even more.[40]

39 For a fuller discussion of the issues presented in this section, *see* Ken Heyer, *Predicting the Competitive Effects of Mergers by Listening to Customers*, 74 Antitrust L.J. (forthcoming 2006).

40 *See* Joseph Farrell, *Listening to Interested Parties in Antitrust Investigations: Competitors, Customers, Complementors, and Relativity,* 18 Antitrust 64 (Spring 2004). For one famous instance of a strategic effort by some firms to raise their rivals' costs as well as their own, *see* Oliver Williamson, *Wage Rates as a Barrier to Entry: The Pennington Case in Perspective*, 82 Q.J. Econ 85 (1968).

In addition, where final demand is inelastic and pass-through is likely to be nearly complete, intermediate goods customers may (correctly) believe that they will not be very much harmed by even a substantial post-merger increase in the price of what they buy. Final consumers, of course, are unambiguously harmed. Moreover, purchasers of intermediate goods who themselves already have substantial stocks of the input—either warehoused, or incorporated into final products not yet sold—may benefit from the higher incremental costs now faced by all from expanding and/or entering their markets. Again, final consumers would be left worse off, even as some (or even all) intermediate good producers benefit.

Finally, in some circumstances pass-through of a cost increase will be greater than one hundred percent,[41] and economists have shown that, depending on final demand conditions, higher marginal costs may actually increase the profits of intermediate goods customers.[42] The takeaway from all of this is not that the views of customers are irrelevant in determining the likely effects of a merger. Rather, it is that the translation from consumer views to implementation of even a consumer welfare standard is often far from a simple task.[43]

All of this having been said, employing a total welfare standard would not be an easy matter either. Moreover, for reasons I now discuss, strict adherence to a total welfare standard would potentially lead to approval of a number of mergers whose likely effect on consumers is significantly negative.

XII. Distributional Considerations

Distributional considerations raise at least two separate and distinct issues. The first is whether pure transfers among groups in society should be considered in merger policy. Again, the statute is silent on this issue.

A second is whether, even if we were to grant that wealth distribution considerations are an appropriate focus of antitrust policy, this provides clear support for use of a consumer welfare standard. Here we need to ask how confident we are of the presumption that consumers impacted adversely by a merger are less wealthy than owners of the firms that may be achieving cost savings in addition

41 For example, this is an implication of Cournot competition.

42 *See* Sheldon Kimmel, *Effects of Cost Changes on Oligopolists' Profits*, 40 J. Indus. Econ. 441 (1992). For an analysis of circumstances under which intermediate goods customers may be harmed even as consumers benefit, *see* Daniel P. O'Brien & Greg Shaffer, *Bargaining, Bundling, and Clout: The Portfolio Effects of Horizontal Mergers*, 36 Rand J. Econ. 573 (2005).

43 It is worth mentioning also, that while final consumers are an excellent source of information about their own demands they cannot generally be expected to opine intelligently on other factors relevant to whether a merger will prove anticompetitive. For example, it is not at all obvious that individual consumers will have reliable knowledge as to a proposed merger's prospects for generating merger-specific efficiencies, or whether in the face of a possible SSNIP (small but significant and non-transitory increase in price) entry can be expected to be timely, likely, and sufficient.

to the consumer surplus they are diverting. Surely not all mergers are likely to have this effect, and we should therefore consider the potentially enormous costs of calculating in each case the net distributional effect from employing a consumer welfare standard.

To see that anticompetitive mergers need not have an adverse distributional impact, consider a proposed merger of all Mercedes-Benz repair shops in a relevant geographic market. Assume that entry into the market is strictly prohibited; perhaps because of zoning restrictions enacted to prevent noise or congestion, or perhaps explicitly to protect these firms from additional competition. Under this fact pattern, should the antitrust authorities challenge the merger if our best economic analysis concluded that, post-merger the profit-maximizing price for repairing Mercedes-Benz automobiles would rise by 25 percent? Typically, one would think, a challenge would be appropriate, and that authorities should not be required first to determine whether the deadweight loss from a merger-induced price increase would be offset by a socially beneficial wealth transfer from rich automobile owners towards service station owners.[44]

Beyond the question of whether it is desirable in theory to take into account distributional effects,[45] it is not clear that one can at low cost confidently predict even the direction, much less the magnitude, of a merger's distributional consequences. The owners of publicly traded corporations proposing to merge in an effort to capture savings not fully passed on to consumers are quite often an ordinary cross section of Americans, and doubtless include the life savings of retirees, as well as the proverbial widows and orphans.[46] Moreover, many consumers who happen to be poor today will be wealthier tomorrow (and vice versa). Although in extreme cases the distributional consequences of a proposed merger may be relatively clear,[47] in most cases it will be far from obvious (and costly to determine) whether consumers of the merging firms' substitute products are disproportionately in the lower wealth brackets. Determining with any degree of confidence the impact on particular consumer groups of proposed mergers in intermediate goods markets is likely to be especially difficult.

44 If station owners are generally poorer than their customers, should station owner cartels be permitted? Should they be encouraged to form?

45 Arnold Harberger, in a paper, offers a plea for economists to accept as part of their "conventional framework for applied welfare economics," the postulate that "when evaluating the net benefits or costs of a given action (project, program, or policy), the costs and benefits accruing to each member of the relevant group (e.g., a nation) should normally be added without regard to the individual(s) to whom they accrue." Arnold Harberger, *Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay*, 9 J. ECON. LIT. 785 (1971).

46 In addition, producer surplus, when not passed through to consumers in the form of lower prices, may be shared to some extent with suppliers of the firm's inputs—including labor.

47 Ralph Winter offers, as one possible example, a proposed merger of duopoly slumlords.

While a policy requiring calculations of not only the likely price effects, but also the likely distributional consequences of mergers, would doubtless contribute significantly to the wealth of economic consultants and experts, the benefits to society from incurring these costs seem highly questionable. And if we are not even sure what a standard's distributional consequences are, it seems hard to justify use of that standard on the basis of its (unknown) distributional consequences.

It is hardly obvious that a decision on whether to block a merger ought to depend in any way on whether, for example, service station owners are, on average, wealthier than their customers (or vice versa). And even if it were costless to determine whether the consumers or the producers affected by a merger are wealthier, surely questions of wealth distribution are better handled through targeted tax and subsidy programs, rather than via antitrust policy. Arguing that government tax and spend policies are a more appropriate means of dealing with issues of income (and wealth) distribution is not equivalent to proposing that government take on the job of neutralizing, in all cases, the consequences of the occasional price-increasing merger that may come to be permitted under a total welfare standard. The wealth of poorer citizens is regularly affected by any number of broad economic factors—including, for example, adverse changes in the supply and demand for their labor. And yet, economists generally deem it more efficient to deal with broad issues of wealth and income distribution at a macroeconomic level, rather than by interfering regularly with the normal, wealth-maximizing functioning of private markets.[48]

Finally, and importantly, a merger policy that contributes to the overall size and growth of the economy generates larger total wealth; and at least part of the proceeds can, if society wishes, be used by fiscal authorities to aid, through targeted taxation and spending programs, those deemed to be most needy, or otherwise most deserving.

XIII. Some Additional Considerations

A. THE COSTS OF RENT-SEEKING BEHAVIOR

Economists have long known that competition can be wasteful at times. That is to say, one really can have too much of a good thing. What does it mean to say, as a matter of economics, that something is "excessive"? Conduct can sensibly be defined as excessive when it is being engaged in past the point at which, from the standpoint of society as a whole, the value it adds is greater than its cost.[49]

48 Indeed, the absolute adverse impact on an individual poor person of a merger permitted under the total welfare standard, but not permitted under the consumer welfare standard, seems likely to be quite modest in comparison with other factors impacting on a poor person's wealth. Even a 20 percent increase in the price paid for a $200 item amounts to a total of only $40.

49 The maintained assumption is that rational actors engage only in conduct whose benefits to them are expected to exceed the costs to them.

Pollution is one commonly used example of where negative externalities may be imposed by manufacturers on residents who are not compensated for harms they suffer. In such cases, it is easy to show that an activity—in this case, production by the manufacturers—may well proceed past the point at which its net value to the economy is positive.

Competition itself can in some cases create a negative externality—an uncompensated harm—that results in excessive entry or excessive product differentiation. The negative externality in these cases is felt not by consumers, but by incumbent producers. Nevertheless, it can result in entry and competition that, while beneficial to consumers in the market, is wasteful to the economy as a whole.[50]

Consider, for example, a market in which incumbent firms are earning significant margins and positive margins will remain even following competitive entry. In such cases, potential entrants will find that the costs of entering will be partly covered by revenues on business that the entrant steals" from incumbents. To the extent that business can be stolen without a substantial cut in price (or improvement in quality), the benefits to consumers may fall far short of the associated costs, in particular, the fixed costs of entry.[51]

Part of what makes entry into a market profitable can be the margins that are shifted from incumbents to the entrant. From the standpoint of society as a whole, however, these are merely transfers from one firm to another. Nevertheless, the prospect of capturing these margins can in some instances make entry profitable, despite the fact that, from the standpoint of the economy

50 Antitrust scholars, most notably Richard Posner, have argued that the drive to obtain market power can transform rents into costs. Richard Posner, *The Social Costs of Monopoly and Regulation*, 83 J. Pol. Econ. 807 (1975). Whatever the power of this logic in general, the argument appears to have questionable applicability to mergers in particular. As Williamson summarizes after responding to this argument in some detail:

> Plausibility standards plainly vary. Those who are easily persuaded that managers enjoy extensive insularity [from stockholder control], that managers fully credentialize on the basis of low probability events, and that the marginal utility of money is fairly constant will conclude that exhaustive ex ante rent transformation occurs in the merger context, as required by Posner's theory. On the other hand, those who are skeptical of any of these assumptions will conclude that rent transformation will be incomplete. As for myself, I believe that the insularity assumption is the most doubtful. Absent this assumption, the entire argument collapses.

See Williamson, *supra* n. 8, at 8.

51 It is hard to imagine there being a serious competitive problem with a merger of two or more of the many high-price coffee shops located within walking distance of my office in downtown Washington, DC, and I admit to wondering whether some of the fixed costs incurred from all of this entry might not be wasteful from the standpoint of society as a whole.

as a whole, entry does not pass a cost-benefit test.[52] In addition to investments in entry, investments in product differentiation can, for similar business-stealing reasons, be wasteful as well.[53]

These considerations suggest a theoretical case for actually prohibiting entry at times, though it is worth mentioning also that, in theory, markets can suffer equally from insufficient entry.[54] The antitrust laws are not employed to keep out competitive entry, nor would many propose that they be used for that purpose.[55] Consumers in markets where even excess entry occurs clearly benefit. The benefits include not only lower prices, but also the prospect for highly valued variety and better products, not to mention the incentive that the entry threat provides for incumbents to minimize costs and identify and satisfy consumer demands. Determining ex ante whether the costs of entry are likely to exceed its many potential benefits seems clearly too daunting a task to take on.[56] As with antitrust's per se prohibition of naked price fixing, courts sensibly eschew case-by-case analysis in circumstances where the expected cost exceeds the expected benefit.

Given how costly—and potentially harmful—it may be to implement a total welfare standard that takes these theoretical possibilities into account, one could argue that it would be best, all things considered, simply to stick with a consumer

52 *See* Gregory N. Mankiw & Michael D. Whinston, *Free Entry and Social Inefficiency*, 17 Rand J. Econ. 48, 55-57 (1986), which discusses various entry biases and examines the tendency toward excessive entry in homogeneous product markets. *See also,* Michael Spence, *Product Selection, Fixed Costs, and Monopolistic Competition*, 43 Rev. Econ. Stud. 217 (1976) and A.K. Dixit & J.E. Stiglitz, *Monopolistic Competition and Optimal Product Diversity*, 67 Am. Econ. Rev. 297 (1977), which discuss the possibility of free entry resulting in too little entry relative to the social optimum in a monopolistically competitive market. Finally, *see* Chiang-Tai Hsieh & Enrico Moretti, *Can Free Entry Be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry*, 111 J. Pol. Econ. 1076 (2003), which attempts empirically to estimate the cost of excess entry into the Real Estate Industry.

53 *See* Michael Spence, *Product Differentiation and Welfare*, 66 Am. Econ. Rev. 407 (1976). *See also,* Steven C. Salop & Jeffrey M. Perloff, *Firm-Specific Information, Product Differentiation, and Industry Equilibrium*, 38 Oxford Econ. Papers 184 (1986) and Dixit & Stiglitz, *id*.

54 *See*, e.g., Chaim Fershtman & Ariel Pakes*, A Dynamic Oligopoly with Collusion and Price Wars*, 31 Rand J. Econ. 207 (2000). There, the authors show that where fixed costs are significant, marginal costs are low, and post-entry pricing is Bertrand, total welfare can be enhanced by permitting post-entry collusion. In certain circumstances total welfare can be enhanced also by having the government subsidize entry or limit the freedom of an incumbent to respond to an entrant's lower prices. *See, e.g.*, Aaron S. Edlin & Joseph Farrell, The American Airlines Case: A Chance to Clarify Predation Policy (Berkeley Competition Policy Center, Working Paper No. CPC02-33, Nov. 2002).

55 The argument that entry should be prohibited on grounds that its costs exceed its benefits has, however, been employed in regulatory proceedings—typically by incumbents claiming to be natural monopolies. *See*, e.g., Paul W. MacAvoy, Daniel F. Spulber & Bruce E. Stangle, *Is Competitive Entry Free? Bypass and Partial Deregulation in Natural Gas Markets*, 6 Yale J. on Reg. 209 (1989).

56 Inefficiencies associated with requirements that potential entrants first obtain regulatory approval have been widely documented. For one discussion of the harmful effects of entry regulation in the trucking industry, *see* Thomas Gale Moore, *Trucking Deregulation,* The Concise Encyclopedia of Economics, *available at* http://www.econlib.org/library/Enc/TruckingDeregulation.html.

welfare standard. And yet, the difficulties of administering a total welfare standard do not seem so large, or so welfare-standard specific, that they offset the strong case for using what is otherwise a far more desirable welfare criterion. Determining whether a merger is likely to raise total welfare requires an estimate of the merger's effect on allocative efficiency via the ability of firms profitably to raise price (or otherwise harm consumers of their products) and the cost savings, if any, that are specific to the merger. In principle, the analyst needs to evaluate these two effects and calculate their net welfare effect under either a consumer, or a total, welfare standard.[57]

B. TRANSPARENCY

One might agree with the policy proposal put forward in this paper—that competition authorities should be employing a total welfare standard—yet also feel that if a change is to be made, it would be best that it not be made very publicly or transparently. A formal policy change of this type could well generate considerable flak and controversy. Might it be best, all things considered, to implement it through the use of prosecutorial discretion not to bring wrong or difficult cases?

Apart from objections in principle to the idea that open debate should be discouraged because the public would simply not support such a change, there are a number of practical adverse consequences from being less than candid about the standard that the competition agencies will apply. For one thing, if parties do not know that certain types of efficiencies are going to be credited by the authorities, they are far less likely to present the types of evidence and analysis required for those efficiencies to be credited. For another, in some circumstances the government may elect to file a case and the defendants will present to the court an efficiency justification that would in theory be credited under a total welfare standard, but that would not be credited under a consumer welfare standard. If total welfare is not the standard officially employed by antitrust officials, one can expect government prosecutors and their experts to argue strenuously that benefits unlikely to be fully passed through to consumers in the relevant markets of concern do not count. This would be, in my view, unfortunate, and may help lead a court to wrongly rule in favor of the plaintiff, or at least to rule in favor of the plaintiff for the wrong reasons. It may also complicate the ability of authorities to implicitly credit savings in producer surplus and appropriately to exercise prosecutorial discretion in the future.

57 An optimal policy under conditions of uncertainty would take into account the costs of both type I and type II errors. Decisions would be made based on the expected value of a merger's effect on welfare, not simply a point estimate. *See* Ken Heyer, *A World of Uncertainty: Economics and the Globalization of Antitrust*, 72 ANTITRUST L.J. 375 (2005).

XIV. Conclusions

More and more antitrust practitioners, both in the United States and abroad, have expressed an interest in incorporating the effects of merger-generated efficiencies into merger analysis. The analysis in this article argues that if we are going to do it,[58] we might as well do it right, and that use of a total welfare standard appears to be not only the theoretically best standard to employ, but also one that can be employed with no significant increase in administrative costs.

And there is a somewhat broader point worth making. Use of a consumer welfare standard in antitrust inherently casts consumers as those who count, and producers as those who do not. This is unfortunate, in my view, for reasons that go well beyond those laid out in this article. For one thing, producers, it bears remembering, happen also to be consumers. Indeed, they have been seen shopping after work, and on their days off. Moreover, there seems good reason to value the welfare of those who produce what we consume as highly as those who do the consuming.

Defense of a welfare standard that ignores the welfare of producers contributes to a perception that producers exist only to benefit those to whom they sell, and that the welfare of those actually doing the producing—often at significant cost and risk—is itself of no value. It thus contributes to a mindset that favors all manner of efficiency-reducing policies to benefit select groups of consumers at the expense of producers, despite the negative effect such policies have on overall economic welfare, certainly in the long run, and often in the short run as well.

> DEFENSE OF A WELFARE STANDARD THAT IGNORES THE WELFARE OF PRODUCERS CONTRIBUTES TO A PERCEPTION THAT PRODUCERS EXIST ONLY TO BENEFIT THOSE TO WHOM THEY SELL, AND THAT THE WELFARE OF THOSE ACTUALLY DOING THE PRODUCING . . . IS ITSELF OF NO VALUE. IT THUS CONTRIBUTES TO A MINDSET THAT FAVORS ALL MANNER OF EFFICIENCY-REDUCING POLICIES TO BENEFIT SELECT GROUPS OF CONSUMERS AT THE EXPENSE OF PRODUCERS, DESPITE THE NEGATIVE EFFECT SUCH POLICIES HAVE ON OVERALL ECONOMIC WELFARE . . .

Quite apart from whatever economic benefits would result from putting antitrust policy on a more economically sensible footing, these considerations argue in favor of doing so as well.

58 Some serious antitrust scholars, including Richard Posner and Robert Bork, have concluded that explicit case-by-case consideration of merger-specific efficiencies, and by implication the use of an efficiencies defense, is simply too difficult to conduct in practice and should therefore not be a formal part of merger analysis and litigation. *See* RICHARD A. POSNER, ANTITRUST LAW (2nd ed. 2001) and ROBERT H. BORK, THE ANTITRUST PARADOX (1978).

Volume 2 | Number 2 | Autumn 2006

Competition Policy International

Some Reflections on the European Commission's State Aid Policy

Philip Lowe

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Some Reflections on the European Commission's State Aid Policy

Philip Lowe

This article first sets out the main economic and legal requirements for a competition enforcement system to be effective, as well as the main building blocks upon which the effectiveness of any such system depends. It then explains how the Commission is currently implementing these requirements in its reform of the State aid control system. It describes how Commissioner Kroes launched for that purpose the State Aid Action Plan and sets out how, in the context of this plan, the political and economic objective of promoting "less and better targeted aid" must be complemented by improvements as regards both procedure and the economic assessment of aid. It goes on to discuss the economic foundations of the new approach and in particular the so-called balancing test which will underlie both the analysis in individual cases and new legislation. It then explains how the reform objectives have already been implemented in a number of recent legislative or quasi-legislative texts. The article concludes by describing the remaining challenges, on both substance and procedure, in order for the State aid reform to constitute one of the cornerstones of the Commission's strategy for growth and jobs.

The author is Director General of the Directorate-General for Competition at the European Commission.
The views expressed are personal to the author and do not necessarily reflect those of the European Commission. The author would like to thank Thomas Deisenhofer and Harold Nyssens for assisting him in preparing this paper.

I. Why We Launched a State Aid Reform

A. BASELINE REQUIREMENTS FOR AN EFFECTIVE COMPETITION LAW ENFORCEMENT SYSTEM

When designing an effective competition law enforcement system—whether for State aid control, merger control, or antitrust—there are a number of baseline requirements.

First and most importantly, rules and individual enforcement actions must be based on sound law and economics. On the legal side, enforcement must be—and widely seen to be—subject to the rule of law, due process requirements, and effective judicial control. On the economic side, the long-term legitimacy of any competition enforcement system rests on the economic story which it tells. Any competition enforcer should be able to explain why and how its enforcement actions contribute to the wider public interest.

Second, the enforcement system must be designed in a way that guarantees coherence and predictability. Coherence of enforcement ensures equal treatment of businesses. Predictability in the enforcement allows businesses to plan for the long term. There may sometimes be a trade-off between predictability of enforcement and the need to deal with each case on its merits. The objective must nonetheless be to guarantee as much predictability as possible.

A third baseline requirement is that the system should allow the enforcement agency to concentrate its limited resources on specific priorities. The enforcer should be able to determine those priorities on the basis of the expected direct and demonstrative effects of decisions. The system should make it possible to concentrate resources on the potentially most harmful conduct and on precedent-setting cases.

Fourth, as to the length of investigation procedures, any effective enforcement system must enable a public agency to take decisions in a timeframe relevant either to the business transaction or public policy initiative concerned. Decisions must therefore be taken when they still have an impact on the economic effects which they aim to address. Precedents must also be set at a moment when they still have the intended wider policy impact.

Last but not least, enforcement must always go hand-in-hand with an effective competition advocacy policy. The ultimate goal of competition policy is to make markets work better in the interest of consumers and of businesses. Only where enforcement and advocacy are both used in parallel in a mutually reinforcing way can this objective be achieved.

B. BUILDING BLOCKS THAT ALLOW A COMPETITION ENFORCEMENT SYSTEM TO FULFILL THOSE BASELINE REQUIREMENTS

To fulfill the requirements set out above, a competition system should contain a number of building blocks.

First, enforcement actions must be based on clear and agreed enforcement objectives. Otherwise we cannot guarantee any coherent and predictable policy. In the field of antitrust and mergers, for example, we explicitly pursue a consumer welfare objective.

Second, rules and individual enforcement actions must be based on a clear and economically sound agreed assessment methodology. Similar cases must be assessed according to the same tests. The assessment methodology should be used not only for the assessment of individual cases but also for the design of general assessment rules. Without a clear assessment methodology, it is difficult to ensure coherence and predictability.

Third, the enforcement architecture (i.e., all substantive, procedural, and internal rules taken together), must allow the agency to set the correct priorities. Block exemptions, de minimis rules, notification thresholds, as well as filters in substantive rules or procedural rules, should allow us to deal quickly, and with limited resources, with unimportant or easy cases. For the remaining cases, there should be the right mix between rules ensuring predictability of outcome (per se rules)[1] and rules ensuring predictability of assessment methods (effects-based analysis).

Fourth, as regards the timing of procedures, it is necessary to have procedural rules and internal best practices which ensure a rapid investigation and rapid internal decision making.

Fifth, the system must provide for tools such as explanatory guidelines which ensure the transparency of the applicable rules and the enforcement policy.

C. IS THERE ROOM FOR IMPROVEMENT IN THE AREA OF STATE AID CONTROL?

The quality of the EC State aid control system has been substantially improved over the last decade. Just to name the most significant achievements, in 1999, the Council adopted the first procedural regulation[2] in the field of State aid. This breakthrough on the procedural front was soon followed by the first de minimis

1 *See infra,* Section III.B.

2 Council Regulation 659/1999 of 22 March 1999 laying down detailed rules for the application of Article 93 (now Art. 88) of the EC Treaty, 1999 O.J. (L 83) 1.

Regulation,[3] and by a number of block exemption Regulations[4] that have significantly reduced the number of notifications of unproblematic cases and thereby reduced red tape. Texts like the *R&D Guidelines* or the *Regional Aid Guidelines* have been constantly adapted to economic and technical progress, enlargement and ever-deeper economic integration within the internal market.[5] Those texts have been designed on the basis of extensive consultations of stakeholders and are accepted to be based on sound policy and economic principles. Nonetheless, there is still room for further improvements.

One key area is the setting of enforcement priorities. Existing jurisprudence on the application of the concepts of distortion of competition and effect on trade in Article 87(1) of the EC Treaty[6] does not allow these concepts to act as a filter for prioritizing workload. Measures with little or no real economic impact on cross-border trade or competition remain within the scope of the notification obligation.

One consequence of this is that unimportant complaints add substantially to the background tasks of the European Commission (the Commission). From the perspective of the individual complainant, every complaint is of course important. As a public enforcement authority we are however responsible for safeguarding the interests of *all* European citizens and are given only limited resources to do so. Ideally an authority should thus be able to choose the complaints that it wishes to pursue on the basis of the interest that they present for the European Community as a whole.

As regards the notification obligation, the block exemption regulations have significantly reduced the number of notifications. However, the Directorate-General for Competition (DG COMP) still receives on average 322 notifications per year.[7] Similar to what happened in the antitrust field before modernization,

3 Commission Regulation 69/2001 of 12 January 2001 on the application of Articles 87 and 88 of the EC Treaty to De minimis aid, 2001 O.J. (L 10) 30.

4 Commission Regulation 68/2001 of 12 January 2001 on the application of Articles 87 and 88 of the EC Treaty to training aid, 2001 O.J. (L 10) 20; Commission Regulation 70/2001 of 12 January 2001 on the application of Articles 87 and 88 of the EC Treaty to State aid to small and medium-sized enterprises, 2001 O.J. (L 10) 33; Commission Regulation 2204/2002 of 12 December 2002 on the application of Articles 87 and 88 of the EC Treaty to State aid for employment, 2002 O.J. (L 337) 3.

5 Community guidelines on State aid for rescuing and restructuring firms in difficulty, 2004, O.J. (C 244) 2 and Guidelines on national regional aid for 2007-2013, 2006, O.J. (C 54) 13.

6 Article 87(1) of the EC Treaty states:

> Save as otherwise provided in this Treaty, any aid granted by a Member State or through State resources in any form whatsoever which distorts or threatens to distort competition by favouring certain undertakings or the production of certain goods shall, in so far as it affects trade between Member States, be incompatible with the common market.

7 Calculation reflects the average number of notifications from 2001 to 2005.

resources have to be used as a matter of priority for notified cases, which are often not the most distortive ones. As a consequence, non-notified new aid or distortive existing aid measures are often only discovered and investigated if there is a complainant and if there are residual resources available.

Another area for possible improvement is the average length of our procedures. The approval of the large number of straightforward notified aid cases takes on average about five to six months. In cases involving the opening of a formal investigation, the time taken to reach a final decision in notified aid cases is on average twenty months and in non-notified aid cases on average thirty months. The main reason for such long delays is the time it takes for Member States to respond to requests for information. The provisions of the current procedural regulation[8] limit the Commission's possibilities for speeding up procedures. In such a context, some Member States—in particular those with complex internal decision-making processes—regularly exceed the indicated time limits, which we set. In addition, under the current procedural regime, the Commission only has the power to formally request information from the Member States even if the information ultimately needs to come from the companies concerned. If the Member State concerned is slow in providing the information requested, the Commission investigation can be delayed substantially.

A third area for possible improvement relates to the type of information that we gather to make our decisions. As already indicated, under the current rules, the aid-granting Member State is the main—and sometimes the only—source of information at the basis of a Commission investigation. This may lead to a lack of information regarding the impact of the aid on competition and trade, mainly because the national authority concerned does not normally have that sort of information available itself. This is also the reason why the notification forms currently used request either little or no information on the affected markets and competitors.

A fourth and final area for possible improvement is the notification obligation. It is clear that the fundamental obligation, provided by the treaty, for Member States to notify new aid before its implementation, is still not respected in a large number of cases. Between 2001 and 2005, DG COMP has had to open procedures in 259 cases that had not been notified. The current procedural framework, or at least the way in which it is interpreted and applied, does not appear always to ensure equal treatment of Member States. Thus, illegal State aid continues to disrupt market incentives and flows of trade within the European Community.

8 *See*, in particular, Council Regulation 659/1999, art. 5 & 10.

To sum up, the long-term credibility of State aid control in a globalizing world requires decisions to be taken more rapidly and to better reflect economic reality in a systematic manner. The legitimacy of State aid control rests on the presumption that we will effectively address those forms of aid which have predominantly negative effects on competition and on the functioning of the internal market. These concerns led Commissioner Kroes to decide on a wide-ranging reform of State aid policy, dubbed the "State Aid Action Plan."

II. The State Aid Action Plan

A. A GLOBAL RESPONSE TO THE CHALLENGES OF STATE AID POLICY

On the initiative of Commissioner Kroes, the Commission launched the State Aid Action Plan (SAAP)[9] on June 7, 2005.[10] The SAAP presents an indicative roadmap for State aid reform for the period 2005 to 2009. It is being implemented in close cooperation with Member States and other stakeholders, as is shown by the numerous consultations organized in the context of this project.[11]

THE LONG-TERM CREDIBILITY OF STATE AID CONTROL IN A GLOBALIZING WORLD REQUIRES DECISIONS TO BE TAKEN MORE RAPIDLY AND TO BETTER REFLECT ECONOMIC REALITY IN A SYSTEMATIC MANNER.

The Commission was already committed to reviewing a number of legal instruments such as guidelines which were due to lapse during 2005 and 2006. Instead of proposing piecemeal improvements which would increase the risk of inconsistency between the various instruments, Commissioner Kroes decided to address all aspects of the reform in a comprehensive and consistent manner. This opportunity was seized on to clarify the objectives of State aid policy, and to call for a new partnership with Member States to make it a success.

The SAAP was adopted in June 2005 as a consultation document intended to launch a political debate about the future of State aid control. We are trying to involve stakeholders at every stage of our reform project, just as we did with the modernization of antitrust and the reforms to merger control.[12]

---

9 European Commission, State Aid Action Plan: Less and better targeted State aid: a roadmap for State aid reform 2005–2009, COM(05)107 final, *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/saap_en.pdf.

10 *See* European Commission, Press Release No. 680, State Aid: Commission outlines comprehensive five year reform of state aid policy to promote growth, jobs and cohesion (2005).

11 See European Commission, State Aid Reform, at http://ec.europa.eu/comm/competition/state_aid/overview/sar.html (visited Aug. 21, 2006).

12 Green Paper on the Review of Council Regulation (EEC) 4064/89 COM(01)745 final at 6; Council Regulation 1/2003 of 16 December 2002 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty, 2003 O.J. (C 243) 3.

The SAAP looks at the underlying philosophy of State aid control, and presents the challenges ahead and the guiding principles that we intend to implement to address these challenges. It outlines four guiding principles for the reform of State aid policy:

1. less and better targeted State aid;
2. better procedures and administration;
3. a shared responsibility between the Commission and Member States; and,
4. a refined economic approach.

These principles will establish a consistent policy that will be simpler in its overall architecture and easier to grasp for stakeholders. The second section of the SAAP reviews the individual areas that will be affected by the reform and sets out proposed changes based on these four principles.

B. THE FOUR PILLARS OF THE STATE AID REFORM

1. Less and Better Targeted State Aid

The European Council of March 2005 restated the aim of "less and better targeted State aid."[13] When governments decide to intervene in the market they should not use taxpayers' money for piecemeal support of undertakings in difficulty or branches of industry in steady decline. The governments that met in the European Council agreed instead to redirect their aid towards horizontal objectives of common interest.

Some forms of State aid clearly distort competition and prevent the market from providing the right incentives for business to become more efficient and to innovate to the final benefit of customers. This should not, however, be misunderstood as a drive against state intervention in whatever form. Many areas of government activity, for instance, fall outside the sphere of State aid control. State intervention in sectors such as education, security, or social security, for example, often does not constitute State aid as the activities at stake are may not be economic in nature. Aid can, of course, also be granted to individuals, for instance, for social reasons, without their being considered as undertakings falling under the scope of State aid rules.

In areas where the State aid rules are relevant, the idea of less State aid means that governments have agreed to reduce to a minimum those support measures not targeted at commonly agreed Lisbon objectives. That it is the governments of the Member States themselves that have set and agreed this objective is

13 *See* Presidency conclusions of the European Council Brussels, 22 and 23 March 2005, *available at* http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/ec/84335.pdf.

important: reducing and better targeting State aid is not something that can be achieved by Commission intervention alone. Unfortunately, the day-to-day actions of Member States are all too often at variance with their stated policies in the European Council. Member States regularly disregard notification obligations and provide illegal and incompatible aid, as already indicated above.

Whether less aid will indeed be provided in the next years will, therefore, largely depend on the Member States' commitment and discipline. The Commission intends to do whatever it can to help, particularly by facilitating the possibilities of granting aid that works in favor of the Lisbon objectives. We are starting to make progress: the "State Aid Scoreboard" shows that increasing amounts of aid are being redirected towards horizontal objectives.[14]

2. More Effective Procedures, Better Enforcement, Higher Predictability, and Enhanced Transparency

The SAAP also includes proposals to improve the efficiency of State aid procedures and to speed up decision making. Better procedures and administration means focusing on the crucial cases, and acting swiftly in those cases in a transparent manner. We need to make State aid rules simpler, clearer, and more user-friendly. A first strand of action in this respect is to provide for larger State aid areas to be covered by so-called block exemption regulations, which authorize the granting of aid without the need for notification to the Commission. With more Lisbon-targeted State aid covered by such regulations, the Commission intends to both reduce the administrative burden on Member States while at the same time freeing up its own resources to focus on the cases that are most distortive of competition. Our intention is to consolidate the many existing regulations into one overarching regulation, so that coherence is increased and the architecture of State aid policy is easier to grasp for all involved.

In addition, the Commission intends, as a first stage, to issue best practices guidelines to streamline individual treatment of the cases under the current legal regime. At a later stage, it may issue proposals to the Member States in order to revise the procedural regulation[15] with a view to improving the speed of decision making and to increase deterrence mechanisms in order to tackle State aid granted illegally.

Finally, to increase transparency, we intend to engage in more advocacy to enhance the overall public awareness of State aid policy. A specific network is being set up with Member States in order to enhance the policy dialogue. Moreover, we will try to intensify the involvement of national courts, especially regarding the treatment of illegal aid granted in violation of the notification

14 See EUROPEAN COMMISSION, STUDIES AND REPORTS, *at* http://ec.europa.eu/comm/competition/state_aid/overview/studies.html (visited Aug. 21, 2006).

15 *See infra,* Section V. Conclusions and Next Steps.

obligation. After a recent study realized by scholars and practitioners about the way in which State aid rules are applied in national jurisdictions,[16] we intend to issue a revised notice on cooperation between national courts and the Commission.[17]

3. A Shared Responsibility Between the Commission and Member States

The Commission cannot improve State aid practice without the effective support of Member States: their commitment to comply with their obligations to notify State aid and to provide the market information required for the Commission to realize its assessment is crucial if we are to achieve "less and better targeted aid" as requested by the European Council. On the other hand, the Commission's responsibility lies in designing new instruments to support Member States in granting clearly compatible aid with the minimum of red tape. The de minimis rule and the expanded general block exemption, as well as more pragmatic guidelines, which stick more closely than before to business reality—such as the recently adopted guidelines on risk capital[18]—will play a central role in this respect. In such a revised environment, Member States can then more easily design aid measures fulfilling the Lisbon agenda.

4. Refined Economic Approach

Proper economic analysis is at the heart of all competition policy. By refining the economic approach underlying State aid policy, the Commission can better balance the economic effects of aid with the common interest objectives such as regional cohesion or environmental protection. Whereas an assessment of the negative effects of aid has much in common with the analysis in antitrust and mergers, the assessment of the positive effects raises questions as to the equity and efficiency benefits of the aid. The latter aspect is most clearly spelled out when an aid measure addresses a market failure, which hampers the market to deliver a Pareto-optimal output. The following section discusses the refined economic approach in more detail.

16 Jones Day, Lovells, Allen & Overy, Study on the Enforcement of State Aid Law at National Level (2006), *available at* http://ec.europa.eu/comm/competition/state_aid/overview/studies.html.

17 Notice on cooperation between national courts and the Commission in the State Aid field, 1995 O.J. (C 312) 8.

18 *See* European Commission, Press Release No. 1015, State Aid: Commission adopts Guidelines on state aid to support risk capital investments in SMEs (2006).

III. Sound Economic Foundations: The Refined Economic Approach

A. THE CONTRIBUTION OF ECONOMICS TO STATE AID

Economic analysis is not new in the field of State aid. State aid rules are already based on economics, and economic concepts—like market failures or incentive effect[19]—are mentioned in our guidelines. The Commission has already applied a robust economic analysis not only in the important areas of rescue and restructuring, the multi-sectoral framework and research and development (R&D) but also in other individual cases.[20] However, the common economic concepts and methodologies underlying the largely form-based State aid rules have, in the past, not always been adequately spelled out, which in turn may limit the clarity and predictability of State aid policy.

The traditional per se or form-based approach to rule-making typically relies on quantitative thresholds and limitative lists of cumulative conditions which attach automatic legal consequences to the fulfilment of these conditions. Under such an approach, legal consequences are thus triggered independently of the economic effects of a conduct or measure. A more effects-based approach to rule-making seeks to provide a methodology on how to establish, in an individual case, the actual or likely economic effects of a measure or conduct on the markets in order to trigger the legal consequences foreseen by the rule. The latter approach typically relies on explanations of the objectives of the rules and their underlying economic concerns, the different steps and elements of the tests to be applied and the positive and negative assessment criteria or presumptions which need to be taken into account. Under such an approach, the rules also set out how to balance these different criteria in order to reach a final assessment and decision.

In analyzing the compatibility of State aid under Article 87 (3) EC, the actual or potential economic effects of State aid should be assessed. A more refined economic assessment of State aid is intended to improve the analysis of both the negative and positive economic impact of State aid and will formalize the balancing exercise of those different effects. Analyzing the effects of State aid requires an understanding of how economic behavior, and consequently the market equilibrium, may be affected—positively and negatively—by a given aid measure.

19 *See, especially,* State Aid and Risk Capital, 2001 O.J. (C 235) 3; EUROPEAN COMMISSION, DRAFT COMMUNITY FRAMEWORK FOR STATE AID FOR RESEARCH AND DEVELOPMENT AND INNOVATION, *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/rdi_frame_en.pdf; and Community Guidelines for State Aid in Environmental Protection, 2001 O.J. (C 37) 3.

20 *See, e.g.*, Commission decision of 9 November 2005 on the State aid which the Federal Republic of Germany has implemented for the introduction of digital terrestrial television (DVB-T) in Berlin Brandenburg, 2006 O.J. (L 200) 14.

There are many situations where the key assumptions of the basic model of competitive markets are not met. Imperfect competition, coordination problems, incomplete information, externalities, and public goods, all present hypotheses where the market fails in its role of producing economic efficiency. Economists refer to these problems as "market failures," situations where a market fails to produce efficient outcomes. Looking at such market failures provides guidance as to the potential for government intervention. In particular, State aid may be an appropriate response to a market failure, provided its benefits outweigh its negative impact on competition and trade.

To do this balancing—traditionally dubbed the "balancing test"—it is necessary to identify—and as far as possible to measure—the positive and negative aspects of the aid. In the fields of mergers and antitrust, the negative and positive effects are assessed essentially on the basis of the benchmark of the consumer welfare standard.[21] It is, however, not possible to transpose this consumer welfare standard directly to the world of State aid, not least because State aid can be justified on the basis of non-economic grounds such as social or regional cohesion which consumer welfare does not measure. To that extent, the correct welfare standard for State aid policy—expressed in economic terms—appears to be the social welfare of the European Union, which is equivalent to the notion of common interest found in Article 87(3) of the Treaty.

Social welfare takes into account the wellbeing of all citizens and includes their appreciation of how welfare is divided among them. Therefore, as a standard, social welfare of the European Union also requires a degree of political understanding, as citizens' preferences may evolve over time. Indeed, State aid can affect both the way in which the economic pie is made larger by a given policy (efficiency) and how the pie is then divided between citizens (equity).

In order to assess the economic impact of the aid, the basic methodology is one of ex ante counterfactual—comparing what would happen to the market with State aid with what would be likely to happen without it.

B. THE BALANCING TEST: THE CORNERSTONE OF COMPATIBILITY ASSESSMENT

In the perspective of this welfare standard, the SAAP proposes to formalize the balancing test in three steps. The first two steps address the positive effects of State aid, and the third the negative effects and resulting balancing of the positive and negative:

21 *See*, *e.g.*, Guidelines on the assessment of horizontal mergers under the Council Regulation on the control of concentrations between undertakings, 2004 O.J. (C 31) 5 and DG Competition, Discussion Paper on the application of Article 82 of the Treaty to exclusionary abuses (19 December 2005), *available at* http://ec.europa.eu/comm/competition/antitrust/others/discpaper2005.pdf.

1. Is the aid measure aimed at a well-defined objective of common interest? (e.g., growth, employment, cohesion, environment)

2. Is the aid well-designed to deliver on this objective? (i.e., does the proposed aid address the market failure or other objective?)

   a. Is State aid an appropriate policy instrument?

   b. Is there an incentive effect? (i.e., does the aid change the behavior of firms?)

   c. Is the aid measure proportional? (i.e., could the same change in behavior be obtained with less aid)

3. Are the distortions of competition and effect on trade limited so that the overall balance is positive?

Clarifying the methodology for assessing an aid in this formalized manner will bring a number of benefits.

First, it should lead to more effective enforcement, through a more rigorous assessment of aid in individual cases. In economic jargon, a refined economic approach will reduce the likelihood of type I or type II errors—often referred to more colloquially as false positives and false negatives. In an exclusively form-based system, such errors are, by the very nature of the system, more frequent. Coming more often to the right result should strengthen the acceptance by stakeholders of decisions adopted by the Commission.

Second, this approach should lead to better State aid legislation. Indeed, an economically sound, coherent, clear and transparent legal framework can only be drawn up, presented to stakeholders, and enforced on the basis of a better understanding, by the Commission, of the economic effects which different types of aid have in different factual settings.

Third, the approach should yield better targeted aid. Under a more refined economic approach, the effectiveness of aid to achieve objectives of common interest will be subject to closer analysis, especially in large cases. A better understanding of the mechanisms underlying the effectiveness of aid will in turn help Member States to devise more effective and better targeted aid. A more effects-based economic approach is thus also essential to achieve the Lisbon objectives.

Fourth, we should be put in a position to realize a better prioritization of enforcement action. The new methodology should allow us to focus our resources on the potentially most distortive aid measures.

A greater focus on the economic approach to assessing State aid cases does not mean that State aid control will become more uncertain. The opposite is more likely to be the case. From the perspective of someone trying to address a genuine problem in an effective and efficient way, pure per se rules, may occasionally look capricious and arbitrary, and dealing with such rules can be frustrating and create disillusion with the system as a whole. At the same time, refining the economics underlying State aid policy does not mean that per se rules disappear entirely to be replaced by pure case-by-case analysis and never-ending economic assessment. As we have found in antitrust and mergers, and as we are already demonstrating in the implementation measures discussed in more detail below, it is perfectly possible to combine clear rules with economic analysis—the safe harbor of the risk capital guidelines, for instance, are ample proof of that.

The challenge is to find the right balance between per se rules based on established and tested economic criteria, and a full effects-based assessment appropriate for selected individual cases. Thus, the more refined economic approach must include the continued use of per se rules where clear and simple rules are required, provided only that they are based on past empirical evidence and tested periodically against economic realities. Future soft law State aid rules, like guidelines and frameworks, should however also leave ample room for individual assessment of aid measures on the basis of their effects.

IV. Implementation

A. ENDORSEMENT BY STAKEHOLDERS

The public consultation showed a broad degree of support for the reform: more than 130 submissions were received, and the European Parliament and the European Economic and Social Committee both issued reports.[22] The Commission has already taken a number of proposals contained in these submissions into account in its ongoing implementation, and will continue to consult both Member States and other stakeholders widely on the different subprojects which constitute the SAAP.

B. IMPLEMENTATION OF THE STATE AID ACTION PLAN

The year 2005 saw the adoption of a new package of rules relating to services of general economic interest, of guidelines on regional aid, and the first proposals for aid to innovation. By the end of 2006, the Commission intends to deliver two essential Lisbon related instruments: the new framework on R&D and innovation (R&D&I) and the new risk capital guidelines. A revised version of the de minimis Regulation will also be adopted, which is essential to reduce the regulatory burden on Member States. It will incidentally also allow the Commission to focus its attention on those aid measures which have the greatest potential to damage competition. All of these measures should support the action of Member States towards making Europe a more attractive place to invest and to do business.

22 See EUROPEAN COMMISSION, RESULTS OF THE CONSULTATION ON THE STATE AID ACTION PLAN (9 February 2006), *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/consult.html.

C. SERVICES OF GENERAL ECONOMIC INTEREST

Services of General Economic Interest (SGEI) are of particular importance for citizens and usually need to be financed through state intervention. Whether such financing constitutes State aid is a very fact specific question. The Commission's objective in this area is, therefore, to set up a clear procedural framework which will ensure that companies can receive public support to cover all costs incurred when providing an SGEI entrusted to them. In particular, the package adopted in July 2005 should provide legal certainty with regard to compensatory measures, while ensuring transparency in order to avoid both overcompensation and cross-subsidization from SGEI earmarked funds into non-public service activities. The latter constitutes a particular concern in liberalizing markets such as telecoms, post, or energy markets.

The cornerstone of the package is a Commission Decision, adopted on the basis of Article 86(3) of the EC Treaty,[23] which specifies the conditions under which compensation to companies for the provision of public services is compatible with State aid rules and does not have to be notified to the Commission in advance. The decision plays a role similar to the block exemptions adopted in other State aid areas, and exempts from the notification obligation any compensation of less than EUR 30 million per year, provided its beneficiaries have an annual turnover of less than EUR 100 million. This decision should seriously reduce red tape with regard to the financing of SGEI by local and regional authorities.

For all forms of compensation not covered by the decision—for instance those whose amount exceeds the ceiling—the SGEI framework[24] specifies the conditions under which compensation is compatible with State aid rules. Such compensation will have to be notified to the Commission due to the higher risk of distortion of competition. Compensation that exceeds the costs of the public service, or is used by companies on other markets open to competition, is not justified, and is therefore incompatible with the State aid rules concerning the SGEI.

Finally, an amendment to the Commission's Transparency Directive[25] clarifies that companies receiving compensation and operating on both public service and other markets must have separate accounts for their different activities, so

23 Commission Decision of 28 November 2005 on the application of Article 86(2) of the EC Treaty to State aid in the form of public service compensation granted to certain undertakings entrusted with the operation of services of general economic interest, 2005 O.J. (L 312) 67.

24 Community framework for State aid in the form of public service compensation, 2005 O.J. (C 297) 4.

25 Commission Directive 80/723/EEC of 25 June 1980 on the transparency of financial relations between Member States and public undertakings, 1980 O.J. (L 195) 35, *as amended lastly by* Commission Directive 2005/81/EC of 28 November 2005 amending Directive 80/723/EEC on the transparency of financial relations between Member States and public undertakings as well as on financial transparency within certain undertakings, 2005 O.J. (L 312) 47.

that the absence of over-compensation can be documented by the company concerned and checked by the Commission.

These texts have been well received by Member States and by the academic world.[26] They should create the legal certainty required in the area of financing of SGEI, by providing for clear exemptions for smaller SGEI and offer, for the rest, a clear framework allowing a refined economic analysis of their positive and negative effects.

D. COMMISSION COMMUNICATION ON INNOVATION

Two months after the adoption of the SGEI package, the Commission adopted a communication on aid to innovation.[27] This addressed a broad range of issues, and was well received.[28] Many of the measures have already been adopted in the new draft framework for Research and Development and Innovation (R&D&I) and in the Risk Capital Guidelines. The proposals for innovation aid cover six broad areas: innovative start-ups; risk capital; the integration of innovation into existing rules on State aid for research and development; innovation intermediaries; training and mobility between university research personnel; and, SMEs and poles of excellence for projects of common European interest.

E. COMMISSION GUIDELINES ON REGIONAL AID

In December 2005, the Commission adopted a first set of comprehensive guidelines covering the whole range of regional aid measures.[29] Reducing disparities between the regions of Europe is beneficial for all EU citizens as it is both a factor of social stability and provides a tremendous potential for economic growth. It was thus essential to adapt the rules governing regional aid, especially to take into account the recent enlargement of the Community. As a result of enlargement, the gap between richer and poorer regions has increased. Consequently, the new guidelines set out the rules for allowing State aid that promotes the development of poorer regions, covering aid such as direct investment grants and tax reductions for companies. The Guidelines specify rules for the selection of regions that are eligible for regional aid, and define the maximum permitted levels of this aid. In line with EU cohesion policy and European Council requests

26 *See, e.g.*, State Aid Group, Economic Advisory Group on Competition Policy (EAGCP), Opinion: Services of General Economic Interest (2006), *available at* http://ec.europa.eu/comm/dgs/competition/sgei.pdf.

27 Communication of the Commission, Consultation document on State aid for innovation, COM(05)436 final, Sep. 2005, *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/cdsai_en.pdf.

28 See European Commission, Results of the consultation on State aid for innovation (5 May 2006), *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/rep_inno.html.

29 *See* European Commission, Press Release No. 1653, State aid: Commission adopts new regional aid guidelines for 2007-2013 (2005) and Guidelines on national regional aid for 2007-2013, 2006 O.J. (C 54) 13.

for less and better targeted State aid, the new guidelines refocus regional aid on the most deprived regions of the enlarged Union. This should help narrow the gap between regions in Europe.

The new guidelines also provide for more flexibility for Member States to decide where and how they want to support regional development. For instance, a new form of aid will be allowed to encourage business start-ups in assisted areas, which will apply to the establishment and expansion phases of small enterprises during the first five years. The new guidelines also contain a number of other changes to clarify and simplify the current rules. In particular, the rules on very large investment projects of over EUR 50 million are included in the regional aid guidelines for the first time, thereby providing for an all-encompassing framework increasing transparency and readability of State aid policy in the area of regional support.

F. SHORT-TERM EXPORT-CREDIT INSURANCE

In December 2005, the Commission also adopted a revised communication on short-term export-credit insurance.[30] The purpose of the communication is to remove distortions of competition due to State aid provided in the sector of short-term (less than two years), export-credit insurance, where there is competition between, on the one hand, public or publicly supported export-credit agencies and, on the other hand, private export-credit insurers. Further to an external study[31] and after consultations with Member States and the private sector, the Commission adopted a revised communication.[32] The modifications take into account recent developments on the market of credit insurance, in order to determine precisely in which sub-markets private insurers are indeed providing services throughout the Community. These modifications are in line with the refined economic approach, in that they focus on the question of whether there is a market failure. The existence of such a market failure was indeed acknowledged in this case.

G. RISK CAPITAL

Risk capital and private equity funding are important for the EU's competitiveness. Insufficient availability of such funding in many parts of the Community

30 Communication of the Commission to the Member States pursuant to Article 93 (1) of the EC Treaty applying Articles 92 and 93 of the Treaty to short-term export-credit insurance, 1997 O.J. (C 281) 3.

31 European Commission, The Report on Market Trends of Private Reinsurance in the Field of Export Credit Insurance (2006), *available at* http://ec.europa.eu/comm/competition/state_aid/others/export_credit_insurance_report.pdf.

32 Communication of the Commission to Member States amending the communication pursuant to Article 93(1) of the EC Treaty applying Articles 92 and 93 of the Treaty to short-term export-credit insurance, 2005 O.J. (C 325) 22.

means that public funding may usefully serve as a means to leverage private funds. After a public consultation in 2005 and a study on the size of the equity gap commissioned by the Commission, we adopted new Guidelines in July 2006 that will help stakeholders to determine when State aid in support of risk capital investment in small and medium-sized enterprises (SMEs) is compatible with State aid rules.[33] This will allow Member States to facilitate access to finance for SMEs in their early stages of development, particularly where alternative means of funding from financial markets are lacking. The Guidelines are a prominent example of the Commission's efforts to encourage Member States to focus State aid on improving the competitiveness of EU industry.

An important change in the Guidelines resides in the fact that they include a safe harbor, set at an investment threshold of EUR 1.5 million per SME over a period of 12 months, an increase of 50 percent on the previous threshold. For these cases, the Commission accepts the presence of a market failure, in that alternative means of funding from financial markets are lacking. Above this threshold, because of the greater potential to distort competition, the Commission will make a detailed assessment, and Member States will have to provide evidence of a market failure. Applying different types of assessment on the basis of economic impact is an important change, and implements the SAAP's refined economic approach.

H. DE MINIMIS

The current de minimis rule is contained in Commission Regulation No. 69/2001 of 12 January 2001.[34] De minimis aid is considered as falling outside the scope of the State aid rules, because of a presumption that it neither affects trade between Member States nor distorts competition. The instrument should be handled with care because, contrary to the other Commission regulations in the State aid area, it does not detail conditions of compatibility under Article 87(3) EC Treaty, but simply defines what is considered as State aid within the meaning of Article 87 (1). The current threshold is set at EUR 100,000 over a three-year period.

After publication of a draft in March 2006 and a first consultation round with Member States and other stakeholders, the Commission issued a revised draft for consultation in June 2006, proposing to raise the ceiling to EUR 200,000.[35] In so doing, it implemented its pledge, announced in the SAAP, to increase the de

33 *See id.*

34 Commission Regulation No 69/2001 of 12 January 2001 on the application of Articles 87 and 88 of the EC Treaty to De minimis aid, 2001 O.J. (L 10) 30.

35 Amended new de minimis block exemption, 2006, O.J. (C 137) 4, *available at* http://ec.europa.eu/comm/competition/state_aid/overview/sar.html.

minimis threshold to take account of developments in the economy. This should result in a greater number of small subsidies being exempted from the notification obligation under EC State aid rules.

The proposal also contains a number of safeguards to prevent abuse of its provisions, especially with regard to so-called non-transparent forms of aid. Indeed, in line with earlier Commission proposals and, more particularly, the new draft for a block exemption covering certain types of regional aid,[36] we proposed limiting the scope of the de minimis regulation to transparent types of aid, defined as being those measures for which it is possible to determine in advance the precise aid amount they include without carrying out a risk assessment.

The final version of this regulation should attempt to cover the largest amount of aid which is unlikely to have any damaging impact on competition—such as, for instance, guarantee schemes in favor of real SMEs—thereby reducing the bureaucratic burden for those support measures, while simultaneously ensuring that the de minimis regulation is not being abused, for instance in favor of large undertakings active on wider geographical markets.

I. RESEARCH AND DEVELOPMENT AND INNOVATION (R&D&I)

Even though effective competition is the best tool to strengthen innovation and competitiveness in the European Community, State aid can occasionally also play a very useful supporting role. State aid can be used to embrace globalization by better targeting public funds, for instance, towards R&D&I, thereby supporting economic reform to deliver long-term competitiveness. With this in mind, the Commission proposed, in April 2005, a draft Community framework for State aid for research and development and innovation.[37]

Regarding the substance of the rules, we prepared a common text covering not only research and development, but also innovation, given the close links between the two areas. The text generally maintains the existing high aid intensities for fundamental and industrial research, while introducing a new category of experimental development, substantially broadened to include innovation activities. The draft also intends to provide increased legal certainty for R&D projects of universities and for public-private partnerships. Most importantly, there will be ground-breaking new rules on support for innovation.

With regard to the procedural treatment of R&D&I aid measures, notified measures will of course be subject to the refined economic analysis set out in the

36 Draft Commission Regulation on the application of Articles 87 and 88 of the EC Treaty to national regional investment aid, 2006 O.J. (C 120) 2.

37 *See* European Commission, Draft Community framework for state aid for research and development and innovation (20 April 2006), *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/rdi_frame_en.pdf.

SAAP. As we propose to authorize new categories of aid for innovation, and higher aid intensities, we also need, in parallel, to concentrate more of our resources on the most distortive cases. Indeed, State aid may well create the correct incentives to increase R&D&I in the right circumstances, but it may also imply serious disruptions of dynamic effects for the competitors of aid beneficiaries.

IT IS WORTH STRESSING THAT A DETAILED ANALYSIS DOES NOT MEAN THAT THE COMMISSION WILL NECESSARILY PROHIBIT THE AID. IT MERELY IMPLIES THAT THE POSITIVE AND THE NEGATIVE EFFECTS WILL BE LOOKED AT IN MORE DEPTH.

It is worth stressing that a detailed analysis does not mean that the Commission will necessarily prohibit the aid. It merely implies that the positive and the negative effects will be looked at in more depth, using assessment methods similar to those which the beneficiaries themselves are using before deciding to embark on any large business project.

J. REGIONAL BLOCK EXEMPTION REGULATION

Simultaneously, with the adoption of the new Guidelines on regional aid,[38] the Commission also proposed to introduce for the first time a block exemption covering regional investment aid.[39] The objective of the draft Regulation is to simplify the administrative handling for Member States, while reinforcing transparency and legal certainty. A first draft of the Regulation was discussed with Member States in April 2006, and a revised version will be discussed again in early autumn. The objective is to adopt the regional Block Exemption Regulation before the end of this year, so that it will enter into force on January 1, 2007, for the new structural funds programming period. That regulation, as a stand-alone instrument, will be of a transitory nature as the areas covered by it are intended to be included in the general, overarching block exemption, to be proposed and discussed in the course of 2007.

K. COUNCIL ENABLING REGULATION

As announced in the SAAP, the European Commission intends to adopt a proposal for a modification of the Council Enabling Regulation by the end of 2006,[40] in order to enable the Commission to adopt block exemptions in new fields like culture, heritage conservation and natural disasters. Indeed, the Commission could, under the provisions of the existing Enabling regulation, only block

38 *See supra,* Point III 1 C.

39 Draft Commission Regulation on the application of Articles 87 and 88 of the EC Treaty to national regional investment aid, 2006 O.J. (C 120) 2.

40 Council Regulation (EC) No 994/98 of 7 May 1998 on the application of Articles 92 and 93 (now 87 and 88 respectively) of the Treaty establishing the European Community to certain categories of horizontal State aid, 1998 O.J. (L 142) 1.

exempt aid in these areas provided the aid was in favor of SMEs. Excluding large companies from the benefit of such block exemptions is difficult to justify. This situation should therefore be changed. Due to the length of the legislative procedure, however, a new Council Regulation cannot be expected before the beginning of 2008.

V. CONCLUSIONS AND NEXT STEPS

The SAAP will adapt State aid policy to the new challenges facing the European Community. Though the reforms that have been implemented recently in the antitrust and merger area are often cited as a reference, the SAAP has to face a number of supplementary challenges. As explained above the welfare standard on the basis of which we have to conduct our assessment is different from the one in mergers and antitrust and the procedural challenges that we face are also considerable.

As has been set out above, we have already made significant progress in implementing the Action Plan. More work remains to be done.

We need to establish an overarching new general block exemption regulation (GBER), which will integrate all the existing block exemption regulations and also cover the new areas identified in the SAAP. The Commission is currently working on a first draft of the GBER, which will significantly expand the scope for Member States to grant aid without having to notify it to the Commission. The current block exemption already covers aid to SMEs, training aid, and employment aid.[41] In addition to those areas, the GBER should include exemptions in favor of regional aid, R&D&I aid to large companies, and environmental aid. The simpler administrative procedure created by the GBER should mostly benefit SMEs.

The Commission is also considering possible modifications to existing guidelines, such as the environmental guidelines.[42] These new guidelines should follow the same structure regarding the economic approach as the already-developed draft R&D&I Guidelines and the Risk Capital Guidelines. The final adoption of these guidelines is planned for the third quarter of 2007.

Reflection on more fundamental procedural reforms which would involve changing the rules has also been initiated. It concerns more ambitious reform ideas intended, first, to save time and increase transparency, second, to ensure that State aid is duly notified or recovered if implemented illegally and third, that greater administrative efficiency be achieved, amongst others by allowing for the relevant sectoral information to be gathered more easily. Reform ideas

41 For an overview, see EUROPEAN COMMISSION, BLOCK EXEMPTION REGULATIONS, *available at* http://ec.europa.eu/comm/competition/state_aid/legislation/block.html.

42 Community guidelines on State aid for environmental protection, 2001 O.J. (C 37) 3.

will be subjected, at an early stage, to a risk-benefit analysis in order to assess possible implications, amongst others, for third party rights, the burden of proof of the respective parties in the procedure and the locus standi of parties in court procedures. These reflections should feed into a consultation document to be discussed with Member States in the course of 2007.

THE PROJECT IS NOT LIMITED TO THE COMMISSION ALONE. COOPERATION FROM THE MEMBER STATES AND PARTICIPATION OF ALL STAKEHOLDERS IS FUNDAMENTAL TO THE SUCCESS OF THE REFORM.

The more refined economic approach and the new procedural framework to be established by the SAAP are not intended to be either more interventionist or more lax. The objective is to provide for a sounder basis for intervention. The project is not limited to the Commission alone. Cooperation from the Member States and participation of all stakeholders is fundamental to the success of the reform. And the success of the reform is in turn fundamental to the success of the Lisbon goals of competitiveness, growth, and jobs.

VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

The State Aid Action Plan: A Bold Move or a Timid Step in the Right Direction?

Frédéric Jenny

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

The State Aid Action Plan: A Bold Move or a Timid Step in the Right Direction?

Frédéric Jenny

This paper explores the current state of discussion on the reform of EC State aid policy. It analyzes the responses to the European Commission's consultation document on State aid, presents the main areas of dissatisfaction, and the extent to which answers to the expressed dissatisfaction can be found in the State Aid Action Plan or the comments of Philip Lowe. The paper then explores the pros and cons of a decentralization of EC State aid policy.

The author is Professor of Economics, ESSEC Business School and Visiting Professor, University College London.

Philip Lowe's article "Some Reflections on the European Commission's State Aid Policy"[1] meets a number of the concerns that have been expressed in the public debate on the reform of EC State aid policy; and there is little doubt that the current thinking of the European Commission is a welcome move toward improving and modernizing an area of competition law that has been somewhat neglected in the past. However, Philip Lowe does not address some of the concerns raised either during the public consultation organized by the Commission or by several authors prior to the consultation. The numerous and very serious criticisms levied against the past enforcement of provisions on State aid in the EC Treaty can be classified in four categories.

First, State aid law enforcement is seen as an area of legal uncertainty and of dubious economic relevance. It is often alleged that what constitutes State aid has not been clearly defined. Indeed, it appears that the Commission and the EC courts do not share the same vision of what constitutes a State aid. The Commission seems to favor a wide interpretation (as it did in the France Telecom case where the Commission argued that statements by the French government between July and December 2002 amounted to illegal assistance for the French telecom operator),[2] whereas the courts seem to have a narrower view of what constitutes State aid.

It is also often observed that, in spite of or because of many guidelines, there is no easy method for distinguishing State aid that needs to be notified from other measures that do not need to be notified.

And it is often noted that what constitutes illegal State aid is generally decided without any assessment of a possible effect on trade between Member States or any assessment of possible distortion of competition. For example, in its response to the Commission's consultation on the "State Aid Action Plan" (the SAAP), the Confederation of British Industries stated:

> "Effects, rather than form, based economic analysis, such as the Commission has in recent years moved towards in its guidance on the application of Article 81, has a fundamental role to play in the assessment of State aid. There is also a need for a much more rigorous approach in defining interstate [*sic*] trade effects. A number of recent decisions (e.g., *Brighton pier; Dorston swimming pool*) have involved State aids which have no real effects

1 Philip Lowe, *Some Reflections on the European Commission's State Aid Policy*, 2(2) Competition Pol'y Int'l 57 (2006).

2 The Commission statement when the decision was published said "the statement created expectations and confidence on the financial markets and helped maintain France telecom's investment rating. If the statements had not been made no reasonable investor would have offered a stakeholder's advance in these circumstances and assumed alone a very large financial risk."

on trade, and where any concerns regarding trade distortions are entirely academic. Decisions of this nature encourage Member States to notify on a precautionary basis schemes which have no real European dimension. This represents a patent waste of scarce resources, which could be better deployed in tackling major rescue and restructuring packages. The Commission therefore must introduce far greater rigour in its analysis of trade effects."[3]

THE COMPLEXITY AND UNPREDICTABILITY OF STATE AID RULES IS LIKELY TO CREATE UNCERTAINTY FOR UNDERTAKINGS (WHETHER BENEFICIARIES OR COMPETITORS) IN ASSESSING WHETHER A MEASURE CONSTITUTED ILLEGAL AID, AND FOR NATIONAL JUDGES IN DECIDING WHETHER OR NOT A MEASURE CONSTITUTES STATE AID.

The complexity and unpredictability of State aid rules is likely to create uncertainty for undertakings (whether beneficiaries or competitors) in assessing whether a measure constituted illegal aid, and for national judges in deciding whether or not a measure constitutes State aid.[4]

The legal uncertainty resulting from the complexity of defining what is a State aid and whether or not a State aid is compatible with the common market, may be reinforced by the fact that national judges do not have the same powers as the Commission in the area of State aid. A national judge has the power to declare the illegality of an aid that has been granted and to pronounce judgment on the consequences of this illegality under national law, but a national judge cannot assess the compatibility of the aid with the common market. The Commission, on the other hand, can assess the compatibility of the aid and order recovery when the aid does not qualify for one of the exceptions set out in paragraphs 2 and 3 of Article 87 of the EC Treaty. Some observers in the business community believe that such differentiated powers can lead to divergent solutions for similar cases. For example, in its submission to the European Union on the SAAP, the American Chamber of Commerce gave the example of two Ryanair actions brought before the Commission and the French administrative courts. "In those two cases, similar State aid measures benefiting Ryanair gave rise to totally distinct solutions." According to the American Chamber of Commerce, "these divergent outcomes undermine legal certainty."[5]

3 CONFEDERATION OF BRITISH INDUSTRY, CBI'S RESPONSE TO THE COMMISSION'S CONSULTATION 6 (2005), *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/consult/37229.pdf.

4 *See*, for example, LINKLATERS, SUBMISSIONS ON EUROPEAN COMMISSION'S STATE AID ACTION PLAN, LESS AND BETTER TARGETED STATE AID: A ROADMAP FOR STATE AID REFORM 2005-2009 (2005), *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/consult/37367.pdf.

5 AMERICAN CHAMBER OF COMMERCE TO THE EUROPEAN UNION, POSITION PAPER ON STATE AID ENFORCEMENT (2005), *available at* http://www.amchameu.be/Pops/2005archive/stateaidenforcement05282005.pdf.

A second major area of criticism concerns the review process at the EU level. The notification system is considered by many to be too cumbersome. Part of the problem comes from the fact that under the EC treaty, Member States are required to notify all aid for assessment by the Commission, and the Commission is the only institution that can authorize State aid under Article 87(2) and (3) EC. Combined with the legal uncertainty mentioned above, this system leads the Commission to spend too much time on non-problematic State aids, leading to delays in the handling of more important cases. Thus, as the American Chamber of Commerce puts it: "The Commission should prioritize its handling of complaints, concentrating on those relating to larger schemes which can have European-wide effects."[6]

A third area of dissatisfaction concerns the large number of State aids which are not notified. The "Study on the Enforcement of State Aid Law at National Level"[7] indicates, for example, that "Of the 400-550 or so cases dealt with by the Commission every year about 15-20% concern cases which are not notified." This study offers three possible explanations to explain this failure of Member States to notify State aid.

First, the study says, it could be that, as mentioned previously, Member States have difficulty in distinguishing between State aid that falls within Article 87 (1) EC and therefore has to be notified, and State aid that does not.

Second, perhaps EC procedures are too cumbersome and too time-consuming for Member States that, in certain cases, have to grant aid quickly to avoid a major political and/or economic problem.

Third, it could be that Member States provide aid to their national industries without any regard to EC rules. Whereas problems identified in the first two possible explanations could be remedied relatively easily by the Commission through technical measures (clarifying the concept of State aid and modifying the notification procedure so as to alleviate the burden of the Commission and thus allowing for a faster treatment of notified cases), the third possible explanation for the failure of Member States to notify their State aid is more challenging. If this last explanation is an accurate description of reality, the Commission must do a better job of advocacy in the area of State aid, the penalties for failure to notify or for granting illegal State aid should be reviewed, and a better system of detection of non-notified aids should be established so that Member States cannot get away with notification lapses.

6 American Chamber of Commerce to the European Union, AmCham EU Response to the Commission's Consultation on the State Aid Action Plan Consultation (2005).

7 Thomas Jestaedt, Jacques Derenne, and Tom Ottervanger, Study on the Enforcement of State Aid Law at National Level, (2006), *available at* http://ec.europa.eu/comm/competition/state_aid/others/study_part_1.pdf.

This leads us to the fourth area of dissatisfaction with the current system of State aid enforcement: the fact that the decentralized recovery mechanism is at best faulty. At present, Article 14, paragraph 3, in combination with Article 14, paragraph 1 of the procedural regulation, serves to ensure that the Member State concerned takes all necessary measures to recover illegal State aid, in accordance with the procedure set out by national law.

However, the fact that the Member State is the competent authority to recover the illegal aid, although it is also often the authority which previously granted this aid, can be a source of difficulties. According to the "Study on the Enforcement of State Aid Law at National Level," while there are encouraging signs that recovery of illegal or prohibited State aid is becoming more satisfactory, at least five types of obstacles exist impeding the recovery process at the national level:

1. In some cases, there is a lack of clarity as to who should issue the recovery decision, who should repay, and the amount to be repaid;
2. In some Member States there is no clear predetermined procedure to recover aid;
3. Interim measures for the recovery of State aid are either unavailable or unused;
4. In several countries recovery proceedings will be stayed while an appeal is pending; and
5. Governmental authorities of Member States may experience difficulties in recovering illegal State aid at the regional or local level.

As a result, the deterrent effect of State aid enforcement remains weak. This is recognized in the SAAP, which states:

> "The effectiveness and credibility of State aid control presupposes a proper enforcement of the Commission's decisions, especially as regards the recovery of illegal and incompatible State aid. Recent experience has shown that the implementation of recovery decisions by Member States is not satisfactory and, moreover, that conditional or positive decisions are sometimes not correctly implemented by the Member States."[8]

8 European Commission, State Aid Action Plan, Less and Better Targeted State Aid: A Roadmap for State Aid Reform 2005-2009 13 (2005), *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/saap_en.pdf.

As Philip Lowe makes clear, the SAAP tries to meet a number of those concerns, and some of his reflections, which go beyond the SAAP, are also very clear indications of the fact that the Commission is responsive to the concerns previously mentioned.

First, it is hard to disagree with the four guiding principles in the SAAP:

- less and better targeted State aid;
- a refined economic approach;
- better procedures and administration; and
- a shared responsibility between the Commission and Member States.

At first glance, these principles seem consistent with the wishes of commentators who seek a clearer and more precise definition of what State aids are, a more economically relevant analysis of State aids, better procedures for the investigation of State aids and fewer notifications of insignificant State aids, and a better compliance of Member States with EC law in this area.

On a more detailed level, one must welcome the desire of the Commission to shift from a form-based approach to an economic analysis of the effects. This can only increase the economic relevance of State aid enforcement. Since "State aid can affect both the way in which the economic pie is made larger by a given policy (efficiency) and how the pie is then divided between citizens (equity),"[9] the Commission proposes to use a social welfare standard (thus recognizing that objectives other than market failures can justify the granting of State aid), and to use a clearly delineated balancing test to evaluate the compatibility of the aid with EC rules by examining three questions: Is the aid aimed at a well-defined objective of common interest? Is the aid well designed to deliver the objective? Are the distortions on trade and competition sufficiently limited so that the overall balance is positive? While this approach represents a clear progress, it is not entirely clear how the Commission will be able to implement the balancing test in cases in which the objective of the State aid is not to correct a market failure leading to inefficiencies but to achieve an equity goal.

Second, the SAAP clearly recognizes that "there are certain shortcomings in the practices and procedures of State aid policy, which can be observed in the long time frame for the treatment of cases" and that "longer time frames are clearly an unacceptable outcome, bearing in mind that a trade off might exist between the duration of the procedure and ensuring an effective control while safeguarding the rights of third parties." The plan suggests various means to "improve its internal practice and administration, and increase efficiency, enforcement and monitoring." In particular, delays could be shortened "within the scope of the current procedural regulations" such as instilling:

9 Philip Lowe, *supra* n. 1, at 67.

> "more predictable timelines, clear intermediary steps in the procedure and ensure higher transparency by providing more information on Internet,...encouraging a higher quality of notifications and by discouraging incomplete notifications by a more systematic use of the information injunction, requesting Member States to provide complete information within a certain period,...issuing best practices guidelines after consulting Member States as well as the public on how procedures could be improved to better administrate State aid control."[10]

For the Commission, the "best practices guidelines together with the general block exemption and the increased de minimis ceiling are expected to reduce both the time it takes before the Commission reaches its decisions and the administrative burden for Member States." All those suggestions could help meet the concern of those who, as we saw previously, argue that the process of State aid enforcement is too burdensome for the Commission, for undertakings, and for the Member States.

Third, the Commission seems intent on ensuring a better compliance with State aid rules by Member States. The SAAP acknowledges that "the effectiveness and credibility of State aid control presupposes a proper enforcement of the Commission's decisions, especially as regards the recovery of illegal and incompatible State aid" and that "implementation of recovery decisions is not satisfactory." In this area, the Commission proposes to follow two tracks. First, to monitor more closely the execution of recovery decisions by Member States and to pursue more actively noncompliance under Articles 88(2), 226, and 228 of the EC Treaty (an implicit acknowledgement that it has not put enough emphasis on this aspect in the past). Second, to "promote advocacy, awareness and understanding of State aid control at all levels to help the granting authorities in designing measures that are compatible with the treaty rules."[11]

However, there is one major area where the SAAP and Philip Lowe's reflections fall short of the expectations.

The SAAP includes one reference to the issue of decentralization and coordination between the Commission and Member States. This topic has been hotly debated both in academic circles and in the replies to the consultation undertaken by the Commission.

10 European Commission, *supra* n. 8, at 12.

11 *Id*. at 11.

In paragraph 51, the SAAP suggests that Member States could (and should) play a crucial role in the implementation of the rules and procedures on State aid by monitoring and screening State aids. The SAAP states that the European Community has already had a positive experience in this area in the context of enlargement where "the screening of State aid measures was conducted by operationally independent monitoring authorities in the new Member States." Drawing on this valuable experience, the Commission will "examine whether independent authorities in Member States could play a role as regards the task of the Commission in terms of State aid enforcement (detection and provisional recovery of illegal aid, execution of recovery decisions)."[12] In addition, as Philip Lowe notes, an attempt will be made to "intensify the involvement of national courts, especially with regard to the treatment of illegal aid granted in violation of the notification obligation."

Thus, it appears that the Commission does not envisage a complete (or even a wide) decentralization of decision making in the area of State aid enforcement, while cautiously considering the possibility that Member States or national independent authorities could play some role in the substantial evaluation of State aid.

Even though one may regret the lack of concrete proposals for decentralization, it must be said that the Commission's stand seems to be partly inspired by the views expressed by Professor Phedon Nicolaides,[13] who has argued that partial decentralization of State aid enforcement at the national level would, under certain circumstances, be preferable to the current complete centralization of enforcement.

Phedon Nicolaides does not favor complete decentralization of the enforcement of State aid at the national level for a reason that is likely to alarm lawyers and business firms and to be considered weak, at best, by economists. He states that:

> "It is obvious that the Commission plays an important role in State aid control... This is not so much because only the Commission has the requisite knowledge or the impartiality to decide what is in the EU's common interest. I believe that it has more to do with the fact that in the end these decisions have an element of arbitrariness. There is no way of telling beforehand where the line should be drawn between exemptable and non-exemptable State aid. Reasonable persons starting with the same assumptions could easily arrive at different conclusions. Hence, if the rules are to be applied uni-

12 *Id.* at 12.

13 See, for example, Phedon Nicolaides, *Decentralised State Aid Control in an Enlarged European Union: Feasible, Necessary or Both?*, 26(2) World Competition 263–276 (2003).

> formly across the European Union it is important to have a single authority that defines the boundaries of the rules."[14]

However, Professor Nicolaides expressed strong support for the idea of entrusting independent national authorities with partial responsibilities in the enforcement of State aid, by stating that:

> "A requirement for national authorities to measure the economic impact of State aid they propose to grant and to demonstrate how it corrects market imperfections would.... make it more difficult for politicians to claim that aid is in the national interests and would correspondingly make it easier for national officials to speak out against aid that demonstrably does not raise overall national welfare. Perhaps ministries of finance would even welcome rules that would facilitate their task of ensuring that public expenditure generates "value for money.""[15]

He summarizes a possible division of labor between the Commission and the independent national authorities along the following lines:

Figure 1

PROPOSED DIVISION OF TASKS BETWEEN THE COMMISSION AND NATIONAL STATE AID AUTHORITIES

Policy foundation	Commission tasks	SAA tasks
Existing State aid rules	• Elaborate State aid policy • Issue regulations • Define guidelines • Deal with aid falling outside guidelines • Monitor decisions of SAAs • Manage a network of SAAs	• Receive all State aid notification • Check their compliance with EU rules • Reject non-compliant schemes • Forward to the Commission schemes that fall outside guidelines (which require assessment of common interest) • Monitor compliance with block exemption and de minimis regulations
Modernized State aid rules (relying more on market analysis and requiring assessment of economic impact of State aid)	As above	As above, plus • Check market analysis and assessment of economic impact

14 *Id.*

15 *Id.*

As Professor Nicolaides does in his previously mentioned article, in the SAAP, the Commission refers to the fact that, in the context of accession, candidates for EC membership had to create State aid monitoring authorities in charge of assessing all State aid granted in their countries and to keep the Commission informed of the cases they handled. Because these authorities seem to have fulfilled a useful role and worked in close cooperation with the Commission, Professor Nicolaides suggests that similar independent institutions could be established in all Member States (or at least in some of them) to analyze the compliance of State aid applications with EC rules. Although considerably more tentative in its expression, the SAAP similarly states that the experience conducted during the enlargement process "has been a valuable experience which should be taken into account when considering further cooperation between the Commission and all Member States." In this context, the Commission will examine whether independent authorities in Member States could play a role facilitating the task of the Commission in terms of State aid enforcement—detection and provisional recovery of illegal aid, and execution of recovery decisions.

In spite of its cautiousness, this proposal by the Commission has opened a lively debate. Some commentators have expressed strong reservations about the idea that national independent authorities could play a role in the assessment of the compatibility of State aid with EC rules. Typical of this line of thought, Eric Morgan de Rivery and Nelly Le Berre-Dodet[16] suggest that such a proposal would raise both legal and practical difficulties. In particular, they point out that it could raise a constitutional issue regarding the Commission's exclusive role in assessing the compatibility of State aid measures, and a practical difficulty of identifying capable entities and ensuring the right level of cooperation from the vast number of entities potentially capable of granting State aids.

Others have been resolutely opposed to the idea that independent national authorities could play a useful role in the recovery of illegal State aid. For example, in its submission on the SAAP, Linklaters, a major law firm,[17] states: "In paragraph 51 SAAP, the Commission proposes that independent national authorities might act as the Commission's agents in enforcing State aid rules." Linklaters is skeptical about this enhancement of the role of Member States. If national authorities were designated to recover aid, burdensome implementation measures would have to be adopted to ensure consistency across the European Community.

16 Eric Morgan de Rivery and Nelly Le Berre-Dodet, *Controlling State Aids*, COMPETITION LAW INSIGHT (12 July 2005).

17 LINKLATERS, *supra* n. 4, at 5.

> "We do not believe that it is appropriate to draw a comparison with the experience of accession Member States. First, it is not clear that the experience of pre-accession aid proceedings was as positive as the Commission suggests. Second, the accession process provided a strong incentive that will be absent in the current Member States notwithstanding their duties under Article 10 EC. Third, the national authorities in question did not have to recover but rather decide on the existence of aid and compatibility."[18]

But not all commentators share those reservations and some have offered their own view of what the role of national independent agencies could be in the State aid area.

For example, Association Française d'Etude de la Concurrence,[19] an organization of French lawyers, suggested that such independent agencies could keep databases; be granted the right to question Member States about aids allocated, their regularity, and the exemption regulations from which they possibly benefit; and have a power to alert or inform the Commission, answer requests for information received from operators, answer other national authorities that are part of the network, assist central state, regional and local authorities, and undertake a calibration of performances.

The submission of the U.K. government on the SAAP[20] envisions two options that would confer a much more active role for national authorities in State aid enforcement with a view to speed up the assessment of the compatibility of State aids with EC law. The first option proposes to offer at least some Member States the possibility of securing an independent review of the distortion of competition that might result from the proposed State aid. It adds: "Such a review, if authoritative, could perhaps obviate the need for detailed investigation by the Commission itself in some cases." The second option, which is very much in line with the proposals of Professor Nicolaides, would be to remove the notification obligation for cases that fall outside the conditions for direct block exemption but are nevertheless within broader safe harbor limits and therefore unlikely to seriously distort competition on a European scale, if Member States obtained an

18 *Id.* at 6.

19 Association Francaise d'Etude de la Concurrence, Comments of AFEC on the State Aid Action Plan of the Commission (2005), *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/consult/37969.pdf.

20 United Kingdom, UK Response to Commission Roadmap for State Aid Reform, *available at* http://ec.europa.eu/comm/competition/state_aid/others/action_plan/consult/37379.pdf.

opinion from an independent national competition authority that the measure is unlikely to distort competition to any significant extent.

The modernization of EC competition law has led to a successful decentralization of the enforcement of Article 81 EC and the elimination of the Commission's monopoly on the interpretation of Article 81 (3) EC. It is already obvious that these have been successful moves. Cooperation between the Commission and national authorities within the European Competition Network is recognized by all as being highly satisfactory; enforcement of EC competition law is now more frequent at the national level and, as a consequence, EC law is better understood; the Commission is now better able to focus its scarce resources on cases of major significance.

Yet the decentralization of enforcement of Article 81 EC was initially resisted by some who feared that national competition authorities could not be trusted because they were neither sufficiently independent nor sufficiently technically competent to enforce Article 81 (3) EC in a consistent and objective manner, or that they were too numerous (and too heterogeneous) for cooperation between them and the Commission to be workable.

When it synthesized the results of the consultation of the SAAP in February 2006, the Commission noted that the principle obstacles mentioned by respondents who questioned the possible role of independent national authorities in the enforcement of State aid laws were: the independence of such national authorities, the risk of increased bureaucracy, the risk of uneven application of the law, a concern about the legality of a full delegation of responsibility in this area to national independent authorities, and the principle of institutional autonomy of the Member States. Most of these concerns had been raised in the discussions leading to the decentralization of the enforcement of Article 81 (3) EC, but either they turned out to be misguided or solutions were found to overcome them.

As we have seen, the supporters of a role for independent national authorities in the enforcement of State aid have never argued in favor of complete decentralization. The arguments of those who oppose giving a role to national independent authorities often seem to assume that the choice is between full decentralization or no decentralization whatsoever. Some respondents to the public consultation on the SAAP suggested that, if it wants to move forward, the Commission "should issue a specific document providing more clarity and a global picture of the powers and obligations of such authorities." Such a document could indeed clarify that no one is pushing for complete decentralization of State aid enforcement and it would make clear that different proposals with different degrees of involvement of independent national authorities are conceivable.

The very cautious position expressed in the SAAP on the issue of decentralization of enforcement of State aid provisions should be considered in light of the fact that at the time of the drafting of the SAAP, it was not yet obvious that the decentralization of the enforcement of Article 81 (3) EC would be a success.

What is disappointing, however, is that in 2006, at a time when it is clear that the alarm of those who opposed the decentralization of the enforcement of Article 81 (3) EC was unjustified, neither the Commission nor Philip Lowe, in his article "Some Reflections on the European Commission's State Aid Policy," have suggested that they intend to actively pursue a discussion on decentralization of enforcement in the State aid area.

A revision as major as the one undertaken by the Commission on EC State aid policy is a golden opportunity to establish conditions that could decrease the misunderstanding in this area between Member States' politicians and the general public, on the one hand, and the Commission, on the other hand, by promoting a public and transparent debate at the national level that would show that the Commission trusts national institutions and wants to cooperate with them. This will undoubtedly make State aid policy better understood and increase its effectiveness. It would be a particularly important result at a time when there is renewed, if misguided, interest in industrial policy measures and the promotion of "national champions" in many Member States.

IF THE SAAP IS IMPLEMENTED WITHOUT ANY FORM OF DECENTRALIZATION OF ENFORCEMENT, THE COMMISSION WILL HAVE ACHIEVED A USEFUL TECHNICAL OVERHAUL OF OUR STATE AID POLICY BUT IT WILL HAVE GIVEN UP THE CHANCE TO BENEFIT FROM A MOMENTUM IN FAVOR OF MAJOR REFORMS THAT MAY NOT REAPPEAR FOR MANY YEARS.

The results of the consultation on the SAAP show that most resistance to this proposal comes from regional authorities. Out of eight regions that commented on the SAAP, seven declared themselves against the proposal and one was in favor. In contrast, opinions were more evenly divided in other categories (six business associations in favor and six against, two law firms in favor and three against, five Member States in favor and eight against). These results suggest that the business community is generally, even if cautiously, in favor of independent national authorities. Thus, by providing a more precise set of options for decentralization and by giving special consideration to the issue of regional State aid, the Commission could reasonably hope to make progress, even if decentralization in the State aid area remains an uphill battle.

If the SAAP is implemented without any form of decentralization of enforcement, the Commission will have achieved a useful technical overhaul of our State aid policy but it will have given up the chance to benefit from a momentum in favor of major reforms that may not reappear for many years.

Volume 2 | Number 2 | Autumn 2006

Competition Policy International

The Economics of State Aid Control: Some Remarks

Mathias Dewatripont

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

The Economics of State Aid Control: Some Remarks

Mathias Dewatripont

This short comment discusses the rationale for State aid control at the level of the European Community and then turns to the State Aid Action Plan and its application to aid to Research, Development, and Innovation. It discusses the merits of a refined economic approach for "better targeting" purposes and stresses the need for State aid control to focus on preventing an excessive reliance on an "innovation-based industrial policy."

The author is Professor of Economics at the European Centre for Advanced Research in Economics and Statistics (ECARES), Université Libre de Bruxelles.

I. Introduction

In his paper,[1] Philip Lowe has offered an excellent discussion of the general objectives and challenges of EC State aid policy, as well as a detailed description of the "State Aid Action Plan" (the SAAP) and of its implementation so far. He has, however, essentially taken as given the existence of a State aid control, a subject of debate among economists, in particular since this policy is unique to the European Community. In this short comment, I will briefly address the rationale for a State aid control before turning to the new approach embodied in the SAAP, considered first in general, and second in the area of Research, Development, and Innovation (R&D&I).

II. Why State Aid Control?

The motto of the SAAP is: less and better targeted State aid. Of course, reducing State aid is not a new policy. And at one level, reducing State aid should not be controversial to economists: it merely reflects the idea of the primacy of competition policy over industrial policy, and experience has taught us how tough it is for public authorities to pick winners. Moreover, State aid controls can be seen as a straightforward consequence of the Single Market program.

On the other hand, State aid policy is not without its critics. Beyond the obviously embarrassing cases of tiny undertakings that have at times occupied European attention, and which should be avoided in the future thanks to more powerful de minimis provisions, there are at least two more substantive arguments to consider. The first one concerns the case where there is only one European firm in the market (e.g., Airbus): isn't Europe hurting its competitiveness, in a world where it is the only jurisdiction to have State aid control? It is true that such cases should ideally be dealt with in a forum like the World Trade Organization. It is clear that State aid control works best in the presence of competitors that act as watchdogs, which means that EC State aid control is most seriously enforced in markets with multiple European-based firms. Beyond this, one should keep in mind the general idea underlying the Single Market program, i.e., that vigorous competition at home does contribute overall to international competitiveness.

The second question concerns the question of European paternalism: should State aid control solely focus on distortions of competition or should it also protect European citizens from their national or regional governments? It is true that, without State aid control, there is a risk of having too much—and unwarranted—aid, because firms are typically better politically organized than taxpayers. And in fact, excessive public funding can even come as a by-product of polit-

1 Philip Lowe, *Some Reflections on the European Commission's State Aid Policy*, 2(2) Competition Pol'y Int'l 57 (2006).

ical accountability: Dewatripont and Seabright show in a theoretical model how inefficient aid can arise as a result of electoral concerns in an environment where politicians have to demonstrate to voters that they are actively trying to attract investment—that can in the end turn out to be good or bad for the region or country as a whole.[2]

This being said, while one can argue that State aid control is a commitment device against such governmental abuse that has moreover been agreed on ex ante by Member States, one should stress that it leads to an image problem for EU institutions, which typically end up playing the role of scapegoat in national politics. While it is obviously beyond the scope of this short comment to discuss in detail the pros and cons of State aid control, it is safe to say that this policy is on more solid ground when it explicitly focuses on distortions of competition.[3] Improving its ability to function along this dimension is the main goal of the SAAP.

III. Better Targeted Aid

As an economist, I am, of course, biased on this question, but an economic approach is in fact a very natural way of streamlining the existing—often ad hoc—case-by-case approach, since economists have, for a very long time, developed analyses trading off market failures and distortions of competition. See for example the paper by Friederiszik, Roeller, and Verouden (2006) for a discussion of the roadmap that can emerge from such an approach.[4]

As stressed quite rightly by Philip Lowe, the challenge for the economic approach comes from the fact that one cannot focus on a simple consumer standard relied on in usual competition policy, because the direct effect of State aid is good for consumers, since, at least in the short run, it typically implies lower prices. This can easily be reversed when one takes into account the tax cost involved in State aid, but trading off this cost with the fall in prices it implies is no easy task.

It makes sense, therefore, to base the approach on the pragmatic idea of compatibility with the single market and thus on the notion of distortion of competition, which means in particular focusing on equal access to aid, for example through tendering processes.

2 Mathias Dewatripont and Paul Seabright, *"Wasteful" Public Spending and State Aid Control*, 4 J. Eur. Econ. Ass'n 513-522 (2006).

3 To justify State aid control, note, however, that one must moreover argue that it adds value to the other competition law provisions in guaranteeing a proper functioning of the market.

4 Hans Friederiszick, Lars-Hendrik Röller and Vincent Verouden, *European State Aids Control: An Economic Framework*, *in* Advances in the Economics of Competition Law (P. Buccirossi ed., 2006).

In any case, I think that attempting to systematically trade off market failures and distortions of competition will not imply a revolution in State aid policy, but rather a (welcome) streamlining: for example, it will still favor Services of General Economic Interest, Small and Medium-Sized Enterprises (SME), innovation, training, poor regions, and environmental measures.

The big questions ahead are: Where will one draw the line? Will better targeted aid necessarily lead to less aid overall? My feeling is that, while economics can help give guidance concerning the determination of desirable relative aid levels, defining the tolerated overall level of State aid will remain pretty much pragmatic, or ad hoc, since economics does not offer any easy rational way to address this issue. Let us, for example, discuss this question in the context of aid to R&D&I.

WHILE ECONOMICS CAN HELP GIVE GUIDANCE CONCERNING THE DETERMINATION OF DESIRABLE RELATIVE AID LEVELS, DEFINING THE TOLERATED OVERALL LEVEL OF STATE AID WILL REMAIN PRETTY MUCH PRAGMATIC, OR AD HOC, SINCE ECONOMICS DOES NOT OFFER ANY EASY RATIONAL WAY TO ADDRESS THIS ISSUE.

IV. Aid to Research, Development, and Innovation

This is an area where economics has given particular attention to the tradeoff between market failures and distortions of competition. Going for an economic approach is, therefore, very natural here as a way to streamline the rules for granting aid.

Beyond this, aid to R&D&I is of course one area where there is a potential tension between the SAAP principle of less aid and the Lisbon strategy, which makes the case for giving preferential treatment to innovation. This case can easily be made, see for example the *Sapir Report*.[5] Better targeting, therefore, can resolve the tension by interpreting it to mean less aid overall but more aid to innovation.

However, it is important, in this Lisbon context, to be careful about the risk of the (re)emergence of an industrial policy captured by big incumbent firms. One way to avoid this is to focus aid to R&D&I on SMEs, which makes sense given that they are more subject to market failures (e.g., in credit markets) and that there is a lower risk of distortions. Moreover, when thinking about promoting European growth, one should keep in mind the key role of new firms in U.S. growth (which has crucially benefited from the ability of successful young firms to grow very fast). On the other hand, it is fair to acknowledge that European innovation may be less entrepreneurial than U.S. innovation. For example, the

5 André Sapir, Philippe Aghion, Giuseppe Bertola, Martin Hellwig, Jean Pisani-Ferry, Dariusz Rosati, José Vinals, and Helen Wallace, AN AGENDA FOR A GROWING EUROPE—THE SAPIR REPORT (2004).

Nokia of today did not start in a garage like Apple, but emerged from a (drastically restructured) large, diversified company. In this sense, ruling out State aid to large companies is not optimal. But it is then crucial to insist on good practices for aid provisions: accessibility to new entrants, limited duration provisions, investment in complementary inputs (like human capital). Current guidelines go in this direction, but doing it even more systematically would not hurt given the potentially excessive enthusiasm for innovation policies in our Lisbon era.[6]

V. Concluding Remarks

The need to streamline the existing case-by-case approach to State aid control was clear to most observers. Economics is a good way to attempt to do it, and trading off market failures with distortions of competition is the natural way to go. This can make policy choices more transparent (for example, by explicitly focusing on questions like: How much aid in total? How much aid for innovation?), and help focus attention to the consistency between policy initiatives, like the Lisbon strategy and the SAAP. Of course, there is a fair number of details still to fill in to make this economic approach operational. But, as Philip Lowe details in his article, progress is already being made.

6 For a discussion of these issues, *see* the Commentary of the State Aid Group of the Economic Advisory Group for Competition Policy (EAGCP), on the Draft Community Framework for State Aid R&D&I, *available at* http://ec.europa.eu/comm/dgs/competition/eagcp.pdf.

VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

EU State Measures against Foreign Takeovers: "Economic Patriotism" in All But Name

Alex Nourry and Nelson Jung

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

EU State Measures against Foreign Takeovers: "Economic Patriotism" in All But Name

Alex Nourry and Nelson Jung

Protectionist tendencies of EU Member States have always been evident in sectors of industry that previously were dominated largely by state-owned companies, namely, the banking, defense, energy, postal, telecommunications, transport, and water sectors where, following privatization, Member States have sought to preserve or create so-called "national champions." However, a new wave of protectionism or economic patriotism by Member States has broken out recently, as exemplified by the long-running saga of the takeover of Endesa, the Spanish electricity group. This article examines the compatibility of special rights and other state measures with the EC's single market objectives within the framework of the EC Treaty and their impact on foreign takeovers and investments. It also examines the initiatives taken by the Commission in order to eliminate such measures and analyzes the potential remedies available to foreign investors when confronted with such measures. Such remedies include the application of the Commission's powers under Article 21 of the EC Merger Regulation that arguably could be used to even greater effect with the abolition or curtailment of the two-thirds rule.

Alex Nourry is a partner and U.K. Head of the European Competition and Regulation group at Clifford Chance LLP. Nelson Jung is an associate at Clifford Chance LLP in the same group. The authors would like to thank their colleagues Dimos Dakanalis, Christos Nifadopoulos, and Trupti Reddy for their valuable contributions to this article.

I. Introduction

Protectionist tendencies of EU Member States have always been a concern to the European Commission in the context of the EU's single-market objectives. Such tendencies have been especially evident in sectors of industry that previously were dominated largely by state-owned companies, namely, the banking, defense, energy, postal, telecommunications, transport, and water sectors where, following privatization, Member States have sought to preserve or create so-called "national champions."

This protectionism has manifested itself either through very direct and blatant means such as the grant of State aid as in the case of the French government's rescue of Alsthom, or through the exercise of other state measures, including special rights in privatized companies or the application of legislative or regulatory powers. These state measures, although now less prevalent as a result of Commission intervention, have tended to deter and, in some cases, have prevented foreign takeovers.

A new wave of protectionism or economic patriotism by Member States has broken out recently, as exemplified by the long-running saga of the takeover of Endesa, the Spanish electricity group. Endesa, having been, originally, the target of a proposed takeover by its Spanish rival, Gas Natural, in a clear attempt by the Spanish government to create a national champion in the energy sector, is also the subject of a proposed takeover by the German energy group, E.ON. While E.ON's proposed acquisition of Endesa received unconditional clearance from the Commission under the EC Merger Regulation (ECMR)[1] on the basis that the two groups do not have competing activities, the Spanish National Energy Commission (CNE) approved the merger on condition that E.ON sell up to 30 percent of Endesa's generation capacity on security grounds, thereby severely undermining the viability of the merger. The Commission reacted very swiftly and strongly to the CNE ruling, requesting that the Spanish government refrain from its protectionist obstruction of E.ON's takeover of Endesa.[2] Meanwhile, Iberdrola, another Spanish utility, also joined the fray by appealing against the Commission's clearance of the E.ON-Endesa deal, motivated by its agreement

1 Council Regulation 139/2004 of 20 January 2004 on the control of concentrations between undertakings, 2004 O.J. (L 24) 1.

2 On the same day of its challenge to the Spanish government's obstruction of E.ON's takeover of Endesa, the Commission also called on the Italian government to explain its move to block the planned merger between Italian highway operator, Autostrade and Spain's Abertis. The Italian government had argued that the merger would be in breach of an Italian law that prohibits construction companies from holding shares in a concessionary. However, the Commission cleared the merger between Albertis and Autostrade on 22 September, 2006, and Italy's Infrastructure Minister Antonio di Pietro has, in the meantime, acknowledged that the Commission's clearance decision overrules the Italian law. *See* European Commission, Press Release No. 1244 (2006).

with Gas Natural to buy up to EUR 9 billion of excess assets following a merger with Endesa.[3]

The takeover battle for Endesa also highlights the inherent limitations of the Commission's exclusive jurisdiction under the ECMR to examine mergers between largely domestic players, even when such mergers may have a substantial impact on the single market. The Commission was unable to assert its jurisdiction over Gas Natural's proposed takeover of Endesa because of the ECMR's so-called "two-thirds rule," as both parties derived over two-thirds of their respective turnover from activities within Spain. As a result, it fell to the Spanish authorities to review the proposed merger, which they ultimately approved subject to certain conditions.

At the time, in a paper to her fellow Commissioners, European Commissioner Neelie Kroes was reported to have called for wide-ranging reform of the EC's merger regime and, in particular, the abolition of the two-thirds rule on the grounds that it no longer reflected "an optimal allocation of competence between the national and the Community level, and even constitutes in some instances, an obstacle to a consistent treatment of cases."[4] The paper also pointed to the merger of E.ON and Ruhrgas, the German energy group, as another merger that the Commission, rather than the national competition authority, should have examined. Although that merger was blocked by the Bundeskartellamt, the German competition authority's decision was overruled by the German government.

The Commissioner's paper also raised the issue of other sectors, in particular, the financial services sector, where some of the largest mergers also fell outside the Commission's competence due to the application of the two-thirds rule, citing the takeover of Credit Lyonnais by Credit Agricole and of Paribas by BNP as examples in the French banking sector.

This article examines the compatibility of special rights and other state measures (other than State aid) with the EC's single market objectives within the framework of the EC Treaty and their impact on foreign takeovers and investments. It also examines the initiatives taken by the Commission in order to eliminate such measures and analyzes the potential remedies available to foreign investors when confronted with such measures. Such remedies include the application of the Commission's powers under Article 21 of the ECMR that arguably could be used to even greater effect with the abolition or curtailment of the two-thirds rule.

3 This issue of *Competition Policy International* went to print on October 2, 2006. The authors, therefore, were not able to discuss new developments after that date.

4 Tobias Buck, *Kroes calls for more powers over mergers*, FIN. TIMES, Nov. 16, 2005.

II. The EU Legal Framework

A. FUNDAMENTAL PRINCIPLES OF EU LAW

The single or internal market freedoms of free movement of goods (Article 28 of the EC Treaty), services (Article 49 EC), capital (Article 56 EC), and right of establishment (Article 43 EC) enshrined in the Treaty constitute fundamental principles of EU law. Over the past decade, the principles governing the application of the Treaty freedoms have converged to a considerable extent. Essentially, any measure by a Member State that is liable to hinder or make less attractive the exercise of any of the freedoms is likely to be in breach of the Treaty unless it can be justified. The range of grounds on which Member States can validly rely to justify such a measure depends on whether the measure constituting a hindrance to the single market discriminates on the grounds of nationality (either directly or indirectly) or if it is indistinctly applicable. Discriminatory measures can be justified only by the limited and narrowly construed grounds set down explicitly in the Treaty, such as public policy, public security, or public health.[5] Non-discriminatory measures, however, are justifiable by overriding requirements in the general interest (a non-exhaustive list of such requirements having been developed by the EC Courts) provided that they comply with the principle of proportionality. Restrictions on the Treaty freedoms can never be justified by purely financial or economic reasons.

State measures restricting foreign takeovers generally come within the ambit of the free movement of capital as they are liable to deter or dissuade cross-border capital transactions. Article 56 EC gives effect to the free movement of capital between Member States, and between Member States and third countries. To that end, it provides that all restrictions on the movement of capital between Member States as well as between Member States and third countries are prohibited as a matter of principle. According to settled case law,[6] Directive 88/361, together with the nomenclature annexed to it, may be used for the purposes of defining a capital movement.[7] By way of example, direct investment in the form of participation in an undertaking by means of shareholding or the acquisition of securities on the capital market constitute capital movements within the meaning of Article 56 EC.[8]

5 In this respect, *see, e.g.*, Articles 30, 46, 55, and 58(1)(b) of the EC Treaty.

6 Case C-222/97, Trummer and Mayer, 1999 E.C.R. I-1661, at paragraphs 20 & 21.

7 1988 O.J. (L 178) 5; *see also*, Communication of the Commission on certain legal aspects concerning intra-EU investment, 1997 O.J. (C 220) 15.

8 The explanatory notes in Council Directive 88/361 of 24 June 1988 for the implementation of Article 67 of the Treaty, Annex I, 1988 O.J. (L 178) 5, state that direct investment is characterized by the possibility of participating effectively in the management of a company or in its control.

Restrictions on acquisitions of controlling stakes in a domestic company by EU investors are also caught by the freedom of establishment as guaranteed in Article 43 EC. In fact, restrictions on the free movement of capital in the context of state measures preventing takeovers by EU investors are usually inextricably linked to the right of establishment.[9]

It should be noted that Article 295 EC, which lays down the Treaty's neutrality between public and private ownership, does not have the effect of exempting the Member States' system of property ownership from the ambit of the Treaty freedoms. Thus, while Member States are not obliged to privatize state-owned companies, once a company is privatized, there is only limited scope for intervention. Member states can also not easily escape the application of the freedoms by invoking security concerns. Article 296 EC sets out the conditions under which Member States may legitimately invoke the protection of their essential security interests, in which case the freedoms do not apply.

B. THE COMMISSION'S EXCLUSIVE JURISDICTION OVER MERGERS WITH A COMMUNITY DIMENSION

Whether a transaction falls within the ambit of the ECMR and, as a consequence, is subject to the exclusive jurisdiction of the Commission, will play a crucial role in determining the powers that the Commission will have at its disposal and the effectiveness of any remedies available to a bidder when faced with state measures which restrict or seek to prevent a foreign takeover. Article 21(3) ECMR provides that no Member State shall apply its national legislation on competition to any concentration that has a Community dimension (i.e., mergers that meet the turnover thresholds set out in the ECMR).

Where the concentration significantly affects competition in a distinct market within a Member State, however, that Member State can request a referral back to its competent authorities under Article 9 ECMR. The notifying parties can also ask for a concentration with a Community dimension to be reviewed at national level by making a submission to the Commission under Article 4(4) ECMR prior to notifying.[10]

Apart from these mechanisms, whereby the whole or part of a concentration may be referred to the national competition authorities on request, Member States cannot intervene in a merger that is subject to the Commission's jurisdic-

9 Special rights in privatised companies in the enlarged Union - a decade full of developments, Annex 1 (Commission Staff Working Document, Jul. 22, 2005).

10 Article 4(4) of Council Regulation 139/2004 of 20 January 2004 on the control of concentrations between undertakings (hereinafter ECMR), 2004 O.J. (L 24) 1, requires the notifying parties to make a reasoned submission to the effect that the concentration may significantly affect competition in a market within a Member State which presents all the characteristics of a distinct market and should therefore be examined, in whole or in part, by that Member State.

tion, unless they can successfully invoke the protection of their "legitimate interests" under Article 21(4) ECMR. Examples of legitimate interests include public security, plurality of the media, and prudential rules. Legitimate interests grounds are interpreted narrowly given that they constitute exceptions to the Commission's exclusive jurisdiction to vet mergers with a Community dimension.

However, there is a limitation to the Commission's exclusive jurisdiction where each of the parties to a merger achieves more than two-thirds of its turnover in the same Member State. Therefore, where the two-thirds rule is met, mergers are not subject to the Commission's scrutiny, but are reviewed instead by national competition authorities, irrespective of what impact they might have on the single market.

There is a limitation to the Commission's exclusive jurisdiction where each of the parties to a merger achieves more than two-thirds of its turnover in the same Member State. Therefore, where the two-thirds rule is met, mergers are not subject to the Commission's scrutiny, but are reviewed instead by national competition authorities, irrespective of what impact they might have on the single market.

III. The Legal and Economic Impact of Special Rights

A. WHAT ARE SPECIAL RIGHTS AND HOW ARE THEY PERCEIVED BY THE COMMISSION?

Special rights, often referred to as "golden shares," are measures used by Member States to retain control over privatized companies, usually to prevent them from being taken over or to prevent the companies' management from taking actions which are contrary to national government policy for the sector in which they operate. While the former constitutes a direct restriction on investment, the latter will only indirectly affect investment decisions by making the investment potentially less attractive.

Golden shares are one of the most commonly used types of special rights, enabling the government to veto specific events or changes in the company's structure. They are usually enshrined in the articles of association of the company and cannot be changed without the government's consent. Special rights may also be conferred on governments by legislation, either under a general framework law covering several economic sectors or specific legislation aimed at an economic sector or a company. Governments may also seek to assert special rights over companies that are awarded concession contracts to provide services of general interest (e.g., in the gambling and broadcasting sectors).

Typically, Member States justify special rights on the grounds that they are necessary to achieve certain policy objectives, usually of a public-interest nature.

However, by restricting (directly or indirectly) cross-border mergers or investments, special rights are liable to infringe the principles of free movement of capital (Article 56 EC) and right of establishment (Article 43 EC) and as such can only be compatible with EU law if they meet the strict criteria that have been laid down by the European Court of Justice (ECJ).

In the 2003 paper, *Capital Movements in the Legal Framework of the Community*,[11] the Commission was critical of the margin of appreciation that Member States had at their disposal on the basis of the Treaty, and was concerned over restrictive practices by the Member States against non-EU bidders targeting companies in the European Union. The paper described the situation as amounting to the creation of a "fortress Europe," hostile to outside investment. The Commission therefore focused its capital movements paper on international trade and the obligations of the Member States towards the Organization for Economic Co-operation and Development and other international organizations.

The Commission is responsible for monitoring the proper and timely application of Treaty rules governing the freedom of capital movements as well as other internal market freedoms. Its objective is to secure the removal of all remaining restrictions to the free movement of capital in the internal market through ongoing cooperation and dialogue with Member States. Where Member States fail to comply with their obligations under the Treaty, this responsibility will require the Commission to initiate infringement procedures against the Member States concerned.

In its staff working document, *Special rights in privatized companies in the enlarged Union—a decade full of developments*,[12] the Commission outlined three principal concerns about special rights, being that they can:

- hinder privatized companies from achieving the full benefits of privatization;
- distort market driven cross-border activity in terms of both direct and portfolio investment in privatized companies; and,
- prove one of the obstacles to achieving a level playing field in the EU market for corporate control.

The Commission also raised concerns about the strong economic implications that special rights can have for the functioning of the single market since the companies in which Member States retain special rights often play a significant

11 *Capital Movements in the Legal Framework of the Community, in* The EU Economy 2003 Review 320, 320-28 (Directorate General for Economic and Financial Affairs, 2003).

12 Special rights in privatised companies in the enlarged Union - a decade full of developments, Annex 1 (Commission Staff Working Document, Jul. 22, 2005).

role in the economy. It identified 141 companies operating in sectors ranging from telecommunications, energy, and postal services to banking and insurance in which Member States retained special rights. Of these 141 companies,[13] 14 percent were accounted for by the fifteen old Member States (EU 15) and 86 percent by the ten new Member States (EU 10).

While special rights are still widely present in the EU 10 and affect a broader range of industries, including alcoholic beverages, food processing, textiles, and pharmaceuticals, in the Commission's view, this is largely due to the timing and scale of the privatization process in these countries and, in most cases, the Member States concerned have put in place a process to deal with the issue. In other cases, the Commission believes the special rights can be justified by exceptions provided in the Treaty or are structured in a way that is compatible with EU law. Overall, the Commission's assessment is that special rights in privatized companies are being phased out and that this can be attributed to two key factors:

- the proactive approach the Commission has taken in engaging in a dialogue with Member States, and
- the impact of the rulings of the ECJ in a series of landmark cases.

The Commission also identified the development of a more robust regulatory environment in specific sectors (e.g., telecommunications and energy) as another relevant factor both at the EU and national level, which has allowed Member States to protect the services of general interest in a relatively less restrictive manner. However, where special rights still persist, whether in the EU 15 or EU 10, that are not compatible with EU law, the Commission will not hesitate to proceed with infringement proceedings against the Member States concerned.

The latest report published by the Commission in November 2005,[14] assesses the microeconomic impact of special rights in privatized EU companies on company performance, investment in the companies, and share values. The report, however, does not address the broader impacts of special rights on economic performance in the industry sectors concerned, or on consumers, or on the European Union as a whole.

In its conclusions, the report found that:

- special rights held by public authorities tend to have a negative impact on the longer-term economic performance of EU privatized companies;

13 This number includes a number of regional companies, including 29 regional waterworks in the Czech Republic and 21 energy companies in Hungary.

14 OXERA, SPECIAL RIGHTS OF PUBLIC AUTHORITIES IN PRIVATISED EU COMPANIES: THE MICROECONOMIC IMPACT (2005) (report prepared for the European Commission).

- both existing and new empirical research provide strong evidence that special rights can constitute important barriers to direct investment;
- special rights have an adverse impact not only on the market for corporate control, by restricting takeover activity and distorting the level playing field in the market, but also on portfolio investors, who would otherwise benefit from increases in the value of their shares during a takeover;
- to the extent that special rights restrict the free movement of capital across EU borders, they impede further financial market integration; and
- regulation may be seen as a potentially less restrictive and more transparent means of achieving public policy objectives, especially if carried out by an arm's-length regulatory authority.

On the basis of this report, it is likely that the Commission will continue and possibly intensify its efforts against special rights and may try to steer national governments more proactively towards entrusting independent regulatory bodies with the protection of public interest concerns rather than resort to special rights or other similar state measures.

B. CONDITIONS FOR COMPATIBILITY OF SPECIAL RIGHTS WITH EU LAW

In its July 2005 document, the Commission regarded the rulings delivered by the ECJ in a series of seven infringement cases initiated by the Commission pursuant to Article 226 EC, as the most significant development in the field of special rights. In these cases, the Court set down very strict criteria for the use of special rights and their compatibility with EU law.

The ECJ's rulings[15] essentially confirm that special rights can be compatible with EU law even when they may be at odds with the principles of free movement of capital (Article 56 EC) and right of establishment (Article 43 EC) provided they are:

- justifiable, based on the exceptions explicitly listed in the Treaty or by reference to requirements in the general interest (where this constitutes a genuine and serious threat);
- necessary to protect the interests concerned and this protection cannot be obtained by less restrictive measures (i.e., proportionate);

15 Case C-58/99, Commission v. Italy, 2000 E.C.R. I-3811; Case C-367/98, Commission v. Portugal, 2002 E.C.R. I-4731; Case C-483/99, Commission v. France, 2002 E.C.R. I-4781; Case C-503/99, Commission v. Belgium, 2002 E.C.R. I-4809; Case C-98/01, Commission v. United Kingdom, 2003 E.C.R. I-04641; Case C-463/00, Commission v. Spain, 2003 E.C.R. I-4581; Case C-174/04, Commission v. Italy, 2005 E.C.R. I-4933.

- objective, non-discriminatory, and transparent; and
- subject to a legal remedy.

The ECJ also confirmed that the Treaty provisions on the free movement of capital do not draw a distinction between private undertakings and public undertakings.[16]

The cases concerned proceedings against Belgium, France, Portugal, Spain, the United Kingdom, and two cases against Italy. They all essentially concerned veto rights in favor of the state relating to the ownership of shares or capital of privatized companies. In all of the cases, with the exception of the case against Belgium, the ECJ ruled in favor of the Commission. While the Commission considers the ECJ's decisions in all seven cases as landmark rulings, those in the cases against Belgium, France, and Portugal delivered in 2002 stand out in particular. However, the cases against Italy, Spain, and the United Kingdom are also briefly examined, as well as the current status of pending cases and ongoing Commission intervention.

1. Portugal

The case against Portugal[17] concerned a general framework law relating to the privatization of undertakings in the banking, insurance, energy, and transport sectors that specified maximum levels of foreign participation (ranging between 5 and 40 percent) in the privatized companies. The ECJ found the law to constitute a restriction on the free movement of capital within the meaning of Article 56 EC. The Portuguese government sought to justify the law on the grounds that it was necessary for the pursuit of national economic policy objectives, for choosing a strategic partner, for strengthening the competitive structure of the market concerned, or modernizing and increasing the efficiency of means of production. However, none of these objectives were found by the ECJ to constitute a valid justification for restrictions on the fundamental Treaty freedoms.

2. France

The case against France[18] concerned legislation that vested in the state a golden share in Société Nationale Elf-Aquitaine, and gave the Minister of Economic Affairs the right to:

- approve any direct or indirect shareholding by a natural or legal person, acting alone or in conjunction with others, in excess of one-fifth, one-tenth or one-third of the capital of, or voting rights in, the company, and

16 Case C-174/04, Commission v. Italy, 2005 E.C.R. I-4933.

17 Case C-367/98, Commission v. Portugal, 2002 E.C.R. I-4731.

18 Case C-483/99, Commission v. France, 2002 E.C.R. I-4781.

- oppose any decision to transfer or use as security the majority of the capital of four subsidiaries of that company.

In addition, the law provided for the appointment of two state representatives on the board of directors of the company, without entitlement to vote.

The ECJ ruled that the French law was incompatible with the free movement of capital, despite the fact that it applied without distinction to French nationals and to nationals of other Member States. It held that:

> "even though the rules in issue may not give rise to unequal treatment, they are liable to impede the acquisition of shares in the undertakings concerned and to dissuade investors in other Member States from investing in the capital of those undertakings. They are therefore liable, as a result, to render the free movement of capital illusory."[19]

The French government's justification for adopting the law was the safeguarding of supplies of petroleum products in the event of a crisis. The ECJ accepted that, in principle, this objective could be a valid public interest justification for the French law. However, the provisions of the law were too wide and "investors concerned were given no indication whatsoever as to the specific, objective circumstances in which prior authorization would be granted or refused."[20] Therefore, the legislation went beyond what was necessary in order to attain the objective indicated and constituted an unjustifiable restriction of the free movement of capital.

3. Belgium

The Belgium case[21] also involved legislation whose objective was to secure energy supplies at a time of crisis. Belgium had passed laws vesting in the state a golden share in, inter alia, Distrigaz. The law provided that:

- advance notice of any transfer, use as security or change in the intended destination of the company's system of lines and conduits that are used, or are capable of being used, as major infrastructures for the domestic conveyance of energy products must be given to the Minister

19 *Id.* at paragraph 41.

20 *Id.* at paragraph 50.

21 Case C-503/99, Commission v. Belgium, 2002 E.C.R. I-4809.

responsible, who can oppose such operations if he considers that they adversely affect the national interest in the energy sector, and

- the Minister may appoint two state representatives to the board of directors of the company who can propose to the Minister the annulment of any decision of the board of directors which they regard as contrary to the guidelines for the country's energy policy, including the government's objectives concerning the country's energy supply.

The ECJ found that the measures were restrictive but allowed them on the basis that they were justified, as being necessary to attain the stated objective (to maintain minimum supplies of gas in the event of a real and serious threat) and because there were no less restrictive measures to attain the general interest objective. In particular, the ECJ found that the measures were acceptable because:

- the law established a system of opposition not prior approval (see the France case discussed above) and the public authorities were obliged to adhere to strict time limits when exercising the right to oppose an acquisition;
- the right was limited to certain decisions in relation to certain strategic assets of Distrigaz (lines and conduits); and,
- any intervention by the Minister had to be supported by a formal statement of reasons and was subject to judicial review by the courts.

Cumulatively, these three factors meant that the law did not grant a wide discretion to the state to intervene and established objective criteria on which the state could do so.

4. Italy

The first decision against Italy, in 2000, concerned investment restrictions contained in a 1992 general framework law on privatization that, although amended by the Italian government in 2001, the Commission still considered to unduly restrict the freedoms of capital movements and establishment, and therefore referred it to the ECJ.

The second decision against Italy, in 2005, concerned a law of 2001 aimed at avoiding anticompetitive attacks on Italian companies operating in the electricity and gas sectors by public entities operating in the same sectors in other Member States, and enjoying a dominant position in their domestic markets. The law was also designed to safeguard energy supplies. In rejecting the Italian government's justifications for the law, the ECJ said that Italy had failed to show how energy supplies would be threatened by the acquisition of an Italian energy company specifically by a buyer dominant in another Member State as opposed to any other type of buyer. The ECJ also ruled that "an interest in generally strengthening the competitive structure of the market in question cannot constitute valid justification for restrictions on the free movement of capital"[22] and, in any case, that objective could be served by the ECMR. Following the judgment, the Commission called on Italy to comply with the ECJ's ruling as it was not convinced that the amendments subsequently made to the law fully implemented the ruling of the Court.[23]

5. Spain

The decision against Spain in 2003 concerned a law of 1995 relating to the privatizations of Repsol SA, Telefónica de España SA, Telefónica Servicios Móviles SA, Argentaria, Tabacalera SA, and Endesa SA., that gave the Spanish government a right of prior approval of certain management decisions.[24] In rejecting the measures, the ECJ ruled that the fact that the legislation concerned introduced regimes that would last only ten years did not make the measures proportionate since an "infringement of Treaty obligations does not cease to be an infringement merely because it is limited in time."[25] Following the ECJ's ruling in 2003, it was not until May 2006 that the Spanish government fully eliminated the restrictions concerned, after the Commission had called on Spain to comply with the ECJ's judgment, and had sent Spain a reasoned opinion as part of infringement proceedings under Article 228 EC.

6. United Kingdom

The decision against the United Kingdom, concerned a golden share held by the U.K. government in British Airports Authority plc (BAA), that limited all interests in the company to 15 percent of voting shares and provided the U.K. government with a veto right over the disposal of shares. Following the ECJ's ruling in 2003, the U.K. government relinquished its golden share, that otherwise might have prevented the recent takeover of BAA by Spain's Ferrovial.[26]

7. Pending Cases

Infringement cases concerning golden shares are also pending before the ECJ against Germany and the Netherlands. The case against Germany concerns the 1960 law privatizing Volkswagen (VW) that to date still prevents any sharehold-

---

22 Case C-174/04, Commission v. Italy, 2005 E.C.R. I-4933, at paragraph 37.

23 European Commission, Press Release No. 439, Free movement of capital: Commission calls on Italy to apply Court of Justice ruling on the law on investment in energy companies (2006); Eur. Commission, Press Release No. 1270, Free movement of capital: Commission calls on Italy to modify law on privatised companies and to apply Court ruling on investment in energy companies (2005).

24 Case C-463/00, Commission v. Spain, 2003 E.C.R. I-4581.

25 *Id.* at paragraph 81.

26 Case COMP/M.4164, Ferrovial/Quebec/GIC/BAA, 2006 O.J. (C 182) 11.

er from acquiring more than 20 percent of voting rights, and confers a special blocking minority right on any shareholder who has 20 percent of voting rights. Traditionally, both the German government and the Land of Lower Saxony (the Land) held 20 percent voting rights in VW, although the Land is now the only shareholder with 20 percent voting rights and two mandatory members of the board. The Commission brought the case before the ECJ in 2004[27] claiming that these provisions of the VW Act make it substantially less attractive for other EU investors to acquire the company's shares with a view to participating effectively in management decisions or controlling it, and so are contrary to the Treaty provisions on free movement of capital and the right of establishment.

The two cases against the Netherlands, initiated by the Commission pursuant to Article 226 EC, concern golden shares held by the Dutch government in Koninklijke KPN N.V. (KPN) and TNT Post Groep N.V. (TPG). Conferring a major influence over KPN's and TPG's financial decision-making and the management of the two companies, the Commission took the view that the golden shares may deter EU investors from investing in the capital of the two companies and, consequently, were contrary to the Treaty rules on free movement of capital and the right of establishment.[28] On April 6, 2006, Advocate General Poiares Maduro came to the conclusion that the Netherlands has indeed failed to fulfill its obligations under Article 56 EC by retaining its golden shares in KPN and TPG and this view was endorsed by the ECJ in its judgment of September 28, 2006.[29]

The Commission continues to monitor compliance with the Treaty provisions governing free movement of capital; and in this context is actively considering infringement procedures against a number of Member States (some of which have already been mentioned), including Denmark (possible obstacles to investment from other Member States in Copenhagen Airports); France (authorization procedure for foreign investments in certain sectors and ban on stock market listing for football clubs);[30] Hungary (privatization framework law); Luxembourg (veto rights over shareholdings in the satellite companies SES Astra and SES Global); and Spain (law amending functions of Spanish energy regulator).

27 European Commission, Press Release No. 1209, Free movement of capital: Commission takes Germany to the Court of Justice on Volkswagen law (2004).

28 European Commission, Press Release No. 1753, Free movement of capital: Commission takes the Netherlands to Court of Justice on special powers in KPN and TNT (2003).

29 Cases C-282/04 & C-283/04, Commission v. Netherlands (not yet reported).

30 France recently lifted its ban on the listing of sports clubs. *See* Adam Jones, *Olympique Lyonnais Leads Listing Field,* Fin. Times, Sept. 25, 2006.

IV. The New Wave of Interventionism

A. "ECONOMIC PATRIOTISM" IN ALL BUT NAME

Based on the Commission's own assessment special rights, or golden shares, are being phased out by Member States, including among the EU 10. This probably also coincides with the fact that privatization programs in a number of Member States have now largely run their course.

The Commission appears disposed to encouraging Member States to entrust independent regulatory bodies with the protection of public interest concerns rather than resort to special rights or other interventionist measures. While Member States may be receptive to such an approach, possibly with the exception of certain sectors, such as defense,[31] this does not necessarily guarantee that Member States will refrain from interventionism in the pursuit of national interests.

RECENTLY, A NEW WAVE OF INTERVENTIONISM BY MEMBER STATES HAS EMERGED. THIS HAS MANIFESTED ITSELF IN THE ADOPTION OR EXERCISE OF LEGISLATIVE AND REGULATORY MEASURES, OR THE BLATANT PROMOTION OF NATIONAL CHAMPIONS IN PREFERENCE TO A FOREIGN TAKEOVER.

Indeed, recently, a new wave of interventionism by Member States has emerged. This has manifested itself in the adoption or exercise of legislative and regulatory measures, or the blatant promotion of national champions in preference to a foreign takeover. Such interventionism is well illustrated, for example, by the Spanish government's actions in relation to the contested takeover of Endesa or by the French government's defense of Danone following rumors that PepsiCo was preparing a takeover bid in a sector that few would credit as being strategic or in the general interest, but nevertheless Danone was deemed by the government "a French icon and off-limits to foreign ownership."[32]

While such government interventions may be perceived to be in the national interest, in the long run they are unlikely to be in the best interests of domestic consumers since they are prone to result in higher prices and lower innovation. They are also undesirable from an internal market perspective as they create an unlevel playing field for business and ultimately are more likely to undermine rather than enhance the competitiveness of EU industry.

31 For example, it is unlikely that the U.K. government would relinquish its golden shares in BAE Systems and Rolls-Royce or France its golden shares in Thales and EADS.

32 Arturo Bris, *Global Growing Pains,* Fin. Times, April 6, 2006.

B. LEGISLATIVE AND REGULATORY MEASURES

Controversially, in December 2005, France adopted a law[33] creating an authorization procedure for foreign investments in certain sectors that could affect public policy, public security, or national defense. The protected sectors include gambling (e.g., casinos); private security services; research, development, or production of chemical or biological antidotes; activities concerning equipment for intercepting communications or eavesdropping; services for evaluation of security of computer systems; dual-use (civil and military) technologies; cryptology; activities of firms that are repositories of defense secrets; research, production, or trade in arms, munitions, explosives, or other military equipment; or any other industry supplying the defense ministry any of the aforementioned goods or services.[34]

On April 4, 2006, the Commission sent a formal request to France, after informal dialogue that led the Commission to believe that the French law could potentially lead to restrictions in the freedom of movement of capital (Article 56 EC) and the right of establishment (Article 43 EC).[35] Although the French law appeared to make wider use of an independent regulator for the purposes of protecting the public interest (along the lines of the proposals contained in the Commission's report of November 2005),[36] the Commission was concerned that the ambit of the powers conferred on the French regulator was considerably wider than appeared necessary for achieving the stated objectives of public policy, public security, and national defense and consequently not proportionate.

The Commission also challenged the inclusion of casinos in the category of companies for which there is public policy concern, especially as the transposition of the Money Laundering Directive could provide adequate public protection. The Commission is therefore likely to require France to amend its legislation, failing which it will be forced to pursue infringement proceedings against France under Article 226 EC.

In February 2006, Spain adopted new legislation extending the powers of the CNE, Spain's National Energy Commission, requiring the authorization of the CNE for the acquisition of over 10 percent of the share capital, or any other percentage resulting in significant influence, in companies engaged directly or indirectly in regulated activities in the energy sector. In exercising such powers, the CNE is able to take into consideration a variety of factors, including any risks in relation to the regulated activities, the inability to perform such activities, and the protection of the general interest and reasons of public security. In the

33 Decree 2005-1739 of 30 December 2005, O.J., December 31, 2005, No. 45.

34 *See id.*

35 European Commission, Press Release No. 438, Free movement of capital: Commission scrutinises French law establishing authorisation procedure for foreign investments in certain sectors (2006).

36 *See* Oxera, *supra* note 14.

Commission's view, these factors are vague and indeterminate and as a result give the CNE wide discretionary powers. In May 2006, the Commission announced that it had opened infringement proceedings under Article 226 EC by requesting in a formal letter that Spain provide more information concerning the new legislation, having already warned the Spanish government that the law was likely to violate EU law, insofar as it hindered or rendered less attractive the free movement of capital and the right of establishment.[37]

The Spanish legislation is particularly sensitive given that it was adopted in the context of competing takeover bids for the Spanish electricity operator, Endesa. On February 21, 2006, the German energy company E.ON announced its intention to launch a public bid for Endesa, countering the bid that Spain's Gas Natural had made for the company in September 2005. Just three days after E.ON's announcement (on February 24, 2006), Spain adopted the law[38] extending the CNE's powers to allow it to block foreign takeovers. This move was widely criticized as a blatant attempt by the Spanish government to deter E.ON from pursuing its bid. Nevertheless, E.ON notified its proposed takeover of Endesa to the Commission on March 16, 2006, and this was cleared unconditionally on April 25, 2006. The CNE subsequently also approved E.ON's proposed takeover of Endesa, on July 28, 2006, but subject to substantial conditions, including, most significantly, the sale of 30 percent of Endesa's domestic generation capacity. In light of the CNE's decision, the Commission announced that it would take the appropriate measures to ensure that its clearance decision in *E.ON/Endesa*[39] was respected and, accordingly, on August 3, 2006, sent the Spanish government a letter based on Article 21 ECMR, requesting "clarifications" regarding the CNE's decision.[40] In its reply of August 10, 2006, the Spanish government accused the Commission of exceeding its authority and countered that the conditions imposed on E.ON do not prevent it from acquiring Endesa. According to the Commission's preliminary conclusions, however, most of these conditions raise serious doubts as to their compatibility with EC law. Competition Commissioner Neelie Kroes vowed to remain vigilant in this case and to continue to take a firm stance in similar cases.[41] Not surprisingly, therefore, on September 26, 2006, the Commission announced that it had

37 European Commission, Press Release No. 569, Free movement of capital: Commission opens infringement proceedings against Spain regarding law amending functions of Spanish electricity and gas regulator (2006).

38 Act of 24 February by which the Functions of the National Commission of Energy are Modified (B.O.E., 2006, 3436).

39 Case COMP/M.4110, E.ON/Endesa, 2006 O.J. (C 68) 09.

40 European Commission, Press Release No. 1096, Statement by Commission Spokesperson: Receipt of letter from Spanish authorities on Eon/Endesa merger (2006).

41 See Competition Commissioner Neelie Kroes' speech on "Cross-border mergers and energy markets" given at the Villa d'Este Forum in Cernobbia, Italy, on 2 September 2006, Speech/06/480.

adopted a decision finding that Spain has breached Article 21 ECMR by virtue of the conditions imposed by the CNE on E.ON's bid for Endesa.[42] The Commission also announced that it had sent Spain a reasoned opinion alleging that the legislation that extended the power of the CNE to authorize mergers in the energy sector infringes Articles 43 and 56 EC. In addition, both E.ON and Endesa have appealed the CNE's decision.[43]

C. PROMOTING MERGERS TO CREATE NATIONAL CHAMPIONS

Member States are not averse to resorting to even more blatant industrial engineering when they cannot rely on legislative or regulatory measures in order to create national champions.

When Italy's Enel expressed interest in the French energy company Suez, the French government orchestrated a defensive merger between Suez and Gaz de France in February 2006. President Jacques Chirac summarized his opposition to Enel's bid for Suez by declaring: "France doesn't want to surrender to a purely financial operation."[44] Although France was cleared in May 2006 of breaching EU internal market rules by promoting the defensive merger, it was criticized by the Commission for considering maintaining a golden share in the merged entity.[45] The Commission decided on June 19, 2006, that it would investigate the merger on the basis of competition concerns raised in the Belgian gas and electricity supply market.[46] It outlined its main concerns in a Statement of Objections issued on August 19, 2006,[47] and is expected to adopt a decision on the compatibility of the merger between Suez and Gaz de France with the common market on November 17, 2006.[48]

A national champion was created in the German *E.ON/Ruhrgas* case. The German Federal Cartel Office prohibited the acquisition of a majority stake in

42 European Commission, Press Release No. 1265, Mergers: Commission Rules Against Spanish Energy Regulator's Measures Concerning E.ON's Bid for Endesa (2006).

43 European Commission, Press Release No. 1264, Free Movement of Capital: Commission Calls on Spain to Modify the Law Amending the Functions of the Spanish Electricity and Gas Regulator (2006).

44 Bris, *supra* note 32.

45 *France to keep control over GdF/Suez assets*, Reuters, May 23, 2006.

46 European Commission, Press Release No. 802, Mergers: Commission opens in-depth investigation into merger between Gaz de France and Suez group (2006).

47 European Commission, Press Release No. 1109, Mergers: The European Commission Adopts a "Statement of Objections" Regarding the Merger Project Between Suez and Gaz de France (2006).

48 In view of remedies proposed by Suez and Gaz de France on September 20, 2006, the deadline for the Commission's decision was extended by 15 working days; MEMO/06/340.

Ruhrgas by E.ON in 2002 on the basis of competition concerns.[49] However, in spite of the merger being blocked by the Bundeskartellamt, the Federal Minister of Economics ultimately granted a special ministerial authorization and cleared the merger. This special authorization could, of course, not have been granted had the transaction triggered the Commission's exclusive jurisdiction under the ECMR. However, the Commission had reached the decision that it was not competent to review the case given that E.ON and Ruhrgas achieved more than two-thirds of their EU-wide turnover within Germany. The U.K.'s energy regulator, Ofgem, argued at the time that the merger should have been reviewed at EU level, claiming that E.ON's subsequent acquisition of U.K. energy company Powergen took it above the two-thirds limit. In Ofgem's view, E.ON had effectively side-stepped the Commission's exclusive jurisdiction over the merger. At the time, the E.ON/Ruhrgas merger attracted widespread criticism that it was incompatible with the spirit of competition and undermined the liberalization of the EU energy market.

D. STATE INTERVENTION IN THE CONTEXT OF ARTICLE 21 ECMR

The ability of Member States to successfully resort to interventionist measures is far more limited when it comes to takeovers with a Community dimension falling within the exclusive jurisdiction of the Commission under the ECMR. Member States do not have the power to intervene in mergers that have a Community dimension unless they can justifiably invoke legitimate interests (e.g., public security, media plurality, or prudential rules) within the meaning of Article 21(4) ECMR. To the extent that this provision constitutes an exception to the Commission's exclusive competence to vet mergers falling within the ECMR, the provision is to be narrowly construed, so that any measures invoked by Member States pursuant to Article 21(4) ECMR must be proportionate and compatible with EU law.

1. The *Champalimaud* Case

In *BSCH/Champalimaud*,[50] the acquisition by Banco Santander Central Hispanico (BSCH), a Spanish bank, of the Champalimaud group of Portuguese banks and insurance companies fell within the ECMR and was therefore notified to the Commission. However, the Portuguese Minister of Finance adopted a decision freezing Champalimaud's shares on the basis that the transaction failed to comply with Portuguese financial rules. While the Portuguese authorities claimed that their decision was based on prudential grounds, the Commission doubted that this was in fact the case and suspected the decision to be driven by protectionist considerations. Moreover, the Portuguese authorities had not communicated the interests they wanted to protect to the Commission in accordance with Article

49 *See* the Bundeskartellamt's press release relating to its decision to prohibit the acquisition *available at* http://www.bundeskartellamt.de/wEnglisch/News/Archiv/ArchivNews2002/2002_02_28.shtml.

50 Case IV/M.1616, BSCH/A.Champalimaud*,* 1999 O.J. (C 197) 5.

21(4) ECMR. As a consequence, the Commission initiated infringement proceedings and adopted an interim measures decision suspending the application of the Portuguese authorities' decision pending further investigation.

The Commission was not only concerned with an apparent attempt to undermine its exclusive jurisdiction under the ECMR, but also raised the issue whether the Portuguese decision violated the principles of free movement of capital and freedom of establishment. However, the Portuguese authorities refused to suspend their decision and instead suspended certain voting rights in the Champalimaud group. The Commission cleared the concentration in August 1999 and initiated further infringement proceedings against Portugal in September 1999 for not complying with the interim measures decision. On October 20, 1999, the Commission adopted a final decision declaring that Portugal had infringed Article 21 ECMR. Following the withdrawal of the Portuguese measures, BSCH went on to acquire control of Banco Totta & Açores, SA. and Banco de Crédito Prédial Português, both subsidiaries of the Champalimaud group.

2. The *Cimpor* Case

This case involved a proposed bid for the Portuguese cement company Cimpor, in which the Portuguese state held golden shares. In 2000, Holderbel, a Belgian subsidiary of Holderbank of Switzerland and Portuguese cement company, Secil, launched a public bid to acquire Cimpor. Given that the transaction had a Community dimension it was notified to the Commission under the ECMR. However, the Portuguese Minister of Finance prohibited the transaction. The Commission issued a decision finding that Portugal did not protect any legitimate interests under Article 21 ECMR, and that Portugal was acting in breach of the Commission's exclusive jurisdiction to review the concentration.[51] The Commission decided that Portugal should withdraw its prohibition measures in order to comply with Community law.

Portugal challenged the Commission's decision claiming that the Commission had no legal basis for it. Portugal had not made a request for the protection of legitimate interests under Article 21 ECMR and argued that non-compliance with the procedure in Article 21 was not sufficient for the Commission to reach its infringement decision. Furthermore, Portugal claimed that, if anything, the Commission should have opened an infringement procedure against Portugal under Article 226 EC rather than relying on Article 21 ECMR to reach its decision. However, in its ruling of June 22, 2004, the ECJ dismissed Portugal's appeal on the basis that the Commission acted legitimately in making a decision under Article 21 ECMR.[52]

51 Case COMP/M.2054, Secil/Holderbank/Cimpor, 2000 O.J. (C 198) 5. The parties subsequently withdrew their notification.

52 Case C-42/01, Portugal v. Commission, 2004 E.C.R. I-6079; for an analysis of the implications of the decision in terms of remedies, *see infra*, Section V.

3. The Polish Banking Case

In March 2006, the Commission began infringement proceedings against Poland under Article 21 ECMR. The Commission considered that Poland had breached the Commission's exclusive jurisdiction to review mergers with a Community dimension by requiring the bank UniCredit to divest its shares in the Polish bank BPH, the acquisition of which had already been approved by the Commission as part of UniCredit's takeover of German bank HVB.[53] The Polish Treasury instructed UniCredit to sell its shares in BPH on the basis that UniCredit was bound by a non-compete clause in a privatization agreement it had entered into when it acquired the Polish bank Pekao in 1999 from the Polish state. Poland insisted that this non-compete clause continued to prevent UniCredit, for a period of ten years, from opening subsidiaries and/or branches in Poland, acquiring control of banks active in Poland, or making any capital investment in any company active in the Polish banking sector. The Commission reminded Poland that Member States can neither apply their national competition law to concentrations with a Community dimension, nor can they adopt measures which could prohibit or prejudice (de jure or de facto) such concentrations unless they can rely on legitimate interests under Article 21(4) ECMR, and the specific measure is proportionate and compatible with EU law.

The Commission considered the Polish government's decision to invoke the non-compete clause to constitute a measure that can de facto prevent, or seriously prejudice, the *UniCredit/HVB* concentration. In addition, the Commission noted that Poland had not communicated to it any other hypothetical legitimate interests under Article 21(4) ECMR, and that, in any event, the non-compete clause itself appeared to be incompatible with the free movement of capital and the freedom of establishment.[54] The Commission emphasized that it could adopt a decision under Article 21 ECMR requiring the Polish government to refrain from invoking the non-compete clause. Moreover, the Commission stressed that such a decision would be directly applicable, meaning that it could be invoked directly before a national court or public authority in Poland by aggrieved third parties.

In April 2006, Poland entered into an agreement with UniCredit according to which UniCredit agreed to sell 200 of the 483 branches of Poland's BPH bank. The Commission, however, announced that this agreement did not necessarily mean an end to proceedings against Poland for violating Article 21 ECMR.

---

53 Case COMP/M.3894, Unicredito/HVB, 2005 O.J. (C 235) 4.

54 Note that the Commission started infringement proceedings under Article 226 EC with regard to a possible violation of Articles 43 and 56 EC. *See* Eur. Commission, Press Release No. 276, Free movement of capital: Commission opens infringement procedure against Poland in context of UniCredit/HBV merger (2006).

4. The Italian Banking Cases

The Italian central bank's handling of foreign takeover bids has also given rise to concerns recently. In one case the Commission cleared the public takeover bid by the Dutch banking group ABN Amro for Italian bank Antonveneta. However, the governor of Italy's central bank, who holds a personal veto over banking mergers, appeared to have favored a rival bid from the Italian bank BPI, and reportedly overruled a number of senior Italian regulators who expressed concerns about the viability of BPI's bid. After a long takeover battle, ABN Amro ultimately succeeded, becoming the first foreign bank to acquire an Italian financial institution, and the governor of the Italian central bank, Antonio Fazio, was forced to step down amid a criminal investigation into allegations of insider trading and abuse of office.

The Italian banking cases have sparked a wider drive by the Commission to combat protectionism in Europe's fragmented financial services sector. The Commission appears particularly keen to clarify the role of central banks and other financial regulators, that are tasked, among other things, with ensuring that cross-border mergers do not undermine the stability of domestic financial markets. The Commission argued that the present regime gives national institutions too much scope to obstruct foreign takeovers. To this end, on March 16, 2006, the Commission announced that it may look to overhaul rules for the policing of mergers in the banking, insurance, and securities sectors with the aim of reducing protectionism and harmonizing supervisory practices.[55] More recently, the Commission said it was going to table a proposal to change the banking directive to the effect that national supervisors will only be able to oppose a merger in this sector if one of the objective and non-discriminatory criteria set out in the directive is met.[56]

V. Remedies against State Measures Available to Foreign Investors

A. OVERRIDING POLITICAL AND COMMERCIAL CONSIDERATIONS

State measures that restrict foreign takeovers or investments and do not meet the conditions for compatibility with EU law laid down by the ECJ will render

55 Charlie McCreevy, EU Commissioner for Internal Market and Services, has commented that "the tendency by national regulators to encourage national champions" is one reason for the currently limited cross-border banking consolidation. Charlie McCreevy, *European Banking - challenges and changes ahead*, Address to Institut International de'Etudes Bancaires (May 20, 2005).

56 According to FIN. TIMES, Sept. 6, 2006, the criteria set out in the draft proposal are (i) the reputation of the acquirer and its ability to meet the standards set out in the banking directive, (i) the reputation and experience of the people charged with steering the merged group, (iii) the financial soundness of the acquirer, (iv) its ability to meet all the criteria and obligations laid out in the directive and other sectoral rules, and (v) suspected links to money laundering and terrorist financing.

Member States in violation of their obligations to comply with the Treaty rules on the free movement of capital (Article 56 EC) and the right of establishment (Article 43 EC).

Violations of the Treaty will in principle give rise to a range of potential remedies for a foreign bidder or investor at both the EU and national levels. The challenge, however, is whether any of these remedies can provide effective and meaningful redress within the context of a contested cross-border merger. Even to the extent that remedies are available that would afford such redress within an acceptable timeframe, aggrieved parties will inevitably have to face up to the political and commercial realities of having to challenge the government of a Member State with which it aspires to do business not only in the immediate future but possibly also over the longer term. These overriding considerations may, in practice, prove to be the greatest obstacle to foreign investors pursuing remedies against Member States, whether at the EU or national level.

B. REMEDIES AT THE EU LEVEL

An aggrieved investor's principal remedy at the EU level will be to lodge a complaint with the Commission against the Member State concerned, with a view to the Commission initiating one or more of the enforcement procedures available to it under either Article 86 or 226 EC. Where the transaction falls within the ambit of the ECMR, the Commission also has the additional option of taking enforcement action under Article 21 ECMR.

1. Article 86 EC

An aggrieved investor could challenge a state measure that it feels confers special rights on a company by complaining to the Commission. Article 86(3) EC imposes a duty on the Commission to ensure that Member States do not keep in place measures that are contrary to Article 86(1) EC,[57] and gives the Commission power to adopt decisions or directives that are legally binding. Failure to comply with such a decision or directive can be the basis for an action before the ECJ. However, while decisions and directives of the Commission are subject to challenge by the Member State or the company that is the subject of the state measure, a failure by the Commission to issue a decision or directive following an investigation cannot be challenged by a complainant because it is does not constitute a legally binding act.[58]

57 In essence, Article 86(1) EC is designed to prevent Member States from evading the competition (and other) rules of the Treaty by maintaining public ownership of undertakings or by granting undertakings special or exclusive rights.

58 *See* Case C141/02, Commission v. Max-mobil, 2005 E.C.R. I-1283.

It has been suggested that the Commission has the power to adopt interim measures under Article 86(3) EC, although it has never done so.[59] This is on the basis that the Commission's powers arising from Article 86(3) EC may be similar to its powers in relation Articles 81 and 82 EC, that allow it to adopt interim measures in urgent cases and on the basis of a prima facie finding of infringement.

2. Article 226 EC

The Commission can (on its own initiative or following a complaint) investigate and issue a reasoned opinion under Article 226 EC in respect of a Treaty violation by a Member State. If the Member State concerned fails to comply within the time limits laid down in the reasoned opinion, the Commission may bring an action before the ECJ. Once the ECJ upholds the Commission's opinion, and if the Member State fails to comply with the finding of the ECJ, the Commission can, after giving the Member State an opportunity to submit its observations, issue a reasoned opinion specifying the points on which the Member State has not complied with the ECJ's judgment. If, following the second reasoned opinion, the Member State fails to take the necessary steps within the time limit laid down, the Commission may then bring a further action before the ECJ, requesting the imposition of a penalty payment against the Member State under Article 228 EC.

It will be readily apparent that the enforcement procedure under Article 226 EC is a long and complicated one, that is not well-suited to the commercial realities of a contested takeover, as is well illustrated by the still unresolved VW case and KPN and TPG cases. Moreover, a final decision under Article 228 EC will be of little practical use to the parties to a prospective merger since they are likely to have been driven to abandon the deal well before the ECJ has issued its decision.

3. Article 21 ECMR

Where a proposed merger falls within the ambit of the Commission's exclusive jurisdiction under the ECMR, Article 21 ECMR probably offers the most effective and timely remedy for an aggrieved investor above any other remedy available at EU or national level.

For the first time in the *Champalimaud* case, and more recently in the *Cimpor* and *UniCredit* cases, the Commission initiated proceedings on the basis of Article 21 ECMR, requiring Member States to refrain from adopting measures (ostensibly on the grounds of legitimate national interest under Article 21(3) ECMR) that would amount to an infringement of the Commission's exclusive jurisdiction under the ECMR to vet mergers with a Community dimension.

In reaching an Article 21 ECMR infringement decision, the Commission initially forms a preliminary view as to the incompatibility of the Member State

59 *See* John Temple Lang, European Union legal rules on State measures restricting competition, Paper Presented at the 2005 Summit at Como (Oct. 2005) (on file with the author).

measures with Article 21 ECMR, and sends the Member State concerned a letter asking it to justify its actions. The Member State then has fifteen working days to reply to the Commission's initial findings. If, following this consultation period, the Commission still believes that Article 21 ECMR has been infringed, it can adopt a decision requiring the Member State to withdraw the infringing national measure. This decision is addressed to the Member State, is binding on it and, crucially, is directly applicable against the Member State to which it is addressed and, as such, is enforceable in the national courts by any party affected by the decision which, in this context, would include an aggrieved foreign investor. Moreover, the Commission may also adopt interim measures in the course of Article 21 ECMR proceedings (as it did in *Champalimaud* case, where it provisionally lifted Portugal's suspension of the merger).

In its 2004 ruling in *Portugal v. Commission*,[60] the ECJ crucially confirmed that Article 21 ECMR proceedings could be initiated by the Commission regardless of whether the Member State concerned had invoked legitimate national interests under Article 21(3) ECMR. In other words, the fact that a Member State had chosen not to follow the Article 21(3) ECMR procedure did not preclude the Commission from using Article 21 to secure the withdrawal of protectionist measures that are not justified or proportionate under EU law.

In its decision, the ECJ emphasized that:

> "if the Commission were reduced, in the absence of any communication by the Member State concerned [under Article 21 ECMR] to the sole option of bringing an action for failure to fulfill obligations under Article 226 EC, it would be impossible to obtain a Community decision within the short time-limits laid down by the Merger Regulation, with a consequent increase in the risk that such a decision may be taken only after the national measures have already irretrievably prejudiced the merger with a Community dimension."[61]

Apart from the political pressure that Article 21 ECMR proceedings bring to bear on the Member State concerned, that alone may result in the withdrawal of an offending national measure, an Article 21 ECMR infringement decision may also be a powerful weapon in the hands of an aggrieved investor before the national courts.

60 *See supra* note 52.

61 *See id.* at paragraph 55.

C. REMEDIES AT A NATIONAL LEVEL

It is a well-established principle of EU law that national courts are under a general duty to disregard any national measure that is inconsistent with EU law.[62] Article 10 EC stipulates that Member States "shall abstain from any measure which could jeopardize the attainment of the objectives of the Treaty," and, as such, requires all Member States, including their courts and competition authorities, to take all necessary measures to guarantee the application and effectiveness of EU law.[63] Article 10 EC embodies the so-called "effet-utile" of EU law and the EC courts have held that it places numerous practical duties and obligations on Member States in the context of complying with both the letter and the spirit of the Treaty.

National courts are under a duty to enable companies to challenge state measures that are contrary to EU law, for example, through judicial review or a declaration that the national measures are contrary to EU law. While the procedures for mounting such a challenge are governed by national law, they must not make it impossible or excessively difficult to exercise the rights granted by EU law.[64]

In *Fiammiferi*,[65] the ECJ extended the duty to set aside national provisions conflicting with Community law to non-judicial bodies such as competition authorities. According to the ECJ, since a national competition authority is responsible for ensuring, inter alia, that Article 81 EC is observed, then Article 81, in conjunction with Article 10 EC, imposes a duty on Member States to refrain from introducing (or to withdraw) measures contrary to the EC competition rules, as otherwise those rules would be rendered less effective if the competition authority were not able to disregard a national measure that is contrary to the combined provisions of Articles 10 and 81 EC.[66]

Where the proposed merger does not fall within the Commission's exclusive competence (i.e., because the parties each achieve two-thirds of their turnover in one and the same Member State), an aggrieved investor will usually be left with a claim that the protectionist measure infringes the rules on free movement of capital (Article 56 EC) and the right of establishment (Article 43 EC). In this scenario, the national court will have to determine whether the national measure at issue is in breach of the Treaty freedoms, and, if so, if it is justified under

62 *See* Case 106/77, Amministrazione delle Finanze dello Stato v. Simmenthal SpA, 1978 E.C.R. 629; Cases C-46 & 48/93, Brasserie du pecheur / Bundesrepublik Deutschland and The Queen / Secretary of State for Transport, *ex parte* Factortame and others, 1996 E.C.R. I-1029; Case 66/86, Ahmed Saeed Flugreisen and others / Zentrale zur Bek_mpfung unlauteren Wettbewerbs, 1989 E.C.R. 803.

63 *See, e.g.*, Case 68/88, Commission v. Greece, 1989 E.C.R. 2965, at paragraph 23.

64 *See, e.g.,* Case C-312/93, Peterbroeck, Van Campenhout & Cie / Belgian State, 1995 E.C.R. I-4599.

65 Case C-198/01, Consorzio Industrie Fiammiferi (CIF) v. Autorità Garante della Concorrenza e del Mercato, 2003 E.C.R. I-8055.

66 *See id.* at paragraph 50.

the criteria set down by the ECJ. If the matter is not clear-cut, it is likely to be referred by the national court to the ECJ for a preliminary ruling under Article 234 EC.[67] Even if there is no referral to the ECJ, inevitably, there will be a risk of several time-consuming appeals before the finding of an infringement of EU law becomes final. Such an action could be pursued in parallel with Article 226 EC infringement proceedings against the Member State concerned.

As already mentioned, where a proposed merger falls within the ambit of the ECMR, then Article 21 ECMR is likely to offer the most effective and timely remedy for an aggrieved investor. Ideally, an aggrieved investor bringing an action before a national court would be armed with both a Commission clearance decision under the ECMR in respect of the proposed merger, and a Commission Article 21 ECMR infringement decision condemning the infringing state measure, in which case the national court would have little choice but to set aside the state measure concerned. Furthermore, in such a scenario, it is unlikely that questions of EU law would arise that the national court would feel compelled to refer to the ECJ under Article 234 EC, and the scope for a Member State to appeal an adverse judgment of the national court is also likely to be considerably diminished. Overall, this should also result in a much speedier resolution of any claim.

On bringing an action before a national court, it would also be open to an aggrieved investor to apply for interim measures. In the *Factortame* case,[68] the ECJ held that national courts are obliged to provide interim relief against the state even if such a remedy was not available under national law. This means that a national court would have a duty to suspend the application of any state measures contrary to EU law. Insofar as a blocked bid may result in losses for a frustrated bidder, it would also be open to the bidder to bring an action for damages against the Member State concerned, as the national courts also have a duty to award compensation for breaches of EU law.[69]

D. DO THE OVERRIDING INTERESTS OF THE INTERNAL MARKET JUSTIFY THE ABOLITION OF THE TWO-THIRDS RULE?

The Commission has shown in its prohibition decision in *EDP/ENI/GDP*[70] that it will not hesitate to block the creation of a national champion where it considers

67 Under Article 234 EC, any court or tribunal of a Member State may (and in certain circumstances must) refer a case to the ECJ for a preliminary ruling concerning the interpretation of EU law. The ECJ rules on the issues referred to it and sends the case back to the national courts which then apply the EU law in question as interpreted by the ECJ to the case at hand.

68 Case 213/89, R v. Secretary of State for Transport*, ex parte* Factortame, Ltd.,1990 E.C.R. I-2433.

69 Cases C-6 & 9/90, Francovich and Bonifaci v. Italy, 1993 E.C.R. I-5357.

70 Case COMP/M.3440, EDP/ENI/GDP, 2005 O.J. (C 288) 2.

that such a merger would significantly impede effective competition in the European Union. Moreover, the Commission appears equally determined to ensure that its clearance decisions and its exclusive jurisdiction under the ECMR are not obstructed by interventionist measures, from Member States, designed to prevent foreign takeovers. European Commissioner Neelie Kroes recently stated that "the EU's single market will descend into chaos" if Member States stand in the way of mergers falling within the Commission's exclusive competence.[71] The Commission's determination to tackle protectionist measures is well illustrated by the vigorous stance it has taken against Spain's attempts to obstruct E.ON's proposed takeover of Endesa as well as its infringement decisions under Article 21 ECMR against the Portuguese government in the *Champalimaud* and *Cimpor* cases.[72]

However, the Commission is effectively powerless to prevent the creation of national champions, regardless of what impact this might have on the internal market, when it comes to mergers which fall outside its exclusive jurisdiction under the ECMR because of the application of the two-thirds rule. Much to Commissioner Kroes' frustration, this was the case with Gas Natural's proposed takeover of Endesa, that led the Commissioner to call for the abolition of the two-thirds rule.[73]

Arguably, the two-thirds rule, which was introduced as part of the EC's first merger control regime in 1989, is outdated and inconsistent with its overriding internal market objectives. However, it is worth noting that, at the time of its review of the ECMR in 2003, the Commission was of the opinion that the two-thirds rule should be retained on the grounds that it applies to less than 10 percent of filings and is a reasonable expression of the principle of subsidiarity.

Arguably, the two-thirds rule, which was introduced as part of the EC's first merger control regime in 1989, is outdated and inconsistent with its overriding internal market objectives.

Aside from the two-thirds rule, the ECMR provides for mechanisms whereby concentrations with a Community dimension, that affect competition in a distinct market within a Member State, can be reviewed at the national level. Under Article 9(2)(a) ECMR a Member State can request a referral back to its competent authorities where a concentration "threatens to affect significantly" competition in a distinct market within that Member State. Where the Commission considers that such a threat exists, it can either assert its exclusive jurisdiction and deal with the case itself, or refer the whole or part of the case to the Member State's competition authorities. In this scenario, the Commission

71 European Commission, Press Release No. 277, Mergers: Commission launches procedure against Poland for preventing Unicredit/HVB merger (2006).

72 *See supra* notes 50 & 51.

73 *See supra* note 4.

has a wide margin of discretion in deciding whether or not to refer the case to the competent national authorities.

According to Article 9(2)(b) ECMR, a request for a referral back can also be made by a Member State where a concentration "affects competition in a distinct market which does not constitute a substantial part of the common market" (i.e., in a local or regional market within that Member State). If the Commission considers such a local or regional market to be affected by the concentration, it is obliged to refer the whole case or that part of the case which relates to this market, to the Member State's competition authorities.[74]

Given that amendments to the ECMR require unanimous approval in the Council,[75] all Member States would ultimately have to agree to abolish the two-thirds rule." This may, of course, be difficult to achieve given the protectionist tendencies of Member States. Even those Member States that may be willing to give more discretion to the Commission may not be keen to make further changes to the ECMR relatively soon after its recent overhaul.

However, rather than completely abolishing the two-thirds rule, a more palatable solution for Member States might be to include the two-thirds rule as an additional ground for the referral of mergers back to Member States under Article 9(2)(a) ECMR. This would allow the Commission to assess the cross-border impact of any merger, even when it concerns essentially domestic players, in exercising its discretion to accede to a Member State request for referral back. This would effectively eliminate the possibility of Member States frustrating foreign takeovers or the objectives of the internal market by the creation of national champions from mergers falling outside the Commission's exclusive jurisdiction under the ECMR.

74 There is also a referral mechanism under Article 4(4) ECMR which can be initiated by the notifying parties. *See supra* note 10.

75 As the legal basis for the ECMR itself is Article 83 EC in conjunction with Article 308 EC, which requires unanimity, any amendments to the ECMR can only be adopted unanimously on the same legal basis.

Volume 2 | Number 2 | Autumn 2006

Competition Policy International

Commerce Clause Constraints on State Business Location Incentives

Peter D. Enrich

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Commerce Clause Constraints on State Business Location Incentives

Peter D. Enrich

Over the past several decades, states and municipalities in the United States have engaged in an accelerating competition to reward business location and investment through the use of a wide range of financial incentives, notwithstanding overwhelming evidence of the minimal efficacy and the high costs of such incentives. This interstate competition for economic activity is reminiscent of the eighteenth-century tariff wars among the states that were a primary impetus behind the crafting of the U.S. Constitution and its assignment of responsibility for the regulation of interstate commerce to the federal government. Over the ensuing centuries, the courts have consistently applied the Constitution's Commerce Clause to constrain parochial state measures that interfered with the free flow of commerce in a national common market.

This article considers whether, and to what extent, the Commerce Clause limits the ability of states and localities to engage in the incentive competition that has proliferated in recent decades. In particular, I argue that well-established Commerce Clause principles forbid a wide range of the location-based tax incentives that states and localities offer to businesses. At the same time, it is important to recognize that judicial application of the Commerce Clause offers, at best, a blunt instrument for addressing the challenges of interstate competition for business investment. This article will also canvass a range of limitations and shortcomings of this constitutional constraint on governmental efforts to intervene in business location decisions.

The author is Professor of Law, Northeastern University School of Law.

I. The Context: Interstate Competition for Business Investment

By the late 1990s, virtually every state, and a great many localities, was aggressively engaged in offering a wide range of incentives for businesses that located their investment and employment in the jurisdiction.[1] While the array of incentives offered by the states often includes expenditures of governmental funds (e.g., for infrastructure improvements, worker training, or, in some cases, cash grants or low-cost loans) or regulatory accommodations, the primary focus of governmental intervention has been through a broad palette of tax incentives. Investment tax credits, job creation credits, and property tax abatement programs have become virtually universal, and in recent years more and more states have been adjusting their rules for the apportionment of corporate income to focus primarily, if not exclusively, on the location of sales, rather than of payroll and property.

Some of these tax incentives are designed as discretionary programs whose benefits are awarded by negotiation between businesses and state officials; others are specifically crafted and enacted to respond to particular industries or even to particular businesses; while still others are structured as entitlements that can be utilized by any company that satisfies their broad criteria. Most states offer long lists of different types of tax incentives geared toward different business situations, and scarcely a week goes by without reports of at least one state proposing or enacting a new tax incentive program. Indeed, a whole business has emerged of consultants who help businesses to keep track of the available incentives and to ensure that they claim all of the incentives to which they are entitled. Reports of incentive packages measuring in the hundreds of millions of dollars have become increasingly frequent, and it is not uncommon for the incentive package for a large plant to be big enough to excuse the company from any state or local tax liability in the jurisdiction for a period of years.

Putting a price tag on all of this activity is not easy. But one careful scholarly effort, extrapolating from the few states providing solid data, estimated that the total cost of state and local incentives, both tax and non-tax, in 1996 approximated US$50 billion.[2] The number a decade later is surely substantially larger and represents a significant fraction of state and local revenues.[3] Perhaps equal-

1 Under U.S. constitutional principles, local governments are political subdivisions of the states, and local actions are, in general, subject to the same constitutional constraints as state actions. So, in this article, I will not distinguish between state and local measures, and will generally use the term "state" as shorthand for "state and local."

2 *See* Kenneth Thomas, Competing for Capital – Europe and North America in a Global Era 158-59 (2000).

3 For the sake of comparison, total state and local own-source revenues in 1995-96 were US$988 billion. U.S. Census Bureau, U.S. State and Local Government Finances by Level of Government: 1995-96, *available at* http://www.census.gov/govs/estimate/9600us.html.

ly importantly, the proliferation of corporate tax breaks is one of several factors that have contributed to a dramatic shift in the distribution of state and local tax burdens between businesses and individual taxpayers. Between 1979 and 1999, businesses' share of total state income tax revenues declined from 29 percent to 15 percent, and the business share of local property tax revenues has likewise dropped sharply.[4]

Despite these high costs, state and local tax incentives do not appear to exert a significant influence on the location of business activity. As business tax incentives have proliferated, so have econometric analyses of the effects of state and local tax burdens and specifically of state and local tax breaks on business activity and investment. Recent surveys of the dozens of empirical studies conclude that state and local tax breaks have been shown to have, at best, only marginal effects on business location.[5]

Several factors contribute to this somewhat surprising conclusion. First, state and local taxes are generally too small to be a major factor in the economics of business location decisions. For the typical business, state and local taxes represent only 1.2 percent of the cost of doing business in the United States.[6] So, even an incentive package that completely eliminates a company's tax obligation will only have a modest effect on overall costs. Variations in other factors, such as costs and skills of labor, access to resources and markets, and utility costs, are likely to overwhelm any small potential savings from tax reductions. Second, since all U.S. jurisdictions are offering a wide array of tax breaks, the potential savings from one state's incentives will be largely offset by those available from competing locations. Moreover, even the modest positive effects of lower taxes on business location found in the research assume that all other factors, including levels of state and local government spending, are held constant. But, in fact, states, unlike the federal government, are subject to balanced-budget requirements, and the evidence reveals robust positive relationships between spending on public services and economic activity.[7] So, if tax breaks reduce the funding available for services, any positive effects of reduced taxes are liable to be offset by the negative effects of reduced services.

4 Robert Tomsho, *In Toledo, a Tension between School Funds and Business Breaks*, WALL ST. J., July 18, 2001, at A1 (reporting U.S. Census data). *See* also Laird Graeser & Al Maury, *Business Taxes – Quo Vadimus*, 7 ST. TAX NOTES 917, 918 (1994) (summarizing data reported by the U.S. Advisory Commission on Intergovernmental Relations indicating that state business taxes had declined from 50 percent of state taxes in the 1950s to 25 percent by 1990).

5 *See* ROBERT G. LYNCH, RETHINKING GROWTH STRATEGIES – HOW STATE AND LOCAL TAXES AND SERVICES AFFECT ECONOMIC DEVELOPMENT (2004); Alan Peters & Peter Fisher, *The Failures of Economic Development Incentives*, 70 J. AMER. PLANNING ASSOC. 27 (Winter 2004).

6 *See* LYNCH, *supra* n. 4, at 4.

7 *See id.* at 43-46.

Even if a state's location incentives do have some positive effect on business investment in that state, the effect of the incentive competition from a national perspective is at best a zero-sum game. Despite occasional suggestions by defenders of state and local incentives that they might help the United States compete in the international market for investment, there is no credible evidence that these incentives are of a scale to have such effects, nor that they were ever designed with such a purpose in mind. At best, these incentives affect the location of economic activity among the states, not its overall level. Indeed, to the modest extent that state incentives are effective and influence businesses to site their activities at locations that would otherwise be economically disfavored, the incentives are likely to reduce, rather than enhance, national economic efficiency.[8] Thus, the primary effect of the states' incentive competition, from a national perspective, is not to encourage or expand economic activity, but rather to lower the general level of state and local taxation of businesses, in a "race to the bottom" that either shifts tax burdens to other taxpayers or reduces the resources available for state and local governmental services.

Notwithstanding the evidence of the minimal efficacy, and deleterious effects, of state tax incentive proliferation, state policymakers have shown little inclination to walk away from the competition. As many analysts have observed, the states are caught in a version of a prisoners' dilemma, where it is irrational for any one state to stop offering incentives while other states remain free to continue providing them.[9] Indeed, even to the extent that state officials recognize the virtual irrelevance of incentives to business location decisions, they are reluctant to forego the use of a tool which, regardless of its actual effect on business behavior, is a powerful way to communicate to voters their commitment to the state's economic vitality. Absent some external constraint, the competition among states and localities to offer ever-larger incentive packages appears unlikely to abate.

II. The Legal Framework: The Commerce Clause's Role

Most U.S. lawyers will be quite surprised to find a discussion of these issues in a journal devoted to competition policy. Unlike the EC framework, within which the problem of state aid is conceptualized as one among many forms of interference with the functioning of competitive market forces, U.S. antitrust law focuses almost exclusively on anticompetitive activities of private actors. Indeed, the presence of state action ordinarily suffices to immunize conduct, even by private

8 For a description and critique of an argument that state incentives might actually help to optimize the efficiency of business location decisions, *see* Peter D. Enrich, *Business Tax Incentives: A Status Report*, 34 Urb. Lawyer 415, 418-22 (2002).

9 *See, e.g.,* Thomas, *supra* n. 2, at 33-40.

parties, from antitrust scrutiny.[10] The notion that state or local use of incentives to reward local business activity might raise antitrust concerns would find no toehold in U.S. law.

Instead, the U.S. legal system's limitations on state interference with an open national common market are conceptualized as aspects of the constitutional framework of federalism, which establishes the respective scope and limits of national and state authority, a body of law that arose more than a century before the emergence of antitrust law's constraints on private anticompetitive conduct. As the U.S. Supreme Court has repeatedly observed, the U.S. Constitution's federalism, and indeed the Constitution itself, were designed largely as a response to the destructive interstate competition for economic activity, most notably in the form of "customs barriers and other economic retaliation,"[11] that characterized the pre-Constitutional period: "If there was any one object riding over every other in the adoption of the constitution, it was to keep the commercial intercourse among the States free from all invidious and partial restraints."[12]

THE U.S. LEGAL SYSTEM'S LIMITATIONS ON STATE INTERFERENCE WITH AN OPEN NATIONAL COMMON MARKET ARE CONCEPTUALIZED AS ASPECTS OF THE CONSTITUTIONAL FRAMEWORK OF FEDERALISM.

At the heart of the Constitution's response to state interference with an open economy is the Commerce Clause, which delegates to the federal government the power to "regulate Commerce . . . among the several states."[13] Although the words of the clause speak only of an affirmative grant of authority to the U.S. Congress, the framers understood,[14] and the Supreme Court has consistently recognized,[15] that it served an equally important negative or dormant function, as a prohibition against state measures that interfere with or seek to constrain interstate economic activity for local advantage. Indeed, while the Commerce Clause has proven a fertile source of a very wide range of federal legislative activity (including, among a great many others, the federal antitrust laws), only rarely has

10 *See, e.g.,* City of Columbia v. Omni Outdoor Advertising, Inc., 499 U.S. 365 (1991).

11 Baldwin v. G.A.F. Seelig, Inc., 294 U.S. 511, 522 (1935) (quoting 2 The Records of the Federal Convention of 1787, at 308 (Max Farrand ed., revised ed. 1937)).

12 Camps Newfound/Owatonna v. Town of Harrison, 520 U.S. 564, 571 (1997) (quoting Gibbons v. Ogden, 22 U.S. (9 Wheat.) 1, 231 (1824) (Johnson, J., concurring in judgment)).

13 U.S. Const., art. I, § 8, cl. 3.

14 *See* 3 The Records of the Federal Convention of 1787, at 478 (James Madison explaining that the Commerce Clause "was intended as a negative and preventive provision against injustice among the States themselves").

15 *See, e.g., Baldwin*, 294 U.S. at 522.

Congress exercised its Commerce Clause authority specifically to rein in state interference with interstate commerce. By contrast, the courts have applied the dormant Commerce Clause with great frequency to invalidate state measures that impermissibly infringed the free flow of interstate commerce, whether through regulation or through taxation.

Among the primary subjects of the courts' attention, dating back into the nineteenth century, have been the recurrent efforts of the states to use their tax systems to provide preferential treatment for in-state economic activity. Tariffs are, of course, the paradigm for such measures, although, as the Supreme Court has observed, "tariffs . . . are so patently unconstitutional that our cases reveal not a single attempt by any State to enact one."[16] Instead, the Court has reviewed a vast array of other types of state tax measures that touched on interstate commerce to determine whether they impermissibly interfered with the Commerce Clause's common-market goals.

In this long history, the Court's efforts to set appropriate limits on state taxation have deployed a wide range of different, and at times inconsistent, approaches.[17] Nevertheless, amidst this complexity,

> "there emerge . . . some firm peaks of decision which remain unquestioned. Among these is the fundamental principle . . . : No state, consistent with the Commerce Clause, may impose a tax which discriminates against interstate commerce . . . by providing a direct commercial advantage to local business. The prohibition against discriminatory treatment of interstate commerce follows inexorably from the basic purpose of the Clause. Permitting the individual States to enact laws that favor local enterprises at the expense of out-of-state businesses would invite a multiplication of preferential trade areas destructive of the free trade which the Clause protects."[18]

For the past thirty years, the Court has adopted a relatively stable analysis for Commerce Clause challenges to state tax measures, which assesses a challenged measure's practical effects against a four-prong test. The Court's long-standing,

16 West Lynn Creamery v. Healy, 512 U.S. 186, 193 (1994).

17 For a helpful overview of this tortuous history, *see* Walter Hellerstein, *State Taxation of Interstate Business: Perspectives on Two Centuries of Constitutional Adjudication*, 41 Tax Law 37, 38-50 (1987).

18 Boston Stock Exch. v. State Tax Comm'n, 429 U.S. 318, 329 (1977) (internal quotations and citations omitted).

anti-discrimination principle is one of the four prongs, and remains the one of primary relevance for measures that favor or reward in-state economic activity.[19] Indeed, over the past three decades, the Court has deployed the anti-discrimination principle to invalidate more than a dozen different state tax strategies that provided preferential treatment for in-state activity or in-state actors.

The types of measures that the Court has struck down as discriminatory are diverse. *Bacchus Imports, Ltd. v. Dias* invalidated a Hawaii provision that exempted certain locally produced alcoholic beverages from an otherwise generally applicable liquor excise tax.[20] *Boston Stock Exchange v. State Tax Commission* found unconstitutional a New York stock transfer tax that provided preferential rates for transfers that were executed on New York stock exchanges, rather than on out-of-state markets.[21] *Westinghouse Electric Corp. v. Tully*, struck down a measure, again from New York, that granted a credit against the state's corporate income tax measured by the share of the company's export business that was conducted in New York.[22] *Fulton Corp. v. Faulkner* overturned a North Carolina property tax which reduced the tax on ownership of corporate shares as the percentage of the corporation's business that was located in North Carolina increased.[23]

The Court's concept of what constitutes discrimination is straightforward: "'discrimination' simply means differential treatment of in-state and out-of-state economic interests that benefits the former and burdens the latter."[24] A primary focus in determining whether a particular tax provision runs afoul of the anti-discrimination principle is a practically oriented analysis of the provision's purposes and effects. If an underlying purpose of a provision is to advantage local commerce or local activities, or if a natural consequence of the measure is to distort tax-neutral decisions about where to do business or to "exert [ ] an inexorable hydraulic pressure" favoring in-state activity, these are each strong indicia of discrimination.[25] In this analysis, the Court directs a particularly critical eye towards

19 *See* Complete Auto Transit, Inc. v. Brady, 430 U.S. 274, 279 (1977). The *Complete Auto* test requires that a tax "is applied to an activity with a substantial nexus with the taxing State, is fairly apportioned, does not discriminate against interstate commerce, and is fairly related to the services provided by the State."

20 Bacchus Imports, Ltd. v. Dias, 468 U.S. 263 (1984).

21 Boston Stock Exchange v. State Tax Commission, 429 U.S. 318 (1977).

22 Westinghouse Electric Corp. v. Tully, 466 U.S. 388 (1984).

23 Fulton Corp. v. Faulkner, 516 U.S. 325 (1996).

24 Oregon Waste Sys., Inc. v. Dep't of Envtl. Quality, 511 U.S. 93, 99 (1994).

25 *See, e.g., Bacchus Imports*, *supra* n. 20, at 263, 270-73; *Boston Stock Exch.*, *supra* n. 21, at 331; American Trucking Ass'ns v. Scheiner, 483 U.S. 266, 286 (1987).

provisions that discriminate “on their face,” which the Court deems “virtually per se invalid.”[26] While the Court has never offered a precise definition of facial discrimination, the measures to which it has assigned this label are those where the differential treatment of in-state and out-of-state activity is evident in the language of the operative tax statute.

It is important to observe that the Court’s anti-discrimination principle requires two distinct elements: both differential treatment and a resultant benefit for in-state activity. If a measure does not provide for distinctive treatment of in-state and out-of-state activity, then the mere fact that the measure may have the effect or purpose of encouraging in-state activity does not render it discriminatory. For instance, a generally applicable reduction of business tax rates or an exemption of business personal property from property taxation surely has the effect, and likely the purpose, of encouraging local investment, but such measures do not treat out-of-state activity in less favorable ways, and nothing in the Court’s Commerce Clause jurisprudence suggests that such measures raise any hint of forbidden discrimination. Indeed, the Court has repeatedly observed that “it is a laudatory goal in the design of a tax system to promote investment that will provide jobs and prosperity to the citizens of the taxing State.”[27] The Commerce Clause only forbids those efforts that seek to achieve these legitimate ends by improper means, means which discriminate in their treatment of in-state and out-of-state activity. As the Court explained in *Boston Stock Exchange*, “in the process of competition no State may discriminatorily tax the products manufactured or the business operations conducted in any other State.”[28]

III. The Commerce Clause Applied to State Location Incentives

To what extent does the Supreme Court’s anti-discrimination jurisprudence set limits to the proliferating efforts of the states to use their tax systems to provide incentives for in-state investment? At the least, many of the tax incentive measures which have become commonplace in recent years invite serious questions of their validity under the Commerce Clause. So, it is perhaps surprising that the Supreme Court has not had the occasion to address the constitutionality of any of the characteristic tax incentives that have been broadly adopted by the states.

The likely explanation lies in the fact that the Court can only address cases that parties bring to it. And the typical parties who litigate Commerce Clause challenges to state tax measures are out-of-state or interstate businesses who are

26 *See, e.g., Fulton Corp., supra* n. 23, at 325, 331.

27 Trinova Corp. v. Michigan Dep’t of Treasury, 498 U.S. 358, 385-86 (1991).

28 *Boston Stock Exch.*, 429 U.S. at 337.

disfavored by the benefits that a challenged measure provides to their in-state competitors. But such businesses are typically receiving the benefit of similar incentives in the states where their plants or other activities are located; so, they would be ill-advised to bring a challenge which, if successful, might well kill the goose that is laying their golden eggs.

In its recently concluded 2005 term, the Supreme Court finally did take a case, *DaimlerChrysler Corp. v. Cuno*, in which the U.S. Court of Appeals for the Sixth Circuit had found that Ohio's investment tax credit violated the Commerce Clause.[29] But this case was brought, not by a business competitor, but by a group of local taxpayers, both individuals and small businesses, who challenged the tax credit because of its negative impact on the state's tax revenues and the resultant impact on them in the form of higher taxes and reduced state services. The Supreme Court concluded that such plaintiffs did not have the requisite personal stake in the litigation to satisfy the requirements for standing to sue in the federal courts, and therefore the Court did not reach the merits of the Commerce Clause claim. Although the plaintiffs remain free to—and intend to—pursue their claim in the state courts, which apply their own, more permissive, rules concerning standing to sue,[30] the outcome in the Supreme Court indicates some of the hurdles (to which we return below) that stand in the way of a definitive ruling on the constitutionality of the more widespread state incentives.

Nonetheless, a number of state tax measures whose purpose or effect is to encourage or reward in-state business location decisions have been reviewed by the courts, and they have repeatedly been found to violate the anti-discrimination principle. The Supreme Court has struck down New York's incentives for locating export activity in the state,[31] North Carolina's property tax breaks for shareholders of companies that expand their in-state presence,[32] and a Louisiana severance tax credit that favored in-state mineral extraction,[33] in each case focusing on the fact that the preference for in-state activity would impermissibly encourage businesses to locate new investment in the state. In addition to the

29 126 S. Ct. 1854 (2006) (vacating 386 F.3d 738 (6th Cir. 2004)). In the interests of full disclosure, I note that I represented the citizen plaintiffs in this case before both the U.S. Court of Appeals for the Sixth Circuit and the U.S. Supreme Court, and continue to represent them in their forthcoming state court suit.

30 In fact, the plaintiffs had initiated the original suit in the Ohio state courts, partly because of concerns about federal rules concerning standing. The case was removed to federal court by the defendants, over plaintiffs' objections. Defendants only attacked plaintiffs' standing after the plaintiffs' victory on the merits before the Court of Appeals.

31 *Westinghouse Electric, supra* n. 22.

32 *Fulton Corp., supra* n. 23.

33 Maryland v. Louisiana, 451 U.S. 725, 756-57 (1981).

Sixth Circuit's decision invalidating Ohio's investment tax credit,[34] other appellate courts have struck down a New York City provision providing accelerated depreciation limited to assets placed in service in the city,[35] an exemption from Pennsylvania's capital stock tax that was designed to encourage in-state location of manufacturing facilities,[36] and a Nevada sales tax exemption limited to air carriers that located their central offices in the state.[37]

As these cases suggest, many of the characteristic state tax incentives used to reward in-state investment are vulnerable to a straightforward and compelling application of the Supreme Court's anti-discrimination reasoning. Consider, for example, an investment tax credit (ITC), like the one challenged in the *Cuno* case, one of the most ubiquitous forms of location incentive. An ITC allows a business to reduce its state income tax by a specified percentage of the cost of new facilities, machinery, or equipment acquired or placed in service in the taxing state. States impose a variety of restrictions on the classes of property or types of businesses eligible for the credit, but they all restrict the credit to investments in property located and used within the state.

Because of this locational restriction, an ITC discriminates in precisely the way that the anti-discrimination principle forbids. Its differential treatment of in-state and out-of-state economic activity is evident. Compare two otherwise identically situated businesses, each of which is subject to the state's income tax on an identical portion of its income. If one now builds a new facility in the state and the other builds an identical facility elsewhere, the first will be entitled to a credit against its state income tax, while the other will not.[38]

---

34 Cuno v. DaimlerChrysler Corp., 386 F.3d 738 (6th Cir. 2004), *vacated on other grounds*, 126 S. Ct. 1854 (2006).

35 R.J. Reynolds Tobacco Co. v. Dep't of Fin., 667 N.Y.S.2d 4 (N.Y. App. Div. 1997).

36 PPG Indus. v. Commonwealth, 790 A.2d 252 (Pa. 1999).

37 Worldcorp. v. Dep't of Taxation, 944 P.2d 824 (Nev. 1997). One arguable exception to this pattern is Caterpillar, Inc. v. Dep't of Treasury, 488 N.W.2d 182 (Mich. 1992), which upheld a Michigan tax preference for in-state capital investment, but the ruling in that case depended heavily on the unique features of Michigan's single business tax.

38 Of course, because of the new investments, the two companies are no longer identically situated. Defenders of state ITCs have suggested that, since the in-state location of a new plant will increase the proportions of the company's property and payroll located in the state and will thereby increase the proportion of the company's income taxable in the state under many states' apportionment rules, the ITC might be justified as merely compensating for the tax increase attendant on the plant location. In fact, however, in almost any realistic scenario, the tax savings from an ITC will vastly exceed any added tax burden from the new plant's in-state location. And, in any case, the ITC's differential effect will remain evident in the different effective tax rates that the two hypothetical companies will pay on the share of their incomes apportioned to the state under the state's (presumably legitimate) apportionment methods, with the in-state company paying a lower effective rate than its out-of-state competitor.

Nor is there any question that the effect of the ITC is to give an advantage to in-state investment, thereby encouraging businesses to locate their new facilities in the state. As the Sixth Circuit explained in *Cuno*, concerning Ohio's ITC:

> "as between two businesses, otherwise similarly situated and each subject to Ohio taxation, the business that chooses to expand its local presence will enjoy a reduced tax burden, based directly on its new in-state investment, while a competitor that invests out-of-state will face a comparatively higher tax burden because it will be ineligible for any credit against its Ohio tax."[39]

Thus, the practical effect of the ITC is to "encourage [ ] the development of local industry by . . . impos[ing] greater burdens on economic activities taking place outside the State than were place[d] on similar activities within the State."[40] In short, the ITC precisely fits the Supreme Court's definition of forbidden discrimination: it constitutes "differential treatment of in-state and out-of-state economic interests that benefits the former and burdens the latter."[41] And, since the discrimination is expressly set forth in the language of the ITC statutes, when they restrict the credit to investments located in the taxing state, state ITCs are apt to be found "virtually per se unconstitutional."

This outcome should come as no surprise, because, in fact, an ITC serves as the functional equivalent of a tariff on out-of-state manufacturers. Were a state to tax two different businesses, each of which sold its products in the state, at different rates based on where the goods were manufactured, with the in-state manufacturer paying a lower rate than its out-of-state competitor, we would have a classic instance of a forbidden tariff.[42] But this is precisely the effect that an ITC accomplishes, albeit by somewhat different means. After all, a credit that is available only on the basis of in-state investment reduces the effective tax rate of those businesses with in-state facilities, just as an explicitly lower tax rate would.

39 *Cuno*, *supra* n. 34, at 743.

40 *Westinghouse Electric*, *supra* n. 22, at 404.

41 *Oregon Waste*, *supra* n. 24, at 99.

42 Perhaps it might be argued that a classic tariff would operate as a tax on the sales (i.e., on the gross revenue from the transactions) rather than on the apportioned net income of the competing companies. But, as the Supreme Court has emphasized in applying the anti-discrimination principle to a corporate income tax, "It cannot be that a State can circumvent the prohibition of the Commerce Clause against placing burdensome taxes on out-of-state transactions by burdening those transactions with a tax that is levied in the aggregate . . . rather than on individual transactions." *Westinghouse Electric*, *supra* n. 22, at 404.

Indeed, since ITCs are often large enough to offset a business's entire state income tax liability for a period of years after they locate a substantial manufacturing facility in the state, an ITC can convert an otherwise neutral income tax on manufacturers into a tariff that applies only to those competitors who manufacture outside the state. Thus, it should come as no surprise that a number of commentators describe ITCs as paradigmatic examples of tax breaks that violate the Commerce Clause.[43]

The same anti-discrimination argument will also reach a wide range of the other tax incentives currently in use by the states. Other locationally based credits against corporate income taxes, such as targeted jobs credits which are measured by a company's new employment in the taxing jurisdiction, are susceptible to precisely the same analysis as ITCs. Or, to take a somewhat different example, consider the increasingly common use of property tax exemptions that are conditioned on a specified level of new employment at the exempted facility or on other forms of in-state activity. As noted earlier, no one would suggest that a simple exemption of a certain class of assets from property taxation would offend the Commerce Clause. Since the state does not—indeed, cannot—tax comparable out-of-state properties, its decision not to tax the in-state properties does not discriminate in favor of the in-state investments.

But, if the property tax exemption is conditioned on some additional form of in-state activity, such as a specified level of in-state employment, the provision becomes discriminatory. Here, the discrimination is not between a business with in-state property and a competitor with out-of-state property, but rather between two businesses with in-state property, one of whom commits to the requisite level of in-state employment (or other in-state activity on which the property exemption is conditioned) and the other of whom concentrates its new employment out-of-state (or is unable or unwilling to commit to the required level of in-state activity). Here, as with the ITC, the tax provision expressly favors the business engaging in in-state activity over a comparably situated competitor who does not, by exempting the property of the one from taxation, while taxing the property of the other.

Several judicial decisions have found such location-based conditions on otherwise non-discriminatory tax exemptions to violate the Commerce Clause. For example, the Supreme Court struck down a Maine property tax exemption for charitable organizations, because the exemption was available only to those organizations which primarily served in-state residents.[44] Similarly, the U.S.

43 *See, e.g.,* Walter Hellerstein & Dan T. Coenen, *Commerce Clause Restraints on State Business Development Incentives*, 81 Cornell L. Rev. 789, 817-18 (1996); Peter D. Enrich, *Saving the States from Themselves: Commerce Clause Constraints on State Tax Incentives for* Business, 110 Harv. L. Rev. 377, 434-37 (1996); Robert D. Plattner, *State Business Tax Incentives: Are They Vulnerable to Constitutional Attack*, St. Tax Today 128-19 (Jul. 3, 2000).

44 Camps Newfound/Owatonna v. Town of Harrison, 520 U.S. 564 (1997).

Court of Appeals for the Fifth Circuit invalidated a Louisiana provision which conditioned property tax exemptions for new industrial facilities on the taxpayer's agreement to give preference to in-state suppliers, contractors, and labor in the construction and operation of the exempted facility.[45] And the Sixth Circuit, in *Cuno*, despite finding that the particular property tax exemption challenged in that case was not unconstitutional, emphasized that "an exemption may be discriminatory if it requires the beneficiary to engage in another form of business in order to receive the benefit or is limited to businesses with a specified economic presence," although it found that the particular conditions imposed on the challenged exemption did not cross those thresholds.[46]

Thus, a wide range of location-based tax incentives appear vulnerable to Commerce Clause invalidation. The primary stratagem of defenders of such incentives in responding to such arguments has been to suggest that the Supreme Court's anti-discrimination case law can, and should, be read more narrowly, in a manner that would not reach typical investment incentives of the sort discussed above. In particular, they argue that the Supreme Court has only invalidated state tax provisions as discriminatory when they either function as tariffs levied directly on interstate transactions or impose tax penalties on businesses for their out-of-state activities.[47] The suggestion, then, is that measures like ITCs or conditional property tax exemptions are unproblematic, since they do not apply against transactional taxes and since they provide tax reductions (i.e., benefits) for in-state activity, rather than tax increases (i.e., penalties) for out-of-state activity.

As I and other commentators have explained at greater length elsewhere,[48] however, this argument's critical distinction between tax benefits and penalties cannot withstand scrutiny. Not only are its proponents unable to cite a single case in which the Court has deployed such a distinction, but in fact the Court has expressly, and quite sensibly, disavowed any meaningful distinction between tax benefits and burdens:

---

45 Pelican Chapter, Associated Builders & Contractors, Inc. v. Edwards, 128 F.3d 910 (5th Cir. 1997). *See also* Worldcorp v. Dep't of Taxation, 944 P.2d 824 (Nev. 1997) (invalidating a Nevada sales tax exemption because it was restricted to purchasers who located their central office in the state).

46 *Cuno*, *supra* n. 34, at 746.

47 This attempted categorization had its origins in Philip Tatarowicz & Rebecca Mims-Velarde, *An Analytical Approach to State Tax Discrimination Under the Commerce Clause*, 39 Vand. L. Rev. 879 (1986), and has been adopted by a number of more recent commentators and litigants.

48 *See* Enrich, *supra* n. 43, at 444-446; Hellerstein & Coenen, *supra* n. 43, at 813-15; Respondents' Brief at 39-43, DaimlerChrysler, Inc. v. Cuno, 126 S. Ct. 1854 (2006) (No. 04-1704).

> "Virtually every discriminatory statute allocates benefits or burdens unequally; each can be viewed as conferring a benefit on one party and a detriment on the other, in either an absolute or relative sense. The determination of constitutionality does not depend upon whether one focuses upon the benefited or the burdened party."[49]

Indeed, several of the Court's anti-discrimination decisions fail to fit within the argument's narrow categories, because they invalidate measures which operate as benefits for in-state activity rather than as burdens on out-of-state activity.[50] And, were the courts to adopt the purported benefits/penalties distinction, they would reintroduce precisely the type of formalistic distinction that the Supreme Court's practically oriented Commerce Clause jurisprudence has sought to eschew, and would invite states to revive a wide range of forbidden measures by simply recasting them in technically different form. In short, the benefits/penalties distinction is not supported by "either the decisions themselves, or the underlying purposes of the Commerce Clause."[51]

IV. Shortcomings of the Commerce Clause as a Constraint

At present, judicial enforcement of the Commerce Clause's anti-discrimination principle appears to be the only viable legal restraint on the states' proliferating competition to offer ever-more generous tax incentives to reward businesses for locating their facilities in the state.[52] Nonetheless, neither the doctrinal parameters of Commerce Clause law nor reliance on the courts to enforce these constitutional limits on the states are without their difficulties. In this closing section, I briefly canvass several shortcomings of this approach, relating both to the content of the applicable doctrine and to the institutional roles implicated in reliance on judicial enforcement.

---

49 *Bacchus Imports*, *supra* n. 20, at 273.

50 *See Westinghouse Electric*, *supra* n. 22, at 404 ("Nor is it relevant that New York discriminates . . . by disallowing a tax credit rather than by imposing a higher tax. The discriminatory economic effect of these two measures would be identical."); Maryland v. Louisiana, *supra* n. 33, at 757; *American Trucking*, *supra* n. 25; *Camps Newfound, supra* n. 12.

51 Hellerstein & Coenen, *supra* n. 43, at 815.

52 One other possible source of legal restraint, which is beyond the scope of this article, lies in U.S. trade treaty commitments, to the extent that they constrain subsidization of domestic industry. For one introduction to the possible arguments, *see* WILLIAM SCHWEKE & ROBERT K. STUMBERG, COULD ECONOMIC DEVELOPMENT BECOME ILLEGAL IN THE NEW GLOBAL POLICY ENVIRONMENT? (Corporation for Enterprise Development, 1999).

A. THE MARKET-PARTICIPANT EXCEPTION

While the courts have enforced the dormant Commerce Clause broadly as a limit on state and local tax and regulatory measures that discriminate against interstate business activity, the Supreme Court has sharply circumscribed the range of state and local actions that are subjected to Commerce Clause scrutiny. In particular, the Court has drawn a bright line between measures it characterizes as market regulation and those it views as market participation. The former category, which encompasses taxation and enactment of governmental rules and standards that apply to private businesses, warrants rigorous judicial review for potential conflicts with the federal power to "regulate Commerce . . . among the several states." But the latter category, which encompasses governmental execution of its own operations and programs, is completely immunized from Commerce Clause analysis, on the theory that governments, just like private market actors, should be free to choose with whom they do business and on what terms. Thus, the Court has found that the Commerce Clause does not apply to a city's decision to require its contractors to employ local residents[53] or to a state's choice to sell cement produced by a state-owned plant at discriminatory rates favoring in-state purchasers.[54]

JUDICIAL ENFORCEMENT OF THE COMMERCE CLAUSE'S ANTI-DISCRIMINATION PRINCIPLE APPEARS TO BE THE ONLY VIABLE LEGAL RESTRAINT ON THE STATES' PROLIFERATING COMPETITION TO OFFER EVER-MORE GENEROUS TAX INCENTIVES TO REWARD BUSINESSES FOR LOCATING THEIR FACILITIES IN THE STATE.

In consequence of this "market-participant exception," a wide range of the non-tax measures that states commonly use to reward business location decisions, such as providing worker training or infrastructure improvements, assembling sites, or offering low-cost loans, are likely not to be susceptible to judicial scrutiny, regardless of the degree to which they may tilt the playing field in favor of in-state investment. Indeed, it can be argued that even a direct cash subsidy paid to a business would fall within the protected sphere of market participation, although the Court has been careful to note that it has never addressed or decided the constitutionality of direct subsidies.[55] Thus, while some types of state location incentives are subject to close Commerce Clause scrutiny, others, indistinguishable in their financial value to the recipient businesses, are not scrutinized at all.

53 White v. Massachusetts Council of Constr. Employers, 460 U.S. 204 (1983).

54 Reeves, Inc. v. Stake, 447 U.S. 429 (1980).

55 *See* West Lynn Creamery, Inc. v. Healy, 512 U.S. 186, 199 n.15 (1994). Despite the Court's disclaimer in *West Lynn*, one of the leading market-participant cases, *Hughes v. Alexandria Scrap Corp*., 426 U.S. 794 (1976), involved state payments of incentive "bounties" to in-state scrap dealers, payments which look quite similar to simple subsidies, although the Supreme Court characterized them as state participation in the market for the processing of abandoned vehicle hulks. *See also* New Energy Co. v. Limbach, 486 U.S. 269, 277 (1988) (noting that *Alexandria Scrap*'s reasoning may not apply to typical subsidy programs).

This differential treatment of financially equivalent incentives constitutes a troubling anomaly, particularly in light of the Court's asserted focus on the "practical effects" of state measures that favor in-state economic activity. Of course, beyond their simple financial valuation, there are a number of significant differences between tax breaks and cash incentives, which make the use of tax breaks more attractive both to governments and to businesses, and which may justify closer scrutiny of the former than the latter. From the governmental perspective, tax breaks are typically far more readily enacted, without the need to compete in the annual appropriations process and with far less transparency; from the business perspective, they smack far less of governmental hand-outs.

Still, defenders of state tax incentives against Commerce Clause invalidation often underscore the foolishness of interpreting the Constitution to ban one kind of measure when states can, and do, deliver precisely the same financial rewards by other means, free of any constitutional constraint. And at least one commentator has suggested that the arbitrariness of the distinction between cash incentives and tax breaks reveals the bankruptcy of the Court's entire anti-discrimination jurisprudence and argues for judicial withdrawal from the field.[56]

Of course, nothing in the Commerce Clause or in the Court's anti-discrimination framework dictates the blanket insulation of all forms of preferential governmental spending from Commerce Clause scrutiny that the market-participant case law suggests. The European Court of Justice, in addressing a comparable issue in defining impermissible "state aid," has deployed a far narrower and more nuanced "market investor" test, which insulates state measures from treatment as state aid only if the resultant advantage for the local business is one which might similarly have been obtained from a private business behaving "under normal market conditions."[57]

Perhaps it is the relative infrequency with which American governments—as contrasted with their European counterparts—have historically participated directly in commercial markets that has invited and rendered viable a broad-brush immunity from Commerce Clause scrutiny for all forms of market participation. And perhaps, with the growing scale of state interventions in the market in favor of local economic activity, the U.S. Supreme Court will move toward an approach to market participation more like the EC Court's. Indeed, the Court's caution not to pre-judge the question of whether direct subsidies constitute market participation, and its recent references to the "narrow exception" for market

56 *See* Edward A. Zelinsky, *Restoring Politics to the Commerce Clause: The Case for Abandoning the Dormant Commerce Clause Prohibition on Discriminatory Taxation*, 29 OHIO N.U. L. REV. 29 (2002).

57 *See* Kelyn Bacon, *The Concept of State Aid: The Developing Jurisprudence in the European and UK Courts*, 2003 EUR. COMPETITION L. REV. 54, 55 (quoting Case C-256/97 Demenagements-Manutention Transport, 1999 E.C.R. I-3913, 1999 C.M.L.R. 1, 22).

participants,[58] may presage such a development. But, in the meanwhile, the Court's bright-line market-participant exception limits the efficacy and cogency of its anti-discrimination jurisprudence.

B. ANTI-DISCRIMINATION'S UNDER- AND OVER-BREADTH

Even within the more limited confines of tax incentives, the Court's focus on whether a measure discriminates in its treatment of in-state and out-of-state activity is an imperfect tool for singling out those tax measures used by the states to lure businesses in violation of common-market norms. In fact, as discussed above, a number of commonly used incentive measures, such as abatements of property taxes for new plants or remission of sales taxes on new machinery and equipment, are generally non-discriminatory (absent any location-based strings attached to the tax benefits) because they involve taxes that simply are not applicable to out-of-state activity. But such provisions are largely indistinguishable, in their economic effects on both businesses and states, from their discriminatory counterparts. Similarly, the Court has concluded that a state's use of a so-called "single sales factor" income apportionment formula, which looks exclusively to the location of the taxpayer's sales in determining what share of the taxpayer's income the state can tax (contrary to what the Court has recognized as the benchmark approach which averages the proportions of a taxpayer's property, payroll, and sales in the state), does not discriminate in favor of in-state production,[59] notwithstanding the rapid proliferation of the single sales factor methodology as a leading location incentive and protectionist device.

Conversely, the anti-discrimination principle threatens to invalidate a range of tax breaks that are intended to serve purposes quite distant from competing for interstate business, but which are nonetheless conditioned on some type of in-state activity. For example, at the oral argument in *Cuno*, Chief Justice Roberts asked whether the Commerce Clause would forbid homestead exemptions from local property taxes, since such exemptions—because they are typically restricted to a taxpayer's primary residence—are limited to homeowners who are in-state residents. Similar concerns might be raised, for example, about state tax credits for the installation of pollution abatement equipment, which are unsurprisingly restricted to equipment installed at facilities in the taxing state. Perhaps it can be argued that such measures, in practical fact, have neither the intent nor the effect of providing an economic advantage to in-state economic activity, or perhaps they can be defended on the basis of their obvious and substantial non-discriminatory purposes. Still, the suggestion that such provisions facially discriminate between in-state and out-of-state economic activity, and thus should be found "virtually per se unconstitutional," is plausible enough to raise serious concerns about whether the anti-discrimination principle is too broad.

58 *See, e.g., Camps Newfound*, *supra* n. 12, at 589.

59 *See* Moorman Mfg. Co. v. Bair, 437 U.S. 267, 277-79 (1978).

These failings of the anti-discrimination principle reflect the difficult challenge that the courts face in attempting to articulate a workable test for impermissible state interference with interstate economic activity. Such a test must draw a reasonably definite line along a murky continuum. On one end of the continuum are tariffs, which directly and exclusively tax out-of-state production and which the Constitution surely sought to forbid; at the other end are state choices about what sorts of taxes to impose and at what rates, decisions which surely fall within the legitimate sphere of state autonomy in a federalist system. In between them lies a virtually infinite array of possible ways that a state can modify elements of its tax system to create a more favorable economic climate or to reduce the burdens of taxation on local businesses.

All of the measures along the continuum can potentially be seen as interfering with the free flow of economic activity in the national common market, since all of them can affect (and are commonly designed to affect) business choices about where to locate. To enable the Commerce Clause to serve as a workable judicial constraint on measures at one end of the continuum, while not intruding on those at the other end, the Supreme Court has settled on the anti-discrimination principle as a way of drawing an intelligible line between the presumptively permissible and the presumptively unconstitutional measures, a line that identifies differential treatment of in-state and out-of-state activity as the critical threshold that states cannot cross. And, while other elements of the Court's Commerce Clause jurisprudence have been repeatedly revised and reversed, this element has remained a steady and generally effective standard for more than a century. While far from perfect, it may be as good a simple standard as courts can devise.

C. LIMITATIONS ON THE JUDICIAL ROLE

Aside from these difficulties with the doctrinal framework that the courts have developed to assess challenged state tax incentives, reliance on the courts as the enforcers of constitutional limits on state efforts to favor in-state activity is itself problematic in a number of respects. Courts can only intervene in particular cases, challenging specific state actions or measures, and can only do so at the instigation of parties who are willing and able to invoke the courts' jurisdiction. The result is, at best, a rather episodic and haphazard oversight of state efforts to further their parochial interests.

One particularly significant difficulty, already alluded to above, arises from restrictions on judicial standing, that is, on who has the right to bring a case in the courts. While many state courts take more liberal approaches to standing, the federal courts require plaintiffs to have a direct and personal relationship to the challenged action, in the form of a direct injury, which is distinct from injuries suffered by the general public, and which will be alleviated by the requested judicial intervention. Thus, the federal courts are generally unreceptive to challenges to state taxing and spending policies that are brought by citizens or taxpayers whose interest is simply to protect the state fisc from unconstitutional expenditures or losses of revenue. As was noted earlier, in *Cuno*, for example, the reason that the Supreme Court declined to reach the question of the constitutionality of Ohio's investment tax credit and vacated the Sixth Circuit's finding of unconstitutionality was that the citizens and small businesses who brought the challenge did not allege an injury that was "concrete and particularized," but instead asserted a grievance that they suffered "in some indefinite way in common with people generally."[60]

Limitations on standing do not bar all potential challengers of state tax incentives from access to the courts. Most of the past Commerce Clause anti-discrimination cases in the federal courts were brought by businesses who were not receiving the benefits of favorable tax treatment available to their competitors, and such competitors certainly would have standing to challenge location incentives that favored their in-state competition.[61] In addition, a state would most likely have standing to challenge an incentive measure offered by another state, which threatened to encourage businesses to shift their activity away from the state bringing the suit.[62] But, at present, neither competing businesses nor states seem promising potential plaintiffs. The businesses are typically receiving comparable tax benefits from the states where they have located their activity, and will be cautious about endangering those benefits as well as their continuing ability to obtain new tax breaks in connection with future decisions, while states are hesitant to bring a challenge which would likely invalidate some of their own incentives, along with those of other states. And, while citizen plaintiffs may often have access to state courts, with their typically more permissive standing rules, the restrictive federal standing doctrine will preclude them from appealing an unfavorable state court ruling to the Supreme Court.[63]

These standing barriers—and the disinclination, on the part of those who do have standing, to challenge a status quo from which they benefit—probably explain the paucity of case law applying the well-established, anti-discrimination

60 DaimlerChrysler Corp. v. Cuno, *supra* n. 48, at 1854, 1862. In fact, many of the plaintiffs in *Cuno* had indeed suffered a direct and personal injury, since they had lost their homes and businesses to make way for the new plant, but, because these injuries would not have been redressed by a judicial ruling invalidating the tax incentives, those injuries were irrelevant to their standing.

61 For one recent example of such a case, see *Northwest Airlines, Inc. v. Wisconsin Dept. of Revenue*, 717 N.W.2d 280 (Wis. 2006), where Northwest challenged a property tax incentive limited to airlines with a hub in Wisconsin. If Northwest seeks Supreme Court review of the state supreme court's decision, standing will be no obstacle.

62 *See, e.g.* Wyoming v. Oklahoma, 502 U.S. 437 (1992).

63 If citizen plaintiffs should win a Commerce Clause challenge in the state courts, the Supreme Court would not be barred from reviewing that decision, on the request of the state or an affected business taxpayer. *See* ASARCO Inc. v. Kadish, 490 U.S. 605 (1989). This is one plausible route by which this issue may ultimately find its way to the Supreme Court, although state courts are likely to be less sympathetic than federal courts to challenges to their own states' tax provisions and hence less likely to reach decisions that would open the door to Supreme Court review.

principle to the wide range of state business tax incentives that have become so common in recent decades. And they raise serious doubts about the efficacy of judicial intervention as a way to set meaningful limits to the continued proliferation of state incentives.

Moreover, even when a successful case is pursued through the courts, the result is limited to the invalidation of the particular incentive challenged in that case. A judicial decision does not, by its own force, affect even quite similar measures in place in other jurisdictions, and states will be quick to argue, as they did in the wake of the Sixth Circuit decision in *Cuno* invalidating Ohio's ITC, that their comparable provisions are significantly distinguishable from the invalidated measure. If the judicial decision comes from a court of limited geographic jurisdiction (that is to say, from any court other than the Supreme Court), the arguments for the decision's inapplicability to other states' measures will only be reinforced. Thus, at least the short-term effect of judicial invalidation of a state tax incentive will likely be to take a tool out of the hands of one state while leaving comparable tools in the hands of many others, a result that hardly furthers the Commerce Clause goal of placing the states on a level playing field. In addition, long experience with judicial enforcement of Commerce Clause limits on state tax measures suggests that invalidation of one type of measure only spurs the states to devise new and different techniques to achieve comparable effects, techniques whose unconstitutionality can only be tested when proper parties step forward to bring yet another lawsuit. Case-by-case adjudication is a clumsy tool for enforcing a national free trade tax policy.

D. RESIDUAL CONGRESSIONAL AUTHORITY

The courts' role as enforcers of Commerce Clause limits on state actions is, of course, a derivative one. The Commerce Clause is primarily a grant of regulatory authority over interstate commerce to Congress, and the courts only deploy the dormant Commerce Clause in the absence of congressional action. It is well-established that, whatever the courts may do in the face of legislative silence, Congress retains the power not only to regulate interstate commerce by affirmative measures, and not only to forbid particular forms of state discrimination against interstate commerce, but also to delegate particular aspects of its authority over interstate commerce to the states, and thereby to authorize them to engage in conduct, even discriminatory conduct, which would, absent congressional authorization, be found unconstitutional.[64]

64 *See, e.g.,* Western & Southern Life Ins. Co. v. State Bd. of Equalization, 451 U.S. 648 (1981) (construing McCarran-Ferguson Act as granting states plenary power to regulate insurance industry and thereby as authorizing discriminatory state taxation of insurance companies).

Exercising its Commerce Clause authority, Congress has occasionally enacted specific prohibitions on discriminatory state taxation.[65] And it has included a number of restrictions on state measures designed to influence business location decisions as conditions on participation in federally funded economic development programs.[66] Some critics of the proliferation of state location incentives have called for federal legislation as the preferred way to halt or limit the interstate competition, and bills have occasionally been filed in Congress to forbid certain kinds of location incentives, or to impose federal taxes that would negate their benefits.[67]

In the present political climate, however, congressional intervention is far more likely on the opposite side—to protect the ability of states to offer tax incentives to reward in-state economic activity. In fact, in the immediate wake of the Sixth Circuit's decision invalidating Ohio's investment tax credit, the senators from the states in the Sixth Circuit (Ohio, Michigan, Tennessee, and Kentucky) filed legislation to override the court's ruling, reflecting their concern that the ruling, unless and until reviewed by the Supreme Court, would place their states at a competitive disadvantage by taking out of their hands, but not competing states' hands, some of the key tools for influencing business location decisions. In the subsequent congressional session, a broader coalition of senators and congressmen filed a far more comprehensive bill (S. 1066) that would broadly authorize the states, with limited exceptions, "to provide . . . for economic development purposes tax incentives that otherwise would be the cause or source of discrimination against interstate commerce under the Commerce Clause."[68]

The effect of this legislation would be to turn the present federal approach to discriminatory state tax measures on its head. Now, discriminatory provisions are presumed to be unconstitutional, unless they can be shown necessary to serve an important non-discriminatory state purpose. But under the provisions of S. 1066, the so-called "Economic Development Act," discriminatory measures would be presumed to be permissible, so long as they were intended for economic development purposes, unless they fell within one of the Act's specific exceptions, which were crafted to avoid overruling a number of the Supreme Court's prior cases.

---

65 *See, e.g.,* 42 U.S.C. § 11503 (barring discriminatory state taxation of railroad property).

66 *See, e.g.,* 26 U.S.C. § 1391(d)(1)(F) (imposing "anti-piracy" conditions on federal designation of empowerment or enterprise zones, eligible for favorable federal tax treatment).

67 *See, e.g.,* Melvin L. Burstein & Arthur J. Rolnick, *Congress Should End the Economic War Among the States*, 9 REGION 3 (1995) (Federal Reserve Bank of Minneapolis Annual Report); Daniel Shaviro, *An Economic and Political Look at Federalism in Taxation*, 90 MICH. L. REV. 895, 975-88 (1992).

68 S. 1066, 109th Cong. § 2 (2005). Identical legislation, H.R. 2471, 109th Cong. (2005), was filed in the House of Representatives. For further discussion of this proposed legislation, *see Cuno and Competitiveness: Where to Draw the Line: Hearing Before the Subcomm. on International Trade of the S. Comm. on Finance*, 109th Cong., 2nd Sess. (Mar. 16, 2006), *available at* http://finance.senate.gov/sitepages/hearing031606.htm.

Before the Supreme Court's decision vacating the Sixth Circuit's decision in the *Cuno* case, S. 1066 had attracted wide and vocal support, both from powerful segments of the business community and from major groups representing state policymakers, such as the National Governors Association. It was widely anticipated that, had the Court affirmed the lower court's invalidation of Ohio's investment tax credit, the pressure for speedy congressional enactment would have been intense. Instead, the effect of the Court's decision in *Cuno* was to slow the momentum behind the bill substantially. But its sponsors and supporters remain committed to its eventual enactment, and their political clout remains an important reminder of the severe institutional limits on the use of the courts to constrain the economic competition among the states.

V. Conclusion

Thus, reliance on judicial enforcement of the Commerce Clause's anti-discrimination principle to rein in the states' inevitable tendencies to favor their parochial economic interests suffers from a number of serious shortcomings, both doctrinal and institutional. Nonetheless, over much of the nation's history, the courts and the dormant Commerce Clause have played a central role in combating pressures toward economic balkanization and in reinforcing the growth of an open national common market. In fact, the intensification of interstate competition for economic activity over the past few decades, and the potent political forces favoring the continuation of that competition, serve as reminders of the importance of judicially enforced constitutional constraints on these tendencies. While far from perfect, this tool has proven more effective and more dependable than the available alternatives. But, in light of its limitations, one of the key questions for those who seek to protect both the states themselves and the national economy from the harms of interstate competition over tax incentives is whether use of the courts can serve to widen public understanding of those harms and to build political support for limits on the competition.

VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

Identifying, Challenging, and Assigning Political Responsibility for State Regulation Restricting Competition

Maureen K. Ohlhausen

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Identifying, Challenging, and Assigning Political Responsibility for State Regulation Restricting Competition

Maureen K. Ohlhausen

This paper examines the role of competition advocacy in combating anticompetitive state regulation. Looking at the constraints facing competition officials such as the state action doctrine, the analysis suggests potential avenues for surmounting these constraints. Relying on experience as the Director of the U.S. Federal Trade Commission's Office of Policy Planning, the author uses real-world examples—real estate brokerage and interstate direct shipment of wine—to demonstrate the ability of a competition agency to use a variety of techniques to improve consumer welfare when enforcement is circumscribed due to state activity.

The author is the Director of the Office of Policy Planning at the U.S. Federal Trade Commission. The views expressed here are the author's own and not necessarily those of the Federal Trade Commission.

I. Introduction

Antitrust conjures visions of large corporations conspiring behind closed doors to fix prices or powerful monopolies crushing upstart rivals. Competition officials must be alert to threats to competition from all sources, however, even from activities that are seemingly open to public scrutiny. Specifically, an important but sometimes overlooked source of anticompetitive harm is the enactment of state laws or promulgation of state regulations that restrict business activities or prohibit some business models altogether. Those concerned about promoting competition must not overlook the serious harm that can be wrought by state legislation and regulation—even well-intentioned actions—that hamper competition by setting prices, mandating offerings, or fencing out certain types of competitors, and which can inflict as much harm on consumers as does private anticompetitive action.

There are strong incentives for competitors to seek through legislation and regulation what they cannot lawfully obtain through private actions. If private price fixers run the risk of prison while government regulation fixing prices is legal, rational competitors looking for shelter from competitive pressures will seek government action to implement such regulation. In addition to being less risky to attain, anticompetitive government restrictions can also be more effective at restraining competition than private restraints. Public restraints are typically open; they appear in public statutes and regulations. They also are easier to enforce. The government keeps out those who would introduce more competition, either by law enforcement against mavericks who try to enter anyway or by providing a limited number of licenses, regardless of need. As the economic theory of regulation posits, consumers are ill-prepared to counter these efforts politically.[1] Their interests are diffuse and the costs of the restriction for any individual are often small. By contrast, those seeking the restrictions are organized firms or professional associations that will reap concentrated benefits from reduced competition. Finally, as regulation increases, so do the opportunities to use the mechanisms of regulation to keep out rivals.[2]

In the United States, the state action doctrine protects from antitrust enforcement state government action that limits or eliminates competition. When applied properly, this doctrine is necessary to the operation of a representative

1 *See* Gary S. Becker, *A Theory of Competition Among Pressure Groups for Political Influence*, 98 Q.J. Econ. 371 (1983); Sam Peltzman, *Toward a More General Theory of Regulation*, 19 J.L. & Econ. 211, 213 (1976); George J. Stigler, *The Theory of Economic Regulation*, 2 Bell J. Econ. & Mgmt 3, 11 (1971).

2 *See, e.g.*, Robert Bork, The Antitrust Paradox: A Policy at War with Itself 347 (1978) ("In order to enter the market and vie for consumers' favor, businesses of all types must gain various types of approval from governmental agencies, departments, and officials. Licensing authorities, planning boards, zoning commissions, health departments, building inspectors, public utilities commissions, and many other bodies and officials control and qualify the would-be competitor's access to the marketplace.").

democracy in a federal system. The doctrine is not always applied correctly, however, and thus one avenue for limiting consumer harm is to be sure such protection is not interpreted expansively to shield truly private anticompetitive actions. Even when this protection is properly applied and enforcement is not a possibility, however, there are avenues that a competition official can pursue when faced with a state considering the adoption of an anticompetitive law that is likely to harm consumer welfare.

IN THE UNITED STATES, THE STATE ACTION DOCTRINE PROTECTS FROM ANTITRUST ENFORCEMENT STATE GOVERNMENT ACTION THAT LIMITS OR ELIMINATES COMPETITION.

In this article, I will discuss the constraints facing competition officials in the United States and then identify avenues for combating anticompetitive state regulation despite these constraints, which may be useful for any competition official faced with similar challenges. In doing this, I will use real-world examples drawn from my experience as the Director of the U.S. Federal Trade Commission's Office of Policy Planning, which oversees the Commission's efforts to persuade policymakers, including state legislators and regulators, to design policies that further competition and preserve consumer choice.

II. The State Action Doctrine

The state action doctrine, which was first articulated in a 1943 U.S. Supreme Court opinion, *Parker v. Brown*,[3] protects from the reach of the Sherman Act actions taken by a sovereign state. The Court reasoned that "in light of states' sovereign status and principles of federalism, Congress would not have intruded on state prerogatives through the Sherman Act without expressly saying so."[4] The Court held, therefore, that conduct that could be attributed to the state itself is immunized from antitrust scrutiny. Thus, a threshold inquiry for invoking state action immunity is whether the anticompetitive action was by the sovereign or by a private party.

In *California Retail Liquor Dealers Assn v. Midcal Aluminum, Inc.*,[5] the U.S. Supreme Court set forth two important limitations on the scope of state action immunity that help to ensure that the immunized conduct is truly that of the state itself, rather than private action. First, the defendant claiming the immunity must demonstrate that the conduct in question was in conformity with a

3 Parker v. Brown, 317 U.S. 341 (1943).

4 STATE ACTION TASK FORCE, OFFICE OF POLICY PLANNING, FEDERAL TRADE COMMISSION, REPORT OF THE STATE ACTION TASK FORCE (Sep. 2003), *available at* http://www.ftc.gov/os/2003/09/stateactionreport.pdf.

5 California Retail Liquor Dealers Ass'n v. Midcal Aluminum, Inc., 445 U.S. 97 (1980).

"clearly articulated" state policy. Second, the defendant must demonstrate that the state engaged in "active supervision" of the conduct.

Although this rule seems fairly clear in theory, the parameters of the doctrine become substantially less clear when applied to delegations of state authority to private parties, particularly to industry members regulating the conduct of their competitors. There is little argument that the Sherman Act was not intended to reach the conduct of a state legislature that adopts anticompetitive legislation.[6] A more contested issue is under what circumstances the Sherman Act can reach, for example, the anticompetitive conduct of a board of professional licensure, dominated by members of the profession.[7]

Thus, one course to explore for competition officials concerned about anticompetitive state regulation is an evaluation of whether the shelter from antitrust enforcement given to state action is unnecessarily broad. For example, the U.S. Federal Trade Commission (FTC) convened a State Action Task Force to reexamine the scope of the state action doctrine; and make recommendations to ensure that the exemption remains closely tied to protecting the deliberate policy choices of sovereign states, and is otherwise applied in a manner that promotes competition and enhances consumer welfare. The Task Force issued a report in September 2003, which concluded that, since *Parker*, the scope of the doctrine has increased considerably and that both the clear articulation and active supervision requirements have been the subject of varied and controversial interpretation, sometimes resulting in unwarranted expansions of the exemption.[8] To address these problems with the state action doctrine, the *Report of the State Action Task Force* recommended clarifications to bring the doctrine more closely in line with its original objectives, including reaffirming a clear articulation standard tailored to its original purposes and goals, clarifying and strengthening the standards for active supervision, and clarifying and rationalizing the criteria for identifying the quasi-governmental entities that should be subject to active supervision.

6 *See, e.g.*, Hoover v. Ronwin, 466 U.S. 558, 567-68 (1984) ("when a state legislature adopts legislation, its actions constitute those of the State . . . and *ipso facto* are exempt from the operation of the antitrust laws." citations omitted). The Court also extended this ipso facto exemption to a state supreme court acting in a legislative capacity. *Id.* at 568.

7 This issue is likely to continue to grow in importance as the percentage of the labor force in the United States covered by state licensing laws continues to grow. *See* Morris M. Kleiner, Licensing Occupations: Ensuring Quality or Restricting competition? 1 (2006) ("During the early 1950s, only about 4.5 percent of the [U.S.] labor force was covered by licensing laws at the state level. That number had grown to almost 18 percent of the U.S. workforce in the 1980s, with an even larger number if city and county licenses for occupations are included.").

8 State Action Task Force, *supra* note 4.

A. ENFORCEMENT

Despite the lack of clarity regarding the exact parameters of the state action doctrine, it is not necessarily a bar to antitrust enforcement against actions by self-interested state boards, and U.S. antitrust agencies have sued state regulatory boards made of up competitors for restricting competition in ways that the state did not authorize.[9] The doctrine, however, does present an additional hurdle for enforcers to surmount. For example, in 2003, the FTC brought a complaint against the South Carolina Board of Dentistry, alleging that it violated federal laws by illegally restricting the ability of dental hygienists to provide preventive dental services in schools.[10] After the South Carolina General Assembly amended legislation to make it easier for dental hygienists to provide preventive dental care services to children in schools—by removing the requirement of a pre-examination by a dentist—the Board passed an emergency regulation that contradicted the General Assembly's amendments by reinstating the requirement that a dentist examine a patient before the patient is eligible for treatment in school. The FTC's complaint alleged that the Board was not acting pursuant to any clearly articulated state policy to displace competition, thereby suggesting that the conduct would not be immune under the state action doctrine. The Board raised a state action defense, which the FTC ultimately rejected in an adjudicative opinion.[11] The FTC held that although the Board was created by state statute, courts have consistently declined to extend ipso facto state action protection to non-elected governmental entities, particularly state licensing or regulatory boards composed, at least in part, of members of the regulated industry.[12]

Because the Board was not deemed part of the sovereign, the FTC then evaluated whether its action was taken pursuant to a clearly articulated state legislative policy.[13] The FTC reasoned that although South Carolina's statutory

9 State boards that regulate professions have been a particularly rich area for competition scrutiny. *See,* e.g., FTC v. Monahan, 832 F.2d 688, 689-90 (1st Cir. 1987); Massachusetts Bd. of Registration in Optometry, 110 F.T.C. 549, 612-13 (1988); Kentucky Household Goods Carriers Ass'n, FTC Dkt No. 9309 (2004).

10 South Carolina State Board of Dentistry, FTC Docket No. 9311 (2003) (complaint), *available at* http://www.ftc.gov/os/2003/09/socodentistcomp.pdf.

11 South Carolina State Board of Dentistry, FTC Docket No. 9311 (2004), *available at* http://www.ftc.gov/os/adjpro/d9311/040728commissionopinion.pdf.

12 *Id.* at 18 (citing Southern Motor Carriers, 471 U.S. at 62-63; Goldfarb v. Virginia State Bar, 421 U.S. 773, 790-92 (1975); Earles v. State Bd. of Certified Pub. Accountants of Louisiana, 139 F.3d 1033, 1040-41 (5th Cir. 1998)).

13 *Id.* at 22 ("[W]hile clear articulation does not require a state entity to show 'express authorization' for every specific anticompetitive act, . . . it does anticipate that the anticompetitive action will have a significant nexus to, or degree of 'foreseeability' stemming from, an identifiable state policy.") (citing Southern Motor Carriers, 471 U.S. at 64; City of Columbia v. Omni Outdoor Advertising, Inc., 499 U.S. 365, 373 (1991)).

regime gave the Board broad general authority to regulate the fields of dentistry and dental hygiene in the state—thus necessarily allowing the Board to displace competition in certain ways—it was not foreseeable that this grant of general supervisory authority encompassed the right to re-impose the pre-examination requirement that the state legislature had just eliminated. Accordingly, the FTC denied the Board's motion to dismiss the complaint on state action grounds.

B. BEYOND ENFORCEMENT

Enforcement is a highly effective tool to combat private interests that attempt to cloak themselves in a government mantle to attain anticompetitive ends. It is not the only tool, however, and where competition officials cannot pursue enforcement actions because the conduct is either that of the sovereign or is pursuant to a clearly articulated and actively supervised state policy, there are still avenues to pursue. One option is a form of persuasion called competition advocacy, which can be broadly described as the use of expertise in competition, economics, and consumer protection to persuade government actors to tailor their policies to protect or foster competition. In addition to reaching beyond where enforcement can go, competition advocacy can also be a cost-effective way to deploy resources to safeguard consumer welfare,[14] which makes it particularly appealing to small and newly created competition agencies that may have insufficient means to support more resource-intensive enforcement actions.

ENFORCEMENT IS A HIGHLY EFFECTIVE TOOL TO COMBAT PRIVATE INTERESTS THAT ATTEMPT TO CLOAK THEMSELVES IN A GOVERNMENT MANTLE TO ATTAIN ANTICOMPETITIVE ENDS. IT IS NOT THE ONLY TOOL, HOWEVER, AND WHERE COMPETITION OFFICIALS CANNOT PURSUE ENFORCEMENT ACTIONS BECAUSE THE CONDUCT IS EITHER THAT OF THE SOVEREIGN OR IS PURSUANT TO A CLEARLY ARTICULATED AND ACTIVELY SUPERVISED STATE POLICY, THERE ARE STILL AVENUES TO PURSUE.

In addition to being a cost-effective way to reduce consumer harm from anticompetitive state actions, competition advocacy can also serve an important function in the political process. In a leading state action case, *Federal Trade Commission v. Ticor Title Insurance Co.*, the U.S. Supreme Court observed that "[s]tates must accept political responsibility for the actions they intend to take . . . Federalism serves to assign political responsibility, not obscure it."[15] Competition advocacy, even when unsuccessful in influencing a

14 A 1989 American Bar Association Report observed: "Because ill-advised governmental restraints can impose staggering costs on consumers, the potential benefits from an advocacy program exceed the Commission's entire budget." AMERICAN BAR ASSOCIATION SECTION OF ANTITRUST LAW SPECIAL COMMITTEE TO STUDY THE ROLE OF THE FEDERAL TRADE COMMISSION, REPORT (1989), *reprinted in* 58 ANTITRUST L.J. 43, 116 (1989).

15 FTC v. Ticor Title Ins. Co., 504 U.S. 621, 636 (1992).

particular state regulation, highlights the costs to consumers of the anticompetitive state regulation and helps assign political responsibility to the state policymakers endorsing the action.

Viewed through the lens of the economic theory of regulation, competition advocacy "helps solve consumers' collective action problem by acting within the political system to advocate for regulations that do not restrict competition unless there is a compelling consumer protection rationale for imposing such costs on consumers."[16] It inserts a voice for otherwise overlooked consumer interests in a political debate typically dominated by organized interests with strong incentives to seek government protection from competition.

III. Overview of the FTC Competition Advocacy Program

The FTC has long had an appreciation for the benefits that advocacy can achieve and has conducted an advocacy program in one form or another for quite some time.[17] Through this program, it has often persuaded state policymakers to eschew anticompetitive proposals or to modify them to reduce the impact on competition or at least drawn public, political, and academic attention to competitive restrictions in the states. Competition advocacy can take a variety of forms, with the most common being letters from the FTC or its staff (sometimes joined by the Antitrust Division of the U.S. Department of Justice (DOJ)) to state legislators, regulatory boards, or governors.[18] The Commission has also filed amicus briefs with state supreme courts considering issues involving state professional licensing requirements,[19] and with national professional associations pro-

16 James C. Cooper, et al., *Theory and Practice of Competition Advocacy at the FTC*, 72 Antitrust L.J. 1091, 1092 (2005).

17 For additional history of the FTC competition advocacy program and various views on it, *see id*. and Arnold J. Celnicker, *The Federal Trade Commission's Competition and Consumer Advocacy Program*, 33 St. Louis U. L.J. 379 (1988).

18 The FTC and its staff also file competition advocacy comments with other federal agencies, typically in response to requests for public comment on pending federal regulations. *See, e.g.*, In the Matter of Request for Comment on Proposed Amendments to the Regulations Implementing the Real Estate Settlement Procedures Act, Docket No. FR-4727-P-01 (Dep't Housing & Urban Development, Oct. 28, 2002) (Comments of the Staff of the Bureau of Economics, the Bureau of Consumer Protection, and the Office of Policy Planning of the Federal Trade Commission), *available at* http://www.ftc.gov/be/v030001.pdf; In the Matter of Request for Comments on Consumer-Directed Promotion, Docket No. 2003N-0344 (Food & Drug Admin. Dec. 1, 2003) (Comments of the Staff of the Bureau of Consumer Protection, the Bureau of Economics, and the Office of Policy Planning, Federal Trade Commission) *available at* http://www.ftc.gov/be/v040002text.pdf.

19 *See, e.g.*, Brief *Amicus Curiae* of the United States of America and the Federal Trade Commission, Lorrie McMahon v. Advanced Title Services Co. of West Virginia, 607 S.E.2d 519 (W. Va. 2004) *available at* http://www.ftc.gov/be/V040017.pdf; Brief *Amicus Curiae* of the Federal Trade Commission, Cleveland Bar Ass'n v. CompManagement, Inc., 104 Ohio St. 3d 168 (2004) *available at* http://www.ftc.gov/os/2004/08/040803amicusbriefclevbar.pdf.

posing model rules that would ultimately be promulgated by state regulatory bodies.[20] One of the most influential means of promoting competition has been through in-depth research conducted by FTC legal and economic staff, resulting in staff studies of certain industries, as well as scholarly reports about antitrust doctrines, such as the *Report of the State Action Task Force*, discussed earlier in this paper.[21] Such studies and reports are often the result of workshops that the FTC staff holds periodically, which focus on specific industries[22] or trends affecting competition more broadly.[23] A course of competition advocacy need not follow any particular order—comments may precede or follow workshops and studies may be the starting point or the conclusion of an inquiry. What is crucial to effective competition advocacy is that it be based on a comprehensive understanding of the industry at issue, competition principles, economic theory, and available empirical evidence. In addition to formal actions, informal presentations and contacts can also be helpful. Thus, FTC staff and Commissioners also promote competition principles through a host of activities, such as speeches before associations of state regulators or industry members, interviews with the press, and articles in general interest publications. Finally, the attention competition advocacy brings to a topic often sparks legal and economic research by legal and economic researchers, whose work adds to the body of knowledge about competition issues in a particular industry.

To give a better idea of what competition advocacy may cover and what it can accomplish, I will discuss recent competition advocacy initiatives, describing their genesis, form, and results. These two areas—real estate brokerage and the interstate direct shipment of wine—are particularly good examples of a competition agency using a variety of techniques to improve consumer welfare when its enforcement is circumscribed due to state activity.

A. REAL ESTATE BROKERAGE

The FTC has long been concerned about anticompetitive practices in real estate brokerage, such as efforts by private associations of brokers to disadvantage brokers who use non-traditional listing agreements that are associated with lower

20 *See* Letter from FTC and the Justice Department to the Task Force on the Model Definition of the Practice of Law of the American Bar Association (Dec. 20, 2002), *available at* http://www.ftc.gov/opa/2002/12/lettertoaba.htm.

21 *See also* Staff of the FTC, Report, Enforcement Perspectives on the Noerr Doctrine (forthcoming).

22 *See, e.g.,* Federal Trade Commission & U.S. Dep't of Justice, Workshop on Competition in Real Estate Brokerage (Oct. 2005).

23 Federal Trade Commission, Workshop on Possible Barriers to Competition in E-commerce (Oct. 2002).

commission rates or flat fee services.[24] This focus has included competition advocacy in connection with a number of issues related to real estate transactions, such as laws that restrict non-attorneys from performing certain aspects of real estate closings[25] and minimum-service brokerage laws, which generally require all real estate agents, regardless of their fee structure, to provide most of the services supplied by traditional full-service agents. Also, a number of years ago, the FTC released a comprehensive report on the real estate brokerage industry reflecting years of enforcement activity and industry research,[26] and is currently exploring the feasibility of updating this research.

In recent years, technological developments have spurred a number of substantial changes in the real estate industry. Agents are increasingly incorporating the Internet into their business models in a variety of ways, such as offering potential buyers the option to view detailed property listing information online, or using websites to gather lead information on customers who seek real estate services and then selling those leads to real estate professionals. Still other business

---

24 *See* In the Matter of Austin Board of Realtors, File No. 0510219 (F.T.C. Jul. 13, 2006), *available at* http://www.ftc.gov/os/caselist/0510219/0510219.htm; In the Matter of United Real Estate Brokers of Rockland, Ltd., 116 F.T.C. 972 (1993), *available at* http://www.ftc.gov/bc/realestate/cases/UnitedRealEstateBrokersofRocklandLtd116FTC972.pdf; In the Matter of American Industrial Real Estate Association, 116 F.T.C. 704 (1993), *available at* http://www.ftc.gov/bc/realestate/cases/AmericanIndustrialRealEstateAssociationetal116FTC704.pdf; In the Matter of Puget Sound Multiple Listing Association, Docket No. C-3300 (F.T.C. Aug. 2, 1990), *available at* http://www.ftc.gov/bc/realestate/cases/PugetSoundMultipleListingAssociation113FTC733.pdf; In the Matter of Bellingham-Whatcom County Multiple Listing Bureau, Docket No. C-3299 (F.T.C. Aug. 2, 1990), *available at* http://www.ftc.gov/bc/realestate/cases/Bellingham-WhatcomCountyMultipleListingBureau113FTC724.pdf; In the Matter of Metro MLS, Inc., 115 F.T.C. 305 (1990), *available at* http://www.ftc.gov/bc/realestate/cases/MetroMLS113FTC305.pdf; In the Matter of Multiple Listing Service of the Greater Michigan City Area, Inc., 106 F.T.C. 95 (1985), *available at* http://www.ftc.gov/bc/realestate/cases/MultipleListingServiceoftheGreaterMichiganCityAreaInc106FTC95.pdf; In the Matter of Orange County Board of Realtors, Inc., 106 F.T.C. 88 (1985), *available at* http://www.ftc.gov/bc/realestate/cases/OrangeCountyBoardofRealtorsIncetal106FTC88.pdf.

25 *See* Letter from FTC & U.S. Dep't of Justice to Massachusetts State Representative Paul Kujawski (Oct. 6, 2004); Letter from FTC & U.S. Dep't of Justice to Standing Committee on the Unlicensed Practice of Law, State Bar of Georgia (Mar. 20, 2003); Letters from the FTC & U.S. Dep't of Justice to Speaker of the Rhode Island House of Representatives and to the President of the Rhode Island Senate, *et al*. (June 30, 2003 and Mar. 28, 2003); Letter from the FTC & U.S. Dep't of Justice to President of the North Carolina State Bar (July 11, 2002); Letter from the FTC & U.S. Dep't of Justice to Speaker of the Rhode Island House of Representatives, *et al*. (Mar. 29, 2002); Letter from the FTC & U.S. Dep't of Justice to the Ethics Committee of the North Carolina State Bar (Dec. 14, 2001); Letter from the FTC & U.S. Dep't of Justice to the Supreme Court of Virginia (Jan. 3, 1997); Letter from the FTC & U.S. Dep't of Justice to the Virginia State Bar (Sept. 20, 1996), *available at* http://www.ftc.gov/be/advofile.htm. *See also,* Brief *Amicus Curiae* of the United States of America and the Federal Trade Commission, Lorrie McMahon v. Advanced Title Services Co. of West Virginia, 607 S.E.2d 519 (W. Va. 2004), *available at* http://www.ftc.gov/be/V040017.pdf; Brief *Amicus Curiae* of the Federal Trade Commission and the United States of America, On Review of ULP Advisory Opinion 2003-2 (Ga. July 28, 2003), *available at* http://www.ftc.gov/os/2003/07/georgiabrief.pdf.

26 *See* Staff of the FTC, Report, The Residential Real Estate Brokerage Industry (1983), *available at* http://www.ftc.gov/bc/realestate/workshop/index.htm.

models use the Internet to match home buyers and sellers. The increased ease with which home sellers can perform tasks that once were the exclusive domain of brokers likely has been an important factor in the increased demand for innovative, non-traditional brokerage services. One form of non-traditional brokerage service is limited-service brokerage, pursuant to which a home seller might choose to pay a broker only for the service of listing the home in the local multiple listing services and placing advertisements, and choose to handle negotiations and paperwork himself or herself. This model gives the consumer the choice to save potentially thousands of dollars in commissions in exchange for taking on more work.

As alternative brokerage models have grown in prominence, several state legislatures and real estate commissions—at the urging of state real estate agent associations—have considered or adopted minimum-service requirements, which would have the effect of forcing consumers to purchase a state-mandated bundle of real estate brokerage services that conform more closely to the array of services offered by traditional, full-service brokers.[27]

In 2005, the FTC, along with the DOJ's Antitrust Division, sent letters to the Texas Real Estate Commission,[28] the Alabama Senate,[29] the governor of the state of Missouri,[30] and to a Michigan state senator[31] providing analysis of the likely competitive effects of proposed minimum-service laws. The comments asserted that by effectively eliminating many of the most popular packages offered by limited-service brokers, these minimum-service laws would reduce consumer choice and competition among traditional brokerage models and limited-service models. They further noted the dearth of evidence that such laws are necessary to protect consumers and that staff was never presented with evidence of actual consumer harm from the limited-service brokerage model. In the end, Texas, Alabama, and Missouri adopted minimum-service laws. The advocacy filing

27 It is common for industry specific (and at times even identical) anticompetitive prohibitions on entry by certain types of competitors or restrictions on certain business models to appear in a number of states at the same time. *See generally, supra* note 25; A.C. Pritchard & Todd J. Zywicki, *Finding the Constitution: An Economic Analysis of Tradition's Role in Constitutional Interpretation*, 77 N.C. L. Rev. 409, 486 (1999).

28 Letter from the FTC & U.S. Dep't of Justice to Loretta R. DeHay, Gen. Counsel, Texas Real Estate Comm'n. (Apr. 20, 2005), *available at* http://www.usdoj.gov/atr/public/press_releases/2005/208653a.htm.

29 Letter from the FTC & U.S. Dep't of Justice to Alabama Senate (May 12, 2005), *available at* http://www.ftc.gov/os/2005/05/050512ltralabamarealtors.pdf.

30 Letter from the FTC & U.S. Dep't of Justice to Governor Matt Blunt (May 23, 2005), *available at* http://www.ftc.gov/opa/2005/05/mrealestate.htm.

31 Letter from FTC & U.S. Dep't of Justice to Michigan State Sen. Alan Sanborn (Oct. 18, 2005), *available at* http://www.ftc.gov/os/2005/10/051020commmihousebill4849.pdf.

appears to have had more immediate success in Michigan, where the proposal failed to pass in the legislature's most recent term.

Despite this limited success in directly persuading state policymakers to reject anticompetitive restrictions on non-traditional business models, there are still other avenues to pursue. One path is to conduct a careful analysis of the marketplace that policymakers and opinion leaders—and eventually the public—may come to rely on in evaluating the conduct of the industry and the state officials who have adopted anticompetitive restrictions favorable to the entrenched business interests. Thus, the FTC and DOJ held a workshop addressing competition policy and the real estate industry in late 2005 to provide a forum to discuss current issues affecting the competitiveness of this important market. At the workshop, a variety of panelists, including practitioners, economists, and state administrators, provided their various views on competition in the real estate brokerage industry. In addition, the agencies received almost 400 submissions in response to their request for public comment in connection with the workshop. The FTC and the DOJ plan to release a report in late 2006 based on information gathered in connection with the workshop and research conducted by staff. To aid those interested in following these activities more closely, the Commission also has launched a website that allows the public to find all of the FTC's work in the real estate area through one central portal.[32]

The sustained focus on competition in real estate brokerage has spurred ongoing press interest, with numerous stories in national newspapers.[33] The U.S. Congress has also taken up the issue, with the House Subcommittee on Housing and Community Opportunity holding hearings on competition in the real estate brokerage industry in July 2006.[34] These inquiries raise consumer awareness of their state representatives' actions that may not advance consumer welfare, thereby helping to assign political responsibility to those policymakers.

B. INTERSTATE DIRECT SHIPMENT OF WINE

Another recent area of extensive competition advocacy activity involves the ability of wineries to ship their wines directly to consumers throughout the United States. Alcohol is heavily regulated in the United States, and the 21st Amendment to the U.S. Constitution, which repealed Prohibition, gives the

32 *See* FTC Bureau of Competition, *Competition in the Real Estate Market Place* (last modified Jul. 19, 2006), *available at* http://www.ftc.gov/bc/realestate/index.htm.

33 *See, e.g., Realtor Racket*, Wall St. J., Aug. 12, 2005, *available at* http://online.wsj.com/article/0,,SB112381069428011613,00.html.

34 *See Hearing on the Real Estate Market and the Development of the Internet in the Real Estate Sector Before the Subcomm. on Hous. & Community Opportunity of the House Comm. on Financial Services*, 109th Cong. (2006) (The author testified in her capacity as Director of the Office of Policy Planning at the FTC), *available at* http://financialservices.house.gov/hearings.asp?formmode=detail&hearing=497&comm=5.

states special authority to regulate it. Pursuant to this authority, all fifty states have required wine to pass through a wholesaler and bricks and mortar retailer before reaching consumers. In recent years, however, the Internet has become a popular avenue to buy wine. Consumers can buy literally thousands of varieties over the Internet directly from the winery, often at lower prices than elsewhere. Direct shipment is a particularly attractive channel for small wineries, which often have difficulty getting distributors to carry their offerings.[35] Not surprisingly, some traditional firms—primarily wholesalers—perceived the Internet as a significant threat, and they successfully lobbied a number of state legislatures to prohibit wineries from shipping directly to consumers, largely on the theory that underage drinkers could buy wine online. Seven states even made it a felony to ship wine directly.

In 2002, the FTC held a workshop on possible barriers to e-commerce that, among other topics, examined issues surrounding the interstate direct shipment of wine. At the workshop, FTC staff heard testimony from all sides of the wine issue, including wineries, wholesalers, and state regulators. The staff also gathered evidence from package delivery companies, the U.S. Alcohol and Tobacco Tax and Trade Bureau, and regulators in states that allow direct shipping. In addition, FTC staff conducted the first empirical study of a wine market in a state that banned interstate direct shipping.

In 2003, the FTC staff issued a report (*Wine Report*) on state restrictions on the direct shipment of wine from out-of-state vendors to in-state consumers.[36] The staff report, reflecting the unique interest and sensitivity of the Commission to both competition and consumer protection concerns, concluded that states could significantly enhance consumer welfare by allowing the direct shipment of wine as a purchase option. The report supported this conclusion with a study conducted by FTC economists that showed that many wines available to consumers online are not available in local retail outlets and that consumers could save money if they purchased their more expensive wines online.[37] Using the *Wine & Spirits* list of the "Top 50 Wines" in America, the study found that 15 percent of a sample of wines available online was not available from retail wine

35 The U.S. Supreme Court recently noted that "many small wineries do not produce enough wine or have sufficient consumer demand for their wine to make it economical for wholesalers to carry their products. This has led many small wineries to rely on direct shipping to reach new markets." Granholm v. Heald, 125 S. Ct. 1885, 1892 (2005) (citation omitted).

36 STAFF OF THE FTC, REPORT, POSSIBLE ANTICOMPETITIVE BARRIERS TO E-COMMERCE—WINE (July 2003), *available at* http://www.ftc.gov/os/2003/07/winereport2.pdf.

37 The study appears as an appendix to the FTC staff report. *Id.* It was published separately as ALAN E. WISEMAN & JERRY ELLIG, HOW MANY BOTTLES MAKE A CASE AGAINST PROHIBITION? (FTC Bureau of Economics Working Paper No. 258, Mar. 2003) and later published as Alan E. Wiseman & Jerry Ellig, *Marketing and Nonmarket Barriers to Internet Wine Sales: The Case of Virginia,* 6 BUS. & POL. 4 (Aug. 2004), *available at* http://www.bepress.com/bap/vol6/iss2/art4. The authors explicitly note that a full welfare analysis of the removal of restrictions would require additional data.

stores within ten miles of McLean, Virginia. Given that the wines studied are the most popular wines of many of America's largest wineries, it is likely that the wines of less-popular or smaller wineries are even more difficult to locate in wine retailers. Moreover, the same study suggested that, if consumers use the least expensive shipping method, they could save an average of 8-13 percent on wines costing more than US$20 per bottle, and an average of 20-21 percent on wines costing more than US$40 per bottle. Less expensive wines may be cheaper in bricks and mortar stores, given that fixed shipping costs will be proportionately larger for less expensive wines.

At the workshop, some parties expressed concern and offered anecdotes suggesting that interstate direct shipping might have the unintended effect of increasing underage access to alcohol or undermining tax compliance. To determine whether these concerns were factually grounded, FTC staff contacted numerous officials from states that allow direct shipping to gather systematically information about whether these problems have occurred.

Given that underage drinking is a serious health and safety issue, the *Wine Report* undertook an in-depth analysis of this issue. The report concluded, however, that there is no systematic evidence of problems of Internet-related shipments to minors. The *Wine Report* stated that, in general, state officials report that they have experienced few, if any, problems with direct shipments of wine to minors, especially when compared with the problem of underage access to alcohol through traditional distribution channels. In addition, several states that permit interstate direct shipping have adopted various procedural safeguards and enforcement mechanisms to prevent sales to minors. These include such precautions as requiring labeling of packages containing wine and requiring an adult signature at the time of delivery. For example, the state of New Hampshire developed penalty and enforcement schemes in coordination with its enforcement agencies.

The *Wine Report* also found that some states also have adopted less restrictive means of protecting tax revenues while permitting direct shipping, such as by requiring out-of-state suppliers to obtain permits and to collect and remit taxes. Most of these states reported few, if any, problems with tax collection.

Finally, the report uncovered little actual evidence to support the distinction found in several states that permit intrastate direct shipment of wine but prohibit interstate shipment. While some parties provided theoretical justifications for the distinction, the report found no evidence based on the experience of state law enforcement authorities to justify the distinction in practice.

The issue of whether states could prohibit out-of-state sellers from shipping wine to consumers while allowing in-state wine producers to do so ultimately came before the U.S. Supreme Court.[38] In striking down two state bans on the

38 Granholm, 125 S. Ct. 1885 (2005).

interstate direct shipping of wine, the U.S. Supreme Court relied heavily on the FTC's wine report in its analysis to determine whether such discriminatory treatment of out-of-state and in-state interests was necessary to advance valid state concerns, such as reducing underage drinking and collecting taxes. The Court found that, as the FTC staff *Wine Report* concluded, prohibitions on the direct shipment of wine were not necessary to protecting these interests.

Since the U.S. Supreme Court decision, a number of states are reconsidering their laws regulating the direct shipment of wine. Legislators in Ohio and Florida asked the FTC staff for its views on bills that would permit the direct shipment of wine to consumers in those states. In these advocacy comments, FTC staff stated that allowing interstate direct shipping likely would allow consumers to purchase both a greater variety of wines and many wines at lower prices.[39]

C. THE EFFECTIVENESS OF COMPETITION ADVOCACY

Unlike enforcement actions, where the competition agency either succeeds or fails in stopping the anticompetitive conduct based on a court's decision or a settlement with the defendants, the effectiveness of competition advocacy can be more difficult to measure. Occasionally, a state policymaker stops an anticompetitive measure and gives specific credit to a particular advocacy. For example, in vetoing a bill in 2004, Governor Schwarzenegger of the state of California cited the FTC's arguments about the potential unintended effects of the bill as a key reason.[40] Even without such explicit acknowledgement as the U.S. Supreme Court's extensive reliance on the FTC *Wine Report* or Governor Schwarzenegger's nod to the FTC, it seems likely that certain competition advocacy work has affected the decision of policymakers to reject anticompetitive proposals or to tailor them to reduce their anticompetitive impact. One study, published in 1989, attempted a systematic measurement of the FTC's competition advocacy filings at the state and local levels from June 1, 1985, to June 1, 1987.[41] It surveyed recipients of the filings during this time period and asked them questions about the effectiveness of the advocacy filing, whether it provided information or perspectives not presented by other sources or not well understood by the decision maker; and the weight given to the advocacy filing. The study found that a majority of

39 Letter from FTC Staff to Ohio State Senator Eric D. Fingerhut (Mar. 22, 2006), *available at* http://www.ftc.gov/os/2006/03/V060010CommentReOhioSB179DirectShipmentofWine.pdf. *See* Letter from FTC Staff to New York State Rep. William Magee *et al.* (Mar. 29, 2004) ("New York Letter"), *available at* http://www.ftc.gov/be/v040012.pdf. New York ultimately enacted legislation permitting (interstate and intrastate) direct shipping of wine to its consumers. *See* N.Y. ALCO. BEV. CONT. LAW §§ 79-c, 79-d (McKinney 2005).

40 Letter of Governor Arnold Schwarzenegger to Members of the California State Assembly Returning Assembly Bill 1960 Without Signature (Sep. 29, 2004), *available at* http://www.governor.ca.gov/govsite/pdf/vetoes/AB_1960_veto.pdf.

41 Celnicker, *supra* note 17, at 392-93.

recipients who replied to the survey reported that the advocacy filings had some positive effect: forty percent stated that the advocacy filings were at least moderately effective, meaning that "the governmental entity's actions were totally or in large part consistent with all of the FTC's recommendations, and that any action taken was largely or partly because of those recommendations," and an additional eleven percent reported that the comments were slightly effective, meaning that "the governmental entity's actions were to a small degree consistent with at least some of the FTC recommendations, and that any action taken was largely or partly because of those recommendations."[42] As for providing additional perspectives, one state attorney general's office responded that "state or local entities are often totally *unaware* of any antirust problems."[43]

A more recent examination of competition advocacy at the FTC identified a number of factors affecting the success of competition advocacy.[44] On the state level, these factors include situations in which one industry, or subgroup of an industry, seeks regulation that favors it at the expense of a rival industry or group. The article theorized that the most important factor is if the competition advocacy is consistent with organized opposition by an industry group rather than supporting consumers and possible (currently unidentified) new entrants alone. Another factor the article identified is empirical substantiation for the proposition that the regulation will hurt consumer welfare.

The Office of Policy Planning at the FTC is currently conducting a new survey to measure the effectiveness of its advocacy filings between 2001 and 2006, and also to gain a better understanding of the factors that contribute to the success and failure of advocacies. Thus, in addition to the types of questions posed in the 1987 survey, this new survey also asks whether there was substantial local press coverage of the proposed regulation, whether there was press coverage of the advocacy comment, and whether the FTC comment was influential due to the publicity and press coverage attending the FTC's involvement in the matter.

D. PUTTING IT ALL TOGETHER

The FTC's long experience with challenging competitive restrictions that claim the mantle of state approval, combined with the insights from the studies of advocacy, suggest several guidelines for successfully reducing consumer harm in this area:

- Competition officials should examine closely any anticompetitive restriction, particularly those proposed by regulatory bodies dominated by industry members, to determine whether it is actually an action of

42 *Id.* at 391.

43 *Id*. at 396.

44 Cooper et al., *supra* note 16, at 1106-10.

the state or the product of private conduct that occurs in the shadow of state regulation but is not actively sanctioned by the state.

- To the extent immunities protect certain anticompetitive restrictions, examine whether they are being interpreted expansively to shelter conduct unnecessarily. If there is a problem, work to improve the state of the law through scholarly reports, amicus briefs, and testimony before relevant policymakers.[45]

- In industries that seem to lack competition, competition officials should engage in in-depth inquiries to identify the source and mechanism of competitive problems, whether from government regulation, private conduct, or otherwise. Such inquiries may require empirical economic research; workshops with industry members, state officials, and academic researchers; and consultations with industry-specific regulatory agencies.

- Using expertise gained through enforcement and inquiries, competition officials should seek to persuade policymakers evaluating anticompetitive state restrictions to forgo such restrictions or to modify them to reduce the negative impact on competition. For example, policymakers concerned about lack of consumer understanding about new offerings in the market can consider requiring a consumer disclosure instead of prohibiting the sale of the new offerings.

- In all of these endeavors, competition officials should not neglect the importance of informing the debate on competitive issues—through formal and informal actions—both to serve as the voice of diffuse consumer interests and to help assign political responsibility for state actions that harm consumer interests.

IV. Conclusion

Identifying, challenging, and assigning political responsibility for state regulation that restricts competition requires competition officials to exercise many talents, not the least of which is creativity in crafting ways to attack restraints that are immune from the frontal assault of enforcement. However, judicious enforcement, careful legal and economic analysis, in-depth inquiry, well-reasoned scholarship and advocacy, and sheer persistence have produced many successes for the FTC. Other competition officials concerned about the harm from anticompetitive state restrictions may want to use the FTC's multi-pronged approach as a guide in this area.

45 For example, the FTC staff testified before the U.S. Antitrust Modernization Commission about the state action doctrine. *See* Maureen K. Ohlhausen, Prepared Statement on the State Action Doctrine (Sep. 29, 2005), *available at* http://www.ftc.gov/os/2005/09/050929antitrustmod.pdf.

VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

Illinois Tool Works v. Independent Ink: A Lawyer's Take on Ending Special Suspicion of Patent Tying

Richard G. Taranto

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Illinois Tool Works v. Independent Ink: A Lawyer's Take on Ending Special Suspicion of Patent Tying

Richard G. Taranto

The U.S. Supreme Court's decision in *Illinois Tool Works, Inc. v. Independent Ink, Inc.*, holds that a plaintiff, when asserting a tying claim under the familiar modified per se rule requiring market power for liability, must affirmatively prove such power even if the defendant owns a patent covering the tying product. Answering the specific question presented to it, the Court thus took the important step of abandoning an earlier presumption of market power in such circumstances. The contribution of the Court's opinion to antitrust law, however, does not stop there. The Court's opinion provides several additional lessons for those who live and litigate under the U.S. antitrust laws. The lessons are both substantive (about the content of antitrust rules and standards) and methodological (about the proper approach to deciding antitrust cases). As with any individual Supreme Court decision in a body of ever-changing common-law-like doctrine, aspects of the Court's analysis leading to the specific holding supply material, of varying strength and solidity, that lawyers must consider and use when analyzing and litigating antitrust issues.

The author is a partner at Farr & Taranto in Washington, DC.

I. Background to the Court's Opinion[1]

Trident, Inc., eventually bought by Illinois Tool Works, Inc., made and sold components of inkjet printers specialized for printing barcodes on cartons and other packaging materials. Trident owned patents covering two of the components it sold to the manufacturers of the printers—a printhead and an ink container, described in the case as "printhead systems." Trident also sold ink specialized for use in those printers, but it owned no patent covering the ink. The conduct that eventually became the subject of the U.S. Supreme Court's ruling was a tying arrangement. In selling the printhead systems, Trident obtained promises from the printer manufacturers that they would buy their ink only from Trident, and that they and their end-user customers (the firms using the printers to print barcodes) would not refill the ink containers at all.[2]

According to the district court, Trident faced significant competition. At least two other firms sold competing printheads suitable for printing barcodes directly on certain packaging materials. Further rivalry came from several sellers of equipment for printing barcodes on labels that can then be affixed to packaging—a substitute for printing directly on the packaging. In addition, several competitors, including Independent Ink, sold ink, even ink that could be used with the Trident printhead itself.[3]

Independent Ink sued Trident. It alleged, among other things, that Trident's tying arrangement was illegal under Section 1 of the Sherman Act, 15 U.S.C. § 1, for violation of the modified per se rule against tying established in precedents such as *Jefferson Parish*.[4] Both in its own motion for summary judgment of illegality and in resisting Trident's motion for summary judgment of legality, Independent Ink placed all its eggs in one basket, according to the district court. Relying on the Supreme Court decisions in *International Salt* and *Loew's*,[5] Independent Ink contended that Trident's patents covering its printhead systems (the tying product) themselves sufficed to establish the market power required by the tying rule. The district court concluded, however, that those precedents were no longer good law and that patents neither conclusively nor even presumptive-

1 The case at the center of this paper is Illinois Tool Works, Inc. v. Independent Ink, Inc., 126 S. Ct. 1281 (2006) [hereinafter *Illinois Tool Works*].

2 *Id.*; Independent Ink, Inc. v. Trident, Inc., 210 F. Supp. 2d 1155, 1158 (C.D. Cal. 2002) [hereinafter *Independent Ink I*].

3 *Independent Ink I*, *supra* note 2, at 1158-59.

4 Jefferson Parish Hosp. Dist. No. 2 v. Hyde, 466 U.S. 2 (1984).

5 International Salt Co. v. United States, 332 U.S. 392 (1947); United States v. Loew's, Inc., 371 U.S. 38 (1962).

ly proved market power.[6] It then explained that Independent Ink had not submitted any other evidence that could establish either a market definition for the tying product or market power in that market, that Trident (through the evidence of competition) had in fact rebutted any presumption of market power that did exist, and that Trident therefore was entitled to summary judgment of no tying liability.[7]

On appeal, the U.S. Court of Appeals for the Federal Circuit rejected the district court's treatment of the older Supreme Court precedents, which the appellate court read as establishing a presumption of market power over patent-covered products in "patent tying" cases. It reasoned that those precedents, though subject to substantial criticism, had to be respected until the Supreme Court said otherwise.[8] The Federal Circuit set aside the summary judgment that Trident had been awarded, concluding that, despite Trident's evidence of substitutes for its printhead systems, its economic analysis of market conditions was inadequate to overcome the legally required market-power presumption on summary judgment.[9]

THE ISSUE PRESENTED TO AND DECIDED BY THE SUPREME COURT WAS WHETHER THE PRESUMPTION OF MARKET POWER OVER PATENT COVERED TYING PRODUCTS SHOULD BE TREATED AS GOOD LAW. THE COURT HELD THAT IT SHOULD NOT.

II. The Court's Opinion

The issue presented to and decided by the Supreme Court was whether the presumption of market power over patent covered tying products should be treated as good law. The Court held that it should not. Recognizing that it was changing the law, it then gave Independent Ink an opportunity on remand to make a normal economic case of market power and to address "any other issues that are relevant to its remaining § 1 claims."[10]

---

6 *Independent Ink I*, *supra* note 2, at 1163-67 and 1167-70.

7 *Id.* at 1167-73. The district court case originally involved a patent-infringement dispute, but that was settled before decision. The district court ruling was limited to the tying claim and Independent Ink's claim under Section 2 of the Sherman Act, 15 U.S.C. § 2, involving actual or attempted monopolization of, or conspiracy to monopolize, the market for ink (the tied product). The court granted Trident summary judgment against the Section 2 charge on the ground that Independent Ink failed adequately to define or to prove a relevant monopoly power in that market. *Id.* at 1173-77. That ruling was affirmed on appeal, Independent Ink, Inc. v. Illinois Tool Works, Inc., 396 F.3d 1342, 1353 (Fed. Cir. 2005) [hereinafter *Independent Ink II*], and was not the subject of Supreme Court review.

8 *Independent Ink II*, *id.* at 1346-52.

9 *Id.* at 1352-53.

10 *Illinois Tool Works*, *supra* note 1, at 1293.

The Court began with a general survey of the history of tying law, not keyed to patents. Both the U.S. Congress, through the 1914 enactment of Section 3 of the Clayton Act, 15 U.S.C. § 14, and the Supreme Court, through several decisions from 1917 to 1969,[11] expressed "strong disapproval of tying arrangements" as allowing a firm with power in one market to restrain competition in another and as "serv[ing] hardly any purpose beyond the suppression of competition."[12] Then, in 1977 and 1984, the Supreme Court rejected the premise that tying arrangements rarely serve a purpose beyond suppressing competition—a premise that "has not been endorsed in any opinion since" 1977—and clearly insisted on market power over the tying product as a condition of per se invalidity.[13] Only after providing this non-patent-specific background account did the Court address "the validity of the presumption that a patent always gives the patentee significant market power."[14]

The Court explained that the presumption originated in patent law cases, particularly those defining a "patent misuse" defense to infringement claims, and that the Court, in the 1947 *International Salt* decision, eventually "accepted the Government's invitation to import the presumption of market power in a patented product into our antitrust jurisprudence."[15] The Court then set forth the key affirmative reasons for now abandoning the presumption. Most importantly, Congress changed the underlying patent law. After "chipping away at the assumption in the patent misuse context" even in the 1952 Patent Act, Congress acted in 1986 flatly "to eliminate that presumption in the patent misuse context" by limiting patent-misuse claims based on tying with language making clear that "the mere existence of a patent [does not] constitute the requisite 'market power.'"[16] That change in the patent statute undermined the presumption in the antitrust setting, not only because "it would be anomalous to preserve the presumption in antitrust after Congress has eliminated its foundation" (in patent law), but because the Sherman Act is a criminal statute, and "[i]t would be absurd to assume that Congress intended to provide that the use of a patent that merited punishment as a felony would not constitute 'misuse,'" subject to the

11 *See* Motion Picture Patents Co. v. Universal Film Mfg. Co., 243 U.S. 502 (1917); Fortner Enters., Inc. v. United States Steel Corp., 394 U.S. 495 (1969).

12 *Illinois Tool Works*, *supra* note 1, at 1286 (quoting Standard Oil Co. v. United States, 337 U.S. 293, 305-06 (1949)).

13 *Id.* at 1287-88 (discussing United States Steel Corp. v. Fortner Enters., *Inc.*, 429 U.S. 610 (1977), and *Jefferson Parish*, *supra* note 4).

14 *Id.* at 1288. The Court suggested that it read the "presumption" in earlier cases as irrebuttable – as simply rendering per se illegal any requirement by a patentee that its customers for a patent-covered product also buy unpatented goods from it.

15 *Id.* at 1289.

16 *Id.* at 1290 (discussing 35 U.S.C. § 271(d)(5)).

"significantly less severe" remedy of denying patent enforcement.[17] The Court then declared that any conclusion of illegality for tying "must be supported by proof of power in the relevant market rather than by a mere presumption thereof," dropping a footnote to say that this conclusion "accords with the vast majority of academic literature on the subject."[18]

Only after reaching this conclusion did the Court address the only substantial argument put forth by Independent Ink, with the support of one important amicus brief (submitted on behalf of Professor Barry Nalebuff and a few colleagues). That argument, acknowledging that most patents carry no market power, defended a rebuttable presumption of market power for only some patent tying arrangements—those involving a requirements tie (a promise to purchase unpatented goods over time, not just simultaneously with the patented good)—on the ground that such arrangements effect price discrimination and price discrimination is strong evidence of market power.[19] The Court's first answer was that this was an argument for a different rule from the one created in *International Salt*, which placed no reliance on the fact that the tie was a requirements tie and which, in any event, seemed to involve no price discrimination.[20] The Court's second answer made what is essentially the only contested substantive economics point in the opinion: that even price discrimination "occurs in fully competitive markets" and, therefore, does not suffice to support a presumption of market power.[21] The Court drew the lesson: "Many tying arrangements, even those involving patents and requirements ties, are fully consistent with a free, competitive market."[22]

The Court wrapped up its opinion with one final support for its conclusion. The government enforcement agencies, which had led the Court to the presumption in 1947, have at least since 1995 expressly disclaimed any presumption of market power based on patents (or copyrights or trade secrets) in exercising their prosecutorial discretion. The Court explained that in antitrust law the enforcement agencies' positions do not bind the courts (they're not delegated substantive standard-setting authority), but the agencies' position was nevertheless significant—because the Sherman Act is a criminal statute "it would be

17 *Id.* at 1291.

18 *Id.* at 1291 & n.4. Earlier, the Court advised that its review was "informed by extensive scholarly comment and a change in position by the administrative agencies charged with enforcement of the antitrust laws." *Id.* at 1285.

19 *Id.* at 1291-92.

20 *Id.* at 1292.

21 *Id.*

22 *Id.*

unusual for the Judiciary to replace the normal rule of lenity that is applied in criminal cases with a rule of severity for a special category of antitrust cases."[23]

In sum, because "Congress, the antitrust enforcement agencies, and most economists have all reached the conclusion that a patent does not necessarily confer market power on the patentee," the Court held "that, in all cases involving a tying arrangement, the plaintiff must prove that the defendant has market power in the tying product."[24]

III. Contributions of the Opinion to U.S. Antitrust Law

The Supreme Court's opinion in *Illinois Tool Works*, consistent with the usual constraints on judicial decision-making, confines itself to discussing the specific question presented as the case reached the Court. Accordingly, the opinion does not address more general issues presented by tying law, such as whether the modified per se rule for tying liability should be abandoned, what market effects must be proved to establish liability under that rule, and what market benefits may be considered in applying the rule. Nevertheless, in the course of analyzing the specific issue of a patent's bearing on the market-power requirement, the Court's opinion makes a number of contributions to antitrust analysis that are of more general significance to the direct audience for antitrust decisions—lower courts, private and government lawyers, businesses, and consumers.

The decision presumably also provides grist for the economist's mill, as commentaries other than this one will show. But economists' analyses are one thing and lawyers' (and hence judges') analyses another. The latter unavoidably must filter the former through a series of lenses embodying limits on the capacities and perspectives of legal institutions and the players in them, including the judges and juries that decide cases and the lawyers who advise clients and prepare cases. Indeed, although legal analysis in antitrust does and must depend on sound economic analysis—an antitrust law approach that is either economically senseless or not supported by coherent economic understanding should not survive—the importance of institutional considerations in defining just how economics is used in antitrust law is one of the lessons that at least one lawyer finds suggested in *Illinois Tool Works*.

23 *Id.* at 1292-93.

24 *Id.* at 1293.

A. SUBSTANTIVE

1. Market Power Required

The Court stated clearly, more than once, that market power is a requirement for an antitrust challenge to a tying arrangement. And its usage was unitary: what is required is the market power demanded by the modified per se rule. Those statements undermine any notion that a tying arrangement might be condemned under a rule of reason approach even without market power or with proof of some lesser degree of market power.[25]

2. Tying Not Inherently Suspicious

The Court's opinion pervasively reaffirms that the era of suspicion of tying arrangements is over. That discussion is presented as the backdrop for the Court's analysis of the specific issue presented in the case—the survival of any presumption of market power based on patent rights. The backdrop discussion is significant for what it says in terms: it restates with clarity and gives emphasis to the Court's repudiation of an earlier suspicion of tying arrangements. The discussion is especially significant, moreover, because the Court chose to include it in an opinion that in no way required it. The sole issue presented was whether patents support a presumption of market power. Resolving that question did not require commentary on whether tying arrangements, when they exist, are especially likely to meet proper standards for antitrust condemnation. The Court's choice to reaffirm that they are not, a merely relevant though unnecessary framework for its decision of the question presented, amplifies the message of that discussion.[26]

3. What Market Power Means

The Court implicitly, but necessarily, adopted a robust understanding of the market power that must be proved for an antitrust challenge to a tying arrangement. That concept, as used by the Court, must mean more than a short-term absence of alternatives for purchasers or even the seller's ability profitably to set its own price above cost; it must mean something more than a departure from the classic model of pure competition under which individual sellers face horizontal demand curves. Although the Court does not define the required "market power" (i.e., durability?; degree?; focus on profits, not prices?), the robustness of the concept is implied by the citations that the Court deploys to support virtually the sole economic point in the opinion.

25 *See, e.g.*, Town Sound & Custom Tops, Inc. v. Chrysler Motors Corp., 959 F.2d 468 (3d Cir. 1992) (*en banc*).

26 The Court's opinion does not cite Eastman Kodak Co. v. Image Technical Servs., Inc., 504 U.S. 451 (1992), though that decision discusses the law of tying after *Jefferson Parish* and the *Illinois Tool Works* opinion is otherwise fairly comprehensive in mentioning tying cases. Sometimes such a notable disregard of an earlier precedent reflects skepticism about its soundness, in whole or in part, at least by some on the Court.

The Court's lead citation is to an article by Baumol and Swanson that makes and defends the points that suppliers who are subject to meaningful competitive discipline engage in price discrimination.[27] The page cited by the Court, however, expressly accepts that such a firm must face a negatively sloping demand curve, as a firm in a model of perfect competition would not. Even so, discussing a familiar model of monopolistic competition, the cited page and surrounding pages observe that many such firms are subject to competitive discipline, so that price discrimination is not a reliable indicator of "market or monopoly power in any sense relevant to antitrust policy," even though the price-discriminating firm is in "violation of several of the commonly accepted indicia of market power."[28] The authors specifically favor a focus on levels of profits over time.

The Court's second citation is to a section of the leading legal treatise on American antitrust law, by Areeda and Hovenkamp.[29] The citation is a little curious, because the concern of the section is not to deny that price discrimination is an indicator of market power; rather, the section explains why a tying arrangement that cannot properly be condemned on other grounds should not be condemned because it is used for price discrimination—among other reasons, price discrimination often moves output to levels that a competitive market would produce. In citing this section to make its point about market power, the Court seems to be referring to a brief passage near the start of the treatise section. That passage affirmatively states that "some degree of market power" is required for price discrimination, but it immediately adds that the required degree need not be great—anything other than "intense competition."[30]

The Court's final citation is to an excerpt from a book by Landes and Posner.[31] The discussion cited stresses that a patent holder's price discrimination requires that it face a downward-sloping demand curve, unlike the horizontal curve faced by a firm in a model of perfect competition.[32] The surrounding material criticizes certain arguments against characterizing the result in market power terms,[33] but

27 William J. Baumol & Daniel G. Swanson, *The New Economy and Ubiquitous Competitive Price Discrimination: Identifying Defensible Criteria of Market Power*, 70 Antitrust L.J. 661 (2003). The Supreme Court decision cites page 666.

28 *Id.* at 666.

29 Phillip E. Areeda & Herbert Hovenkamp, 9 Antitrust Law ¶ 1711, at 100-15 (2d ed. 2004). The Supreme Court decision cites this section in toto.

30 *Id.* at 101-02 & n.4.

31 William M. Landes & Richard A. Posner, The Economic Structure of Intellectual Property Law (2003). The Supreme Court decision cites pages 374-75.

32 *Id.* at 375; *see also, id.* at 377.

33 *Id.* at 375-77.

insists that the "fundamental" point is that "market power is pervasive" and whatever market power a price discriminator has should not be a source of legal "worry": it is "too little . . . in a meaningful economic sense to interest a rational antitrust enforcer."[34] The Supreme Court, in citing this discussion, implicitly is adopting a robust legal standard for market power.

B. METHODOLOGICAL

1. Statutory Policy and Antitrust Precedent

The Court's analysis of the market-power presumption issue relies overwhelmingly, though not quite exclusively, on an essentially non-economic analysis. It is based on legal considerations that would find a place in almost any statutory analysis. The Court traces the origin of the market-power presumption to a borrowing for antitrust law of a principle adopted early on in patent law and then relies centrally on the fact that Congress has repudiated the principle in the patent setting where it originated. The Court thus treats the change of statutory policy as critical for the resolution of the question in the context of antitrust law.

That approach reflects a traditional legal analysis in two respects. The first involves a delicate, essentially one-way principle, based on the strong common-law flavor of antitrust law, which comes from statutes that leave so much of the content of decisional rules to elaboration and evolution by the courts.[35] One guide for how courts can fill out the content of common law—or indeed of any open-ended standard—is statutory policy laid down by the legislature in closely related areas.[36] That guide is very limited: statutory duties not serving competition cannot be imported into antitrust law, and even statutory duties that might be characterized as serving competition in some sense (the senses are multifarious) generally should not be imported into the antitrust regime of criminal and treble-damages enforcement through non-expert decision makers.[37] Nevertheless, the "look to Congress" guide does support judicial repudiation of a special antitrust-law presumption when Congress has rejected it in substance by direct statutory enactment. The second, less delicate doctrinal grounding for *Illinois Tool Works*'s use of congressional modification of the patent statute is the law of stare decisis. The established standards for when a precedent should be overruled include, as one of their most solid elements, the consideration whether

34 *Id.* at 377, 374-75.

35 *See* Business Electronics Corp. v. Sharp Electronics Corp., 485 U.S. 717, 731-33 (1988). The Court adverts to this principle in *Illinois Tool Works*, *supra* note 1, at 1291 n.3, in citing State Oil Co. v. Khan, 522 U.S. 3, 20 (1997).

36 *See* Moragne v. States Marine Lines, 398 U.S. 375, 390-91 (1970).

37 The Supreme Court made closely related points in Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, 540 U.S. 398, 410-15 (2004).

the underpinnings of the earlier ruling have been eroded.[38] That is what the Court found to have happened when Congress repudiated a presumption of market power in the patent-law setting.

2. Enforcement Agency Position and the Rule of Lenity

Two additional aspects of the Court's legal analysis leading to its repudiation of the presumption of market power based on patent rights are noteworthy.

- First: The Court's opinion twice invokes the fact that the Sherman Act is a criminal statute, the second time noting that the rule of lenity—directing courts to resolve genuine ambiguities and uncertainties against breadth of prohibition—applies to criminal statutes. The invocation of the rule of lenity, even in a private civil case, is a striking aspect of the Court's analysis. It is of potentially quite general importance in the task of construing the Sherman Act, suggesting that doubts go against antitrust intervention in market activities.

- Second: The Court's opinion places express reliance on the position of the government enforcement agencies. That is not because the agencies have the kind of statute-implementing authority that would entitle them to the formal legal deference that other agencies receive for their substantive views about what statutes mean or how they should be applied. Rather, it is for a different reason, which works only in one direction, namely, leniency. The Sherman Act is a criminal statute, and it would be anomalous for the courts to apply a restrictive standard that the prosecutors have disclaimed. That principle establishes a legal basis for one kind of deference that supplements the simple fact of life that agencies with expertise and experience, though also with institutional biases, will get special attention from the Court when they speak, as they often do, with rigor and clarity.

3. Substantive Economics and Burdens of Proof

As the foregoing indicates, almost all of the Court's affirmative reasoning in support of overruling the market-power presumption is a traditional legal analysis. The Court's opinion says very little about substantive economics. It makes its economic point, that price discrimination does not strongly enough imply market power, only briefly, and only within an analytical structure that has already essentially shifted the burden of affirmative justification to the proponents of anything but general applicability of the requirement that plaintiffs must prove market power in tying cases.

The brevity of the economic discussion, its reliance on very simple points backed by a perceived broad (if not uniform) scholarly consensus, and its placement in an analytic structure so as not to make it bear too much of the weight

38 *See* United States v. Gaudin, 515 U.S. 506, 521 (1995); Rodriguez de Quijas v. Shearson/American Express, Inc., 490 U.S. 477, 480-481 (1989).

for the Court's conclusion are telling. The Court's opinion does not finely parse the terms of the debate—a parsing that sometimes is essential to a proper resolution of an antitrust case, but that can easily strain the institutional capacity of courts, and that was ultimately unnecessary in *Illinois Tool Works*. Here, as in so many areas of the law, the Court relied critically on general considerations to establish a burden of proof that effectively determined the outcome. From the general recognition that tying is very often not an economically harmful practice, and the congressional and executive repudiation of a presumption of market power based on patents, the Court decided to insist across the board on affirmative proof, without aid of shortcuts, of the market power that is one important precondition for finding a particular tie deserving of legal condemnation. This approach reflects the centrality of limits on institutional capacity in deciding what to do, in law, with potentially complex debates among economists.

4. A General Approach?

In an area of law predominantly shaped by judicial opinions, individual decisions about particular issues combine to form patterns that can embody general principles or attitudes to guide legal analysis more broadly. Lawyers and judges reading such decisions can and do infer, even if they have difficulty precisely articulating, decision-shaping lessons for approaching new problems in the area. Such general dispositions, often conveying a message of skepticism or receptivity, play a persistent and powerful role in decision-making and in the evolution of doctrine that strives for overall coherence. Commentary on a particular decision is always at its most tentative when it moves from the particulars of the case to hazard a description of such a more general theme, but *Illinois Tool Works* lends itself to the attempt.

The Court's invocation of the rule of lenity, its decision to give so little emphasis to the substantive economic debate, and its re-assertion of the need for affirmative proof by the plaintiff can collectively be understood to support a more general disposition toward judicial antitrust analysis. At the heart of this disposition is a tenet of institutional epistemology, captured in the Johnny Mercer title, "How Little We Know." Such a disposition—given great prominence in the Supreme Court's 1993 *Brooke Group* decision on predatory pricing and its 2005 *Verizon v. Trinko* decision on unilateral refusals to deal[39]—treats the limited capacity of legal institutions as a primary element, rather than an afterthought, in deciding what antitrust rules to adopt.

The epistemological modesty about how well the lay decision-makers in the legal system can make reliably accurate determinations, and how well lawyers acting as advisers can predict such determinations, has at least two aspects. One is skepticism about how well such decision-makers and advisers can understand and make objectively grounded assessments of economic behavior and its likely effects

39 Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, 223 (1993); *Trinko*, *supra* note 37, at 414-15.

in an individual case—and whether they can do so at a cost that does not swamp the economic benefits of the substantive determination. The other, which is at least as important, is skepticism about how well—and, again, how cheaply—they can predict and control the systemic conduct-altering consequences of an antitrust rule or principle, or even of a seemingly narrow ruling on legality in a particular case (seemingly case-specific rulings in a common-law system shape future rulings and hence private decisions made in light of predictions about such rulings).

These two pillars of doubt about the reliability of legal antitrust decision-making support an ethical prescription, at least when coupled with a comparative assessment of markets, which rely on decentralized guesswork and experiment, as often better suited to identifying, responding to, and adjusting responses to competitively problematic conduct. The prescription would be an analytical starting point, a default position, that is a kind of Hippocratic oath for courts asked to intervene in private market activity: first do no harm. In practical doctrinal terms, a plaintiff's demand for antitrust intervention would be rejected unless and until the plaintiff articulates objective and generalized standards for antitrust intervention, defends those standards based both on how accurately and cheaply they can be applied in individual cases and (as important) how well they avoid systemically deterring desirable private market conduct, and proves that the standards apply to the facts of the case. For some common situations, such as price agreements among business rivals and some kinds of horizontal mergers, the default position is readily overcome by familiar and soundly based doctrine. For other situations, especially unilateral conduct, existing doctrine supplies no such general basis for overcoming a default rule of non-intervention.

The Supreme Court decision in *Illinois Tool Works*, which reads as a modest rather than ambitious pronouncement, does not venture to articulate such broad themes. But in its very cautiousness, and in its choices of premises for justifying its result, both substantive and methodological, the decision can be seen as part of a pattern that suggests a general modesty about judicial intervention under the antitrust laws. Such patterns, when they emerge from individual decisions, unavoidably exert great influence in an area of law where the decision-makers so often find the case-specific empirical analyses of particular problems indeterminate; where the standard jury instructions mix the meaningless, the confusing, and the misleading; and where even some of the common terms of judicial analysis (i.e., "anticompetitive," "exclusionary," "predatory") are so far from having uniform, intuitive, and operationally practical meanings that they provide more in the nature of bottom-line conclusions than thought-clarifying guides to a mind trying to assess why conduct should or should not be condemned.

In its very cautiousness, and in its choices of premises for justifying its result, both substantive and methodological, the decision can be seen as part of a pattern that suggests a general modesty about judicial intervention under the antitrust laws.

Volume 2 | Number 2 | Autumn 2006

Competition Policy International

Review of Hovenkamp, *The Antitrust Enterprise: Principle and Execution*

Randal C. Picker

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2. *Competition Policy International* is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Review of Hovenkamp, *The Antitrust Enterprise: Principle and Execution*

Randal C. Picker

Herbert Hovenkamp, the Ben V. & Dorothy Willie Professor of Law and History at The University of Iowa College of Law, is best known to the antitrust bar for his role as the senior surviving author of the multi-volume *Antitrust Law* treatise originated by Philip Areeda and Donald Turner. The treatise is the standard reference in antitrust, and the common-law nature of antitrust in the United States makes the treatise particularly influential. Hovenkamp has also written more broadly, and my personal favorite has always been his 1991 business history, *Enterprise and American Law 1836-1937*. Now Hovenkamp has written a new single-volume overview of U.S. antitrust law titled *The Antitrust Enterprise: Principle and Execution*.[1]

Of course, the gold standard for this genre is Robert Bork's *The Antitrust Paradox* and I push my students towards Richard Posner's *Antitrust Law*, a second edition of which was issued in 2001.[2] Like those books when they were published, it is easy to say that any serious antitrust participant should buy and read *The Antitrust Enterprise*. The book is a highly readable, integrated perspective on the state of antitrust law in the United States, written by someone who has both a

1 Herbert Hovenkamp, The Antitrust Enterprise: Principle and Execution (2006).

2 Robert Bork, The Antitrust Paradox (1993); Richard A. Posner, Antitrust Law (2nd ed. 2001).

The author is Leffmann Professor of Commercial Law at the University of Chicago Law School.

historian's sense of time and change and regulatory cycling, and an up-to-date knowledge of current doctrinal twists. You should put it on your bookshelf—and on one of the low shelves that you can reach easily while sitting at your desk.

I should map the intellectual space of *The Antitrust Enterprise*. Bork's *The Antitrust Paradox* helped to define the Chicago School and helped push the courts towards a dramatic shift in how they approached antitrust cases. The colonization of antitrust by economics still continues as the U.S. Supreme Court slowly prunes away decades worth of per se rules. The rise of post-Chicago analysis has expanded the possibilities, and how the courts should deal with that richness is one of the early themes of the book. We might think of *The Antitrust Enterprise* as *The Antitrust Paradox* for a post-Chicago antitrust landscape. One of the key features of the new domain is how the court system should manage the complexity of antitrust doctrine itself.

By the standards of modern statutes, the Sherman Act is remarkably brief, but Hovenkamp suggests that we could usefully shorten the statute, sufficiently so that we could put the entire statute on a bumper sticker: "unreasonable restraints on competition are hereby forbidden." This is reminiscent of a 1908 *New York Times* editorial on potential antitrust legislation to address the perceived defects of the Supreme Court's 1897 decision in *Trans-Missouri*.[3] The *Times* wanted a one-word amendment to the Sherman Act ("[t]he insertion of the word 'unreasonable' before the word 'restraint' would take the mischief out of the act, and sufficiently amend it") rather than the power grab being pushed by President Theodore Roosevelt in the form of the 1700-word Hepburn Bill.

That we have not moved the ball forward in a century's worth of work reflects either the genius of the Sherman Act (at least as reinterpreted by the Supreme Court in *Standard Oil*[4] when it abandoned *Trans-Missouri*) or the limits of drafting. Hovenkamp recognizes that law professors lack the power to rewrite statutes—so he spends the rest of the book working with the statutes that we have. The book is divided into twelve chapters organized into three groups: limits and possibilities; traditional antitrust rules; and regulation, innovation, and connectivity.

In the first third of the book, Hovenkamp offers a framework for managing post-Chicago economic analysis. Hovenkamp is acutely aware of the powerful limitations that circumscribe our abilities to implement competition policy regulation. Those limits arise out of the difficulty of identifying antitrust violations; then of proving them ("it has become something of a commonplace that rule of reason antitrust violations are almost impossible to prove, particularly in private

3 United States v. Trans-Missouri Freight Ass'n, 166 U.S. 290 (1897).

4 Standard Oil Co. v. United States, 221 U.S. 1 (1911).

plaintiff actions"); and then coming up with a tractable remedy to improve matters. Antitrust remedies are beset by the central problem faced by a dog chasing a car: what does the dog do with it when it catches it?

Hovenkamp notes the challenges faced by generalist federal judges, who may not be well-versed in economics and the industry in question and its technologies (or even in antitrust for that matter). The frequent solution to those hurdles is for the judge, as Hovenkamp notes, to delegate decision-making to the only group in the courtroom with less knowledge of these issues, the jury that has been dragooned into service for $18 a day. Hovenkamp also notes that the theoretical development in antitrust analysis has put further burdens on antitrust decision-making. Post-Chicago analysis often consists of possibility results, that is, a demonstration of harmful behavior that might—and you cannot overemphasize might—emerge under certain highly-stylized conditions and under particular parameter settings. Operationalizing these insights in a way that we can be confident that we are fixing more mistakes than we are making may be beyond our capabilities.

With that in mind, Hovenkamp offers a series of administrative suggestions—a five-step program for antitrust—designed to inject a level of caution for antitrust decision makers:

1. Not every anticompetitive practice can be condemned;
2. Intent evidence should be used sparingly;
3. Whether intervention is justified may defend depend on the remedy;
4. An antitrust rule that cannot be administered effectively is worse than no rule at all; and
5. Administrative and compliance costs count.

All of this turns Hovenkamp into something of a Chicago School apologist: "As a result the rather tolerant Chicago School rule may be the best one for policy purposes even though substantial anti-competitive behavior goes undisciplined, simply because we cannot recognize and remedy it with sufficient confidence."[5]

The second section of the book addresses traditional antitrust rules. Chapter 5 starts with a discussion of market definition and market power and offers an extended criticism of the Supreme Court's 1992 decision in *Kodak*,[6] which held that a single-brand after-market could count as a separate market. The book then offers a six-step recipe for applying the rule of reason amidst the quagmire creat-

5 HOVENKAMP, *supra* n. 1, at 48.

6 Eastman Kodak v. Image Technical Servs., 504 U.S. 451 (1992).

ed by the Court's decision in *California Dental Association*.[7] That decision rejected categorical lines and held, in Goldilocks fashion, that the extent of antitrust inquiry should be that required under the circumstances. Not too hot, not too cold, but somewhere in the middle. Chapter 6 considers the difficulties of identifying collusion under Section 1 of the Sherman Act given the strong economic incentives towards parallel behavior and Section 1's agreement requirement.

Chapter 7 turns to Section 2 of the Sherman Act and monopolization, exclusion, and foreclosure. Hovenkamp's ongoing concern with the likelihood of mistakes in decision-making appears again in his discussion of predatory pricing. Predatory pricing is the Loch Ness monster of antitrust: there are frequent sightings, but, on further investigation, it is not clear that we've ever found the beast itself. Given that, we should fear that the threat of a predatory pricing action will deter critical competitive behavior. Hovenkamp concludes that "a high error rate gives reason to believe that predatory pricing law does more harm than good."[8]

Chapter 8 focuses on antitrust and distribution, including discussions of vertical restraints and the Robinson-Patman Act. Hovenkamp joins the long line of law professors calling for the repeal of that act, but notes that we have been doing so for more than fifty years without any risk of success. Hovenkamp concludes the chapter with a useful discussion of exclusive dealing especially as applied to franchise agreements. Finally, to conclude the second section, Chapter 9 provides a very serviceable walk-through of current merger doctrine, including the U.S. Federal Trade Commission (FTC) and U.S. Department of Justice (DOJ) merger guidelines and leading cases such as *Heinz*,[9] though the book was written before the decision in *Oracle*[10] or before the publication of the March 2006 DOJ and FTC *Commentary on the Horizontal Merger Guidelines*.[11]

The last third of the book addresses regulation, innovation, and connectivity (think, to oversimplify, *Trinko*,[12] patents and copyrights, and *Microsoft*).[13] Hovenkamp offers a nice discussion of the interaction between regulation and antitrust, and of the variety of ways in which antitrust steps back, including fed-

7 California Dental Assn. v. FTC (97-1625), 526 U.S. 756 (1999).

8 Hovenkamp, *supra* n. 1, at 160.

9 Federal Trade Commission v. H.J. Heinz Co., 246 F.3d 708 (D.C. Cir. 2001).

10 United States v. Oracle Corp., 331 F. Supp. 2d 1098 (N.D. Cal. 2004).

11 U.S. Federal Trade Commission and U.S. Department of Justice, Commentary on the Horizontal Merger Guidelines (Mar. 2006), *available at* http://www.usdoj.gov/atr/public/guidelines/215247.htm.

12 Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, 540 U.S. 398, 410-15 (2004).

13 United States v. Microsoft Corp., 253 F.3d 34 (D.C. Cir. 2001).

eral regulatory immunity; state action immunity under *Parker v. Brown*;[14] and *Noerr-Pennington*[15] immunity for efforts to influence the government. Hovenkamp notes how the role for antitrust has expanded as we have deregulated chunks of the economy, such as trucking, railroads, and aviation. Again Hovenkamp emphasizes limits, as he suggests that "*Trinko* may effectively have brought the era of antitrust essential facility claims to an end, certainly in regulated industries where an agency is actively supervising the conduct that forms the basis of an antitrust claim."[16] Hovenkamp regards that as a decidedly good thing, as "an important step in our recognition that competition is not regulation, and federal courts are not regulatory agencies."[17]

Hovenkamp turns next to the perceived conflict between antitrust and intellectual property rights. He takes a detour into copyright proper to criticize the Sony Bono Copyright Term Extension Act, which extended copyright duration retroactively an additional twenty years. In doing so, he illustrates—inadvertently I suspect—one of his points, namely that we fail to draw clean lines separating intellectual property and antitrust and that there is too much of a temptation to try to use antitrust to clean up IP doctrine seen by some to have run amuck. Hovenkamp is on firmer footing when he returns to considering the Supreme Court's patent licensing cases and the blanket licenses at stake in *Broadcast Music*.[18] Unfortunately, the book was written before the recent interest in patent settlements relating to generic drug entry under the abbreviated new drug application process set forth in the Hatch-Waxman Act.

The book then considers network industries and the *Microsoft* case. This is likely to be the most controversial chapter in the book. Hovenkamp starts with the definition of network as "a market subject to economies of scale in consumption."[19] That seems wrong; that is more a definition of a network externality that of a network itself. The phone system is a network with network externalities (the one-user phone system isn't very valuable), while the electricity grid should count as a network, but I don't benefit directly when you start consuming electricity. Hovenkamp then turns to discussing *Microsoft* case. He suggests that the case should be regarded as one of the "great debacles in the history of public antitrust enforcement" as he believes the consent decree that resolved the case

14 Parker v. Brown, 317 U.S. 341 (1943).

15 Eastern Railroad Presidents Conference v. Noerr Motor Freight, Inc., 365 U.S. 127 (1961).

16 Hovenkamp, *supra* n. 1, at 248.

17 *Id*.

18 Broadcast Music, Inc. v. Columbia Broadcasting System, Inc., 441 U.S. 1 (1979).

19 Hovenkamp, *supra* n. 1, at 277.

is likely to accomplish very little.[20] He argues in favor of allowing judges to retain jurisdiction over cases, not merely to assess whether consent decree terms have been met, but seemingly with a broader, overarching power to attempt to cleanse the market of the consequences of the anticompetitive conduct. That suggestion seems to fly in the face of Hovenkamp's prior concerns about the limits of judicial competence and what I suspect is the settled wisdom of the ongoing regulation of the telecommunications industry in the *AT&T*[21] case by Judge Harold Greene. Hovenkamp goes on to suggest that the government itself could inject more competition into the operating system market through its purchases and in particular could switch to open source software to put pressure on Microsoft.

More importantly, Hovenkamp's book is pre-Google. Google and what it represents—a shift away from the computer desktop in favor of the network, and a move from products to services—have emerged as the core source of competition to Microsoft's position. Indeed, Ray Ozzie, who replaced Bill Gates as Microsoft's Chief Software Architect in June 2006, quickly declared the personal computer era over and saw the future in the shift towards services of the sort provided by Google. But the open questions are causal: did the antitrust consent decree play any role in creating space for Google to operate? Would Google have emerged anyhow? Or would Microsoft have Netscaped Google and rendered still-born the next serious threat to its desktop monopoly?

We should turn to what is missing from the book. Formal economics is largely absent from the book. One equation (the Lerner index, which measures market power), no calculus and maximization, and indeed relatively few extended numerical examples. There are a few graphs: a traditional early discussion of the social welfare harms resulting from a monopolist's reduced output and the Williamson welfare trade-off graph which highlights—usually in merger analysis—the social benefit of more efficient production and the potential harm from reduced output through greater market power. The book will not bring you up to date on recent developments in antitrust economics, including the large literature on the competitive consequences of bundling or the difficult issues associated with two-sided markets. The book is also a U.S. book, so there is very little discussion of international antitrust or of the antitrust rules of other jurisdictions (most notably nothing on the European Community).

These are not quibbles, but it would be a mistake to make too much about what are ultimately choices about the scope and size of the book. Hovenkamp's *The Antitrust Enterprise: Principle and Execution* will have a long life and a hefty market position. Readers would do well to make their contribution to that position by buying and reading the book.

20 *Id.* at 298.

21 United States v. American Telephone & Telegraph Co., 552 F.Supp. 131 (D.D.C. 1982), aff'd sub nom Maryland v. United States, 460 U.S. 1001 (1983).

VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

Review of O'Donoghue and Padilla, *The Law and Economics of Article 82 EC*

Richard Whish

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Review of O'Donoghue and Padilla, *The Law and Economics of Article 82 EC*

Richard Whish

The extent to which EC competition law has been modernized in the last decade is really quite breathtaking. In the first half of the 1990s, the realization that there was something seriously wrong with the way in which the rules were applied in practice began to become widely recognized, including within the European Commission itself. An obvious problem was that the rules were applied with insufficient attention both to economic principles and to quantitative techniques. A paper at the 1996 Fordham Conference by David Deacon, an economist at the Commission's Directorate General IV (as it then was), was a major event, when he effectively denounced the entire approach towards vertical agreements that historically had been taken by the Commission, based on formalism and excessive concentration on the wording of clauses in agreements, rather than the economic impact of those agreements. There followed the reform of the vertical restraints regime, which involved a major repositioning of the law and economics of the subject and which appears to have worked well in practice. Numerous policy initiatives followed—new block exemptions for research and development and specialization agreements, guidelines on horizontal cooperation, a new regime for technology transfer, the recast merger regulation, and the horizontal merger guidelines. Most radical of all, perhaps, was the Modernization Regulation, a product of the Commission being prepared to think the unthink-

The author is Professor of Law at King's College London.

able: to dismantle the notification system established in 1962 and in its place to create a Community-wide system of cooperation and power-sharing.

This brief history of the last decade of change leads to the obvious question: what about Article 82 of the EC Treaty, which prohibits the abuse of a dominant position? The most difficult question in competition law is probably to determine when the individual behavior of a firm with significant market power crosses the line from being legal to being illegal, with all the consequences that that entails (for example, the possibility of lengthy and intrusive investigation, fines, remedies, and damages actions). As the authors of this book frequently state, pro- and anticompetitive behavior often look the same: the price cut that may or may not be predatory, the combination of products that may or may not amount to an illegal tie-in. Depending on where the line is placed, the law may condemn perfectly competitive behavior and so chill competition; but unduly lenient treatment could lead to the exclusion of efficient competitors and, in some cases, to the market tipping in favor of the dominant firm, perhaps with long-term adverse consequences for consumer welfare. Whatever system of rules is adopted needs to take into account this problem of false negatives and false positives, or type I errors and type II errors.

Reforming Article 82 is no easy matter. The reforms of the law of agreements and mergers were all based on legislative changes—for example, the adoption of Regulation 2790/99 on vertical agreements and the recast merger regulation, Regulation 139/2004. These were accompanied by helpful guidelines, but it was the legislation in question that changed the substance of the law. The same is true, of course, of the Modernization Regulation. Article 82 cannot be changed by legislation—unless, of course, the Treaty itself were to be amended, which the events of recent years in France and the Netherlands suggest is all but impossible. It follows that reform of Article 82 is a rather different exercise, in which the existing case law of the Community Courts and the administrative behavior of the Commission need to be analyzed, with a view to a possible change in the Commission's prosecutorial practice and in the jurisprudence of the Courts. The Commission decided in 2004 to conduct such a review, and a huge amount of time and intellectual effort has been expended in trying to make sense of the existing law, its origins, purpose, deficiencies, and strengths. Draft guidelines may become available in 2006, although it is more likely that nothing will issue until 2007.

Many (though not all) commentators share the view that there are numerous problems with Article 82. Market definition—and more specifically the hypothetical monopolist (or the Small but Significant and Non-transitory Increase in Price (SSNIP)) test—is more appropriate for the prospective analysis of markets in merger cases than for the predominantly retrospective analysis of alleged infringements of Article 82: this is where the so-called "cellophane fallacy" is relevant. The threshold of dominance is not easy to determine, and perhaps there has been too much emphasis on market share figures (and a presumption of dominance at levels that are too low) rather than on an overall assessment of market power, including barriers to entry, buyer power, and other factors indicating dominance. The question of whether there is a concept of super-dominance over and above mere dominance is also controversial, as has been the meaning of collective dominance. And then there is the really complex question: when is the behavior of a dominant firm abusive? There are numerous associated questions: for example, are there some types of conduct that are so obviously abusive that they should be illegal per se? Cartels are held to have as their object the restriction of competition under Article 81(1), so perhaps sales at less than cost should be illegal per se under Article 82 when effected by a dominant firm. However it is important to note that even an agreement by object could be saved by an efficiency argument under Article 81(3), provided that there is convincing evidence to support the claim;[1] and that there may be good reasons—that is to say reasons other than a strategic attempt to eliminate a rival—to explain even below-cost sales (start-up prices, end-of-season sales, etc.). However, if per se rules are inappropriate in Article 82 cases, it is necessary to determine what type of rule-of-reason analysis should be undertaken. If every case were to require the demonstration of adverse economic effects, to a high standard of proof, the enforcement of Article 82 might become all but impossible, which would bring one back to the problem of false negatives and false positives. In the post-Chicago era, there is widespread recognition that there is such a thing as abusive behavior, and that, therefore, Article 82 and its various progeny—in Europe and beyond—do have an important role to play in the maintenance of competitive markets. The challenge is to devise administrable rules that capture those types of conduct that may be harmful to welfare; to avoid rules that might inhibit perfectly reasonable types of competitive behavior, including, of course, most price cuts; to do so within a framework which gives a reasonable degree of certainty to companies, professional advisers, competition authorities, and to courts; and, furthermore, to do so in a way that blends together the respective roles of both economics and law. Fortunately, we have evolved to the point where there is now a widespread recognition that both economics and law—and therefore both economists and lawyers—have a part to play in designing a workable system of EC competition law. Legal rules about competition that fail to reflect sound economic principles are likely to be harmful to welfare, but economic theories without predictability or due process may also be. So too are turf wars between these two interest groups, who should be cooperating in order to achieve correct outcomes for clients and/or competition authorities.

This lengthy preamble brings us to Robert O'Donoghue and Jorge Padilla's *The Law and Economics of Article 82 EC*.[2] This book is an admirable achievement, and the authors are to be congratulated on producing a work of high class and great interest. It is handsomely produced, easy to read, and comprehensive in its scope.

1 *See e.g.*, Case T-17/93, Matra Hachette v. Commission, 1994 E.C.R. 595 (CFI).

2 Robert O'Donoghue & A. Jorge Padilla, The Law and Economics of Article 82 EC (2006).

Its publication now is timely, as we await the next stage of the Commission's review of Article 82. The authors anticipate, in their preface, a second edition "sooner rather than later," and this would be most welcome. Not only will the Commission probably provide guidance on the application of Article 82 in the fairly near future, but there are also a number of very important appeals in the pipeline before the Community Courts, including *Microsoft*, *Deutsche Telekom*, *Wanadoo*, *AstraZeneca*, and *British Airways*, in each of which very significant issues regarding the scope of Article 82 are under consideration. These cases provide the Courts with a great opportunity to develop the existing jurisprudence of Article 82. It is a well-known problem that the existing case law is fairly sparse, and that a lot of it is fairly old and somewhat formalistic; furthermore much of the past decisional practice of the Commission was concerned with the old economy, whereas most of the cases just mentioned are about innovation or information technology markets, where different economic analysis may be called for. A great deal of improvement could be achieved if the Community Courts were able, in these cases, to refine some of the earlier jurisprudence in the light of the debate that has taken place as a result of the Commission's review over the last couple of years. Much of that debate has been of a very high standard, with the interested parties—representing a number of different interest groups—genuinely trying to find workable solutions to the complex conundrums raised by Article 82. This book makes an enormous contribution to that debate. However, how the Courts will react to the challenge of the times remains to be seen. The European Court of First Instance's judgments in *Michelin II v. Commission* and *British Airways v. Commission*[3] do not seem to suggest much of an appetite for change. A second edition of this book in 2008 or early 2009, by when judgments in these cases are likely to have been handed down, would be most helpful.

All of the issues discussed above—and many others—are dealt with in *The Law and Economics of Article 82 EC*. It begins with some basic economics and a review of the scope of application of Article 82 and the procedural framework in which it is applied. Chapters on market definition and dominance, and an excellent one on the general concept of abuse, follow. The book then has detailed accounts of different types of abuse—predatory pricing, margin squeeze, exclusive dealing, refusal to deal, tying and bundling, exclusionary non-price abuses, abusive discrimination, excessive pricing, and other exploitative abuses. Each of these chapters is immensely useful, examining the arguments for and against condemnation of the practice in question, discussing the economic theories, and proposing workable solutions. The final two chapters deal with the concept of effect on trade between EU Member States and with remedies. There is no bibliography, which might be a useful addition to the second edition: there is ample reference to academic literature, speeches, and conference papers in the footnotes to the text, but it would be quite helpful to draw these together at some point.

3 Case T-203/01, Michelin II v. Commission, 2003 E.C.R. II-4071. Case T-219/99, British Airways v. Commission, 2003 E.C.R. II-5917.

Throughout the book, the law and economics are appropriately interwoven, as would be expected of a team representing both disciplines. The authors explicitly recognize the need for there to be rules that are administrable, with the consequence that a pure-effects approach would not be workable. They are opposed to per se rules and in favor of a so-called "structured rule of reason." They have produced a first-rate piece of work that will be highly influential in the years ahead, and that will be gratefully referred to by everyone interested in this fascinating but difficult topic. It is very highly recommended.

Volume 2 | Number 2 | Autumn 2006

Competition Policy International

Walter Eucken and Ordoliberalism: An Introduction from a Consumer Welfare Perspective

Christian Ahlborn and Carsten Grave

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

Walter Eucken and Ordoliberalism: An Introduction from a Consumer Welfare Perspective

Christian Ahlborn and Carsten Grave

This article serves two purposes, (i) to introduce "The Comparative Order and its Implementation," a seminal article published in 1949 by Walter Eucken, ordoliberalism's, or the "Freiburg School's," most prominent scholar, and (ii) to compare some ordoliberalist competition policy recommendations to those of a consumer welfare standard. The article provides an overview of the key concepts of ordoliberalism (such as "competitive order," "economic constitution" and "Ordnungspolitik") and outlines its implications for competition policy. It provides examples for the ordoliberal legacy in German and European competition policy, such as, inter alia, the market share thresholds for dominance, and the control of exploitative abuses such as excessive pricing. Finally, the article gives a critique of ordoliberalism from a consumer welfare perspective, and looks, among other things, at the implications of ordoliberalist policies for innovation and dynamic competition, the roots of the structure-conduct-performance paradigm, and the classification of certain forms of unilateral behavior (e.g., tying).

The authors are competition lawyers with Linklaters. The authors would like to thank Bill Allan, David Bailey, Liza Lordahl Gormsen, and Alison Oldale for their valuable comments. Any remaining blunders are, of course, the responsibility of the authors.

I. Introduction

Many of the foundations for German and EC competition policy were laid at an unlikely time, in an unlikely place: in 1933, following Hitler's ascent to power, two lawyers, Franz Böhm and Hans Grossmann-Doerth and one economist, Walter Eucken, met in the sleepy German university town of Freiburg, close to the French and Swiss border.

They discovered that they shared similar views about the failings of the Weimar Republic and equally rejected Nazi totalitarianism and socialism. Their interdisciplinary research of law and economics shaped the Freiburg School ideas which in turn provided the core of ordoliberalism.[1]

Ordoliberalism saw itself as a "third way" between the centrally planned economy of socialism and the unregulated market advocated by laissez-faire liberalism. For ordoliberalism, competition in the market economy would ensure a prosperous and humane society but only where law (and the enforcement of law through a strong state) would create and preserve the conditions under which competition would function properly.

After World War II, ordoliberalism had a significant impact on German economic policy, with many ordoliberals in key positions, most prominently Ludwig Erhard, the Minister of Economics for the first fourteen years of the Federal Republic. Ordoliberalism also gained prominence at European level where its philosophy of open markets fit well with the idea of European integration.[2]

It has been pointed out that "despite its enormous importance, ordoliberal thought—and German neo-liberal thought has received little attention in the English-speaking world."[3] The recent discussion about the modernization of European competition policy, and in particular Article 82 EC Treaty, where ordoliberal ideas clash with the views held by proponents of the consumer welfare approach, has fueled an interest in ordoliberalism.[4]

---

1 Key protagonists of ordoliberalism, in addition to Walter Eucken, Franz Böhm, and Hans Grossmann-Doerth, are Leonhard Miksch, Wilhelm Röpke, and Alexander Rüstow.

2 Walter Hallstein, the first president of the European Commission and Hans von der Groeben, one of the authors of the Spaak report and the first Competition Commissioner, were closely associated with ordoliberalism.

3 David Gerber, Law and Competition in the Twentieth Century Europe: protecting europe 232 (1998). Things have improved since David Gerber made this statement in 1998. In addition to the clear overview of ordoliberalism in Chapter VII of his book on *Law and Competition in the Twentieth Century Europe*, further work on Ordoliberalism has been published in English, including Victor Vanberg, The Freiburg School: Walter Eucken and Ordoliberalism (Walter-Eucken-Institut, Freiburg Discussion Papers on Constitutional Economics 04/11), Niels Goldschmidt & Arnold Berndt, Leonhard Miksch (1901-1950) – A Forgotten Member of the Freiburg School (Walter-Eucken-Institut, Freiburg Discussion Papers on Constitutional Economics 03/2).

4 *See*, for example, Liza Lovdahl Gormsen, *Article 82 EC: Where are we coming from and where are we going to?*, Competition L. Rev (2006).

It is, therefore, an opportune moment to republish, as part of the *Competition Policy International* series of antitrust classics, Walter Eucken's seminal article "The Comparative Order and its Implementation" which first appeared in *ORDO* in 1949.[5] The article has been chosen because it gives a broad overview of ordoliberalist foundations as well as policy recommendations for a number of policy areas such as competition policy and monetary policy, and thus, presents ordoliberalism the way it should be perceived in our view—a school of thought penetrating many areas of policy and not limited to competition. On the other hand, it also touches on a number of individual competition policy problems, which today are as topical as they were at the time Eucken published his article, such as the role of an independent competition authority at the core of competition policy's institutional framework and the (apparent) conflict between competition law and intellectual property rights.

This introduction to Walter Eucken and his work provides an overview of the key concepts of ordoliberalism (such as competitive order, economic constitution, and Ordnungspolitik) and outlines its implications for competition policy (Section II). It provides examples for the ordoliberal legacy in German and European competition policy (Section III) and gives a critique of ordoliberalism from a consumer welfare perspective (Section IV).

II. Ordoliberalism and Competition Policy

A. COMPETITIVE ORDER, ECONOMIC CONSTITUTION AND "ORDNUNGSPOLITIK": THE TRIAD OF ORDOLIBERALISM

1. The Competitive Order

Eucken contrasts two extreme types of "economic orders" (Ordnungen):

1. On the one hand, the market economy or "transaction economy" (Verkehrswirtschaft) where private and autonomous decision making determines economic activity according to incentives created by the competition process;

2. On the other hand, the "centrally administered economy" (Zentralverwaltungswirtschaft) where economic activity is the result of a bureaucratic process.

For Eucken, the transaction economy, and more precisely the competitive order underlying the transaction economy, is the key to a prosperous and humane

5 The article was originally published as Walter Eucken, *Die Wettbewerbsordnung und ihre Verwirklichung* (*The Competitive Order and its Implementation*), 2 ORDO, JAHRBUCH FUR DIE ORDNUNG VON WIRTSCHAFT UND GESELLSCHAFT 1-99 (1949), abridged version translated and reprinted in this issue as 2(2) COMPETITION POL'Y INT'L 219 (2006) (complete translation available at http://cpi.esapience.org). Hereinafter, where the Eucken article is cited, the first set of page citations refer to pages in Eucken's original article and the second set in parenthesis refer to pages in the reprint.

society. Competition and only competition would achieve sustained economic development.

However, not every form of competition can be expected to produce this beneficial outcome but only the form of "complete competition," i.e., competition in a market where no firm has the power to coerce conduct of other firms. According to Eucken, "[i]f there is competition on the supply side, as well as on the demand side and if the economic planning of both sides is based on such competition, then the market form of complete competition is achieved."[6] Eucken suggests two indicators to identify the market form of complete competition: first where "[t]he price is not forced upon the market by way of a market strategy, but is taken from the market," in other words, where all market participants are price takers; second, "certain measures . . . clearly indicate that complete competition does not exist because these measures cannot be implemented under complete competition: for example, obstructions to purchasers or suppliers that have dealings with competitors, or loyalty rebates or predatory pricing or dumping or destruction of stocks," in other words, complete competition is the market form of competition which ensures "performance competition" (Leistungswettbewerb).

Despite the fact that Eucken does not link complete competition to a particular market structure and despite Eucken's criticism of neoclassical economics of perfect competition, the concept of complete competition does have an underlying structural assumption of a polyopol and can best be understood as the real world adaptation of the model of perfect competition, i.e., a concept which is inspired by the model of perfect competition but which does not necessarily rely on its unrealistic assumptions.[7]

At the same time, competition has not only an economic but also a very important political dimension to Eucken and other ordoliberals: "Competition is by no means only an incentive mechanism but, first of all an instrument for the deprivation of power (Entmachtungsinstrument) . . . the most magnificent and most ingenious instrument of deprivation of power in history."

6 Eucken, *supra* n. 5, at 26 (at 230).

7 According to Moschel:

> the scholars of ordoliberalism have also used economic models for the description of their ideas, for instance, the model of perfect competition as it was developed in the traditional theory of competition. Such models, however, served only for the description of general effects of a market system, illustrating them in what might be called a chemically pure form. That did not imply, however, that those partly unreal premises were to be integrated as goals into practical competition policy. Any attempts to disprove or ridicule the ordoliberal concepts of competition as unrealistic miss this point.

Wernhard Moschel, *Competition Policy from an Ordo Point of View*, *in* H Willgerodt & A Peacock, German Neo-Liberals and the Social Market Economy ch. 7 (1989), at 146.

For this reason, the economic characteristics of complete competition also have a political equivalent: "In the same way as a law-based state, so the competitive order should also create a framework in which the individual's freedom to act is limited by the freedom of others, thereby ensuring a balance between every unit of human freedom."[8]

2. The Economic Constitution

The analysis of the competitive order has led Eucken to a number of insights:

- First, the two fundamental orders, the transaction economy and the centrally administered economy, are incompatible. Bringing together elements from these two orders in an actual economic system inevitably harms the functioning of that economic system.
- Second, there is an inherent self-destructive aspect to the competitive process of the transaction economy and to economic freedom:

> "The supplier and the customer always—wherever possible—seek to avoid competition and to acquire or assert monopolistic positions. There is an omnipresent, strong and irrepressible urge to eliminate competition and to acquire a monopolistic position. Everyone espies possibilities of becoming a monopolist. Why should three bakers in a 13th century town compete with one another? They could simply come to an agreement and create a monopoly. This was the situation earlier and the same applies today."[9]

The conclusion which Eucken and other ordoliberals drew from these insights was that the competitive order needed to be protected through a political and legal framework which would safeguard the efficient functioning of the competitive order and which would protect from any self-destructive tendencies. Here, Eucken foresaw a clear separation of roles for the state and the private sector:

> "The policy of competitive order does not leave the choice of market forms and monetary systems to the economy itself because the experience of the era of laissez-faire policy speaks for itself. The development of the framework in which businesses and households can plan and act freely is governed by

8 Eucken, *supra* n. 5, at 27 (at 231).

9 *Id.* at 5 (at 222).

the economic policy under which the framework is supervised. Businesses are free to choose what they produce, what technology they use, what raw materials they purchase and what markets they wish to sell on. . . . Freedom of the consumer exists, but not the freedom to choose how to define the rules of the game or the forms which the economic process takes. This particularly falls within the field of *Ordnungspolitik* (order-based policy)."[10]

3. Ordnungspolitik

The concept of "Ordnungspolitik" has been described as "the untranslatable soul of ordoliberalism."[11] Eucken describes the Ordnungspolitik as follows:

"[T]he crucial issue of modern economic policy should be treated as a crucial issue. This is to be achieved by making the establishment of a functioning price system of complete competition the essential criterion of every economic measure. This is the basic principle of the economic constitution."[12]

Eucken emphasizes that:

""[this] does not only demand the avoidance of certain acts of economic policy: such as state subsidies, the creation of mandatory state monopolies . . . Rather, a positive economic constitutional policy is required and its aim must be to further the development of the market form of complete competition and thus comply with the basic principle.""[13]

This general principle of the Ordnungspolitik means that "the body of doctrine of classical economic philosophy had to be translated from the language

10 *Id.* at 23 (at 227).

11 Gerber, *supra* n. 3, at 246.

12 Eucken, *supra* n. 5, at 33 (at 232).

13 *Id*. at 33 et seq (at 232).

of economics into the language of legal science."[14] Complete competition is, therefore, the guiding and, at the same time, the limiting principle for government policy.

B. COMPETITION POLICY IMPLICATIONS

1. Competition Policy at a Macro Level

For Eucken, "every measure of economic policy" was ultimately competition policy, in the sense that it was intended to safeguard and enhance complete competition. At this macro level of competition policy, Eucken distinguishes between "constitutive principles" (konstituierende Prinzipien) and "regulative principles" (regulierende Prinzipien). Constitutive principles ensure the establishment of the competitive order, regulative principles its continuous functioning.

For Eucken, the constitutive principles were:

(i) the primacy of monetary policy ("all attempts to implement a competitive order will fail, as long as monetary stability is not guaranteed");

(ii) the protection of open markets against state measures (for example, import bans) against private measures (among which Eucken included exclusivity arrangements) and against any combination of state and private measures (Eucken regarded in particular patents as a dangerous threat to open markets);[15]

(iii) consistency of economic policy over time;

(iv) private ownership;

---

14 Franz Böhm, Wettbewerb und Monopolkampf ix (1933).

15 *See, e.g.*:

> Patent law also belongs to the multitude of more recent legal institutions which did not have the effects desired by the legislator. Patent law was intended to promote technical development as well as to protect and reward the inventor. . . . Contrary to expectations, despite certain statutory precautionary measures, patent law has triggered a strong trend towards the establishment of monopolies and concentrations in the industry. This is due to the fact that patents create an exclusive right to manufacture an object, to bring it onto the market, use it and sell it. Although many of the patents do not close supply, these are patents which only cover a minor part of the production process of a commodity and which can be circumvented by using other production methods, substitute products and the like. However, a different type of patent group exists, namely master patents, which close the supply of goods, such as inter alia the well-known Telefunken patents for the production of radio tubes, or the benzopurpurin patent of 1884, which became an important feature for the organization of the chemical industry.

Eucken, *supra* n. 5, at 40 (at 236).

(v) freedom of contract; and

(vi) the need for clear and unambiguous allocation of liability.

2. Competition Policy at a Micro Level

Competition policy at a micro level, in the sense of antitrust, was regarded by Eucken as one of the key regulatory principles.

General aspects

According to Eucken, antitrust policy, and in particular the control of monopolies would ultimately fail, not least for political reasons, if large parts of the industrial sector were monopolized. By contrast, the situation would be different under the competitive order:

> "The creation of monopolistic power entities is prevented. Not only by prohibitions of cartels, but also—and far more importantly—by an economic and legal policy which breaks through the strong forces of competition, as exist in a modern economy, by applying the constitutive principles."[16]

As a result, according to Eucken, the pre-dominant market form in the competitive order is the market form of complete competition: monopolies and oligopolies are exceptions.

Regulation of monopolies

Eucken foresaw an independent competition authority (Monopoly Office) in charge of monopolies whose task it was to break up "avoidable monopolies" and to regulate "unavoidable monopolies." The basic principle for this regulation was the principle of "behavior analogous to competition" (wettbewerbsanaloges Verhalten) reflected in the "as if" approach: "The aim of monopoly legislation and monopoly supervision is to ensure that the bearers of economic power behave as if complete competition prevailed. The behavior of the monopolists should be 'analogous to competition.'"[17]

16 *Id.* at 65 (at 239).

17 *Id.* at 68 (at 241).

According to Eucken, in practice, this would mean the following:

> "Every form of impediment competition by embargos, loyalty rebates, predatory pricing, etc. is prohibited. . . . This creates a condition which would automatically arise in a complete competition situation, where impediment competition would be pointless. Admittedly, in order to achieve a result analogous to competition, it is necessary to introduce an obligation to contract, as here coercion is necessary to achieve the same result as would automatically arise under complete competition.
>
> As is generally known, under complete competition the same prices will become established for the same goods and services. Supply monopolies for example, whilst striving for the highest profit, have a tendency to demand differentiated prices for the same goods or services from individual segments of demand. This price differentiation should be prohibited in the competitive order.
>
> What is most difficult is to implement the fundamental principle within the scope of determining price levels. The price is to be fixed in such a way that offer and demand are in equilibrium at this price, and, at the same time, the marginal costs are just covered."[18]

Antitrust policy towards oligopolies

Eucken regarded oligopolies as a transient market form: "[t]his oligopoly—or part oligopoly—situation often passes by rapidly, and soon leads to the creation of a cartel, i.e., to a collective monopoly or an individual monopoly, by overpowering the opponent." However, as sometimes "the unstable condition of the oligopoly or part oligopoly exists for many years or decades." Ordoliberals still saw the need to address the issue. Eucken records two conflicting views among the ordoliberals:

> "According to the first opinion, as has been put forward particularly impressively by Miksch, a special regulation is necessary for oligopolies and part oligopolies: namely the "tied competition" which takes place under state supervision.
>
> . . .
>
> According to the other view, this is too great a burden for the state. According to this view, an active monopoly supervision is indeed sufficient,

18 *Id*. at 69 (at 242).

> and it also offers something better for such cases. With a decisive monopoly supervision, the oligopolists have no reason to destroy the others by aggressive means or to attain a position of monopoly of their own. This is because it comes up against a rigorous monopoly control. Furthermore, the oligopolists themselves will attempt to behave as if complete competition prevailed, as they will otherwise come to the individual attention of the monopoly office. An example: a cement cartel is dissolved. As a result the seven members become oligopolists. That one company now seeks to overpower the others is unlikely. This is because all measures of impediment competition—predatory pricing, blockades, loyalty rebates etc. are prohibited and punishable. If, however, it becomes a monopolist by using competitive means, it would be subject to the comprehensive, deterring supervision of the monopoly office."[19]

III. Ordoliberal Legacy in German and European Competition Policy

Over time, ordoliberalism has inevitably lost influence over EC competition policy as the number of Member States of the European Union has grown from six to twenty-five and there has been a proliferation of national competition regimes.

Nevertheless, due to the huge impact ordoliberalism had on EC competition policy, many of the ordoliberal concepts have been hard-wired into the system, even when at times the link to ordoliberalism has been obscured or forgotten. During the discussion about the modernization of Article 82 EC these ordoliberal concepts, which are difficult to reconcile with a consumer welfare approach, have frequently been the focus of the debate. Examples are:

DUE TO THE HUGE IMPACT ORDOLIBERALISM HAD ON EC COMPETITION POLICY, MANY OF THE ORDOLIBERAL CONCEPTS HAVE BEEN HARD-WIRED INTO THE SYSTEM, EVEN WHEN AT TIMES THE LINK TO ORDOLIBERALISM HAS BEEN OBSCURED OR FORGOTTEN. DURING THE DISCUSSION ABOUT THE MODERNIZATION OF ARTICLE 82 EC THESE ORDOLIBERAL CONCEPTS HAVE FREQUENTLY BEEN THE FOCUS OF THE DEBATE.

- the low threshold of dominance;
- the "special responsibility" of dominant firms;
- the formalistic (quasi per se) approach towards many types of unilateral behavior; and
- the prohibition of exploitative abuses.

19 *Id*. at 71 (at 245).

A. MARKET STRUCTURE AND THE THRESHOLD OF DOMINANCE

EC and German competition law regimes rely on market shares as an important element in the assessment of dominance. This is by no means unusual. What distinguishes them from some of the other regimes, however, is the intervention threshold. Under German law, a company "is presumed to be dominant if it has a market share of at least one third."[20] Under EC competition law, the (rebuttable) presumption of dominance kicks in where market shares are in excess of 50 percent.[21] This contrasts with an assessment under U.S. law where a market share of 70 to 75 percent for at least five years is required to lead to a presumption of monopoly power.[22]

The low threshold and the absence of any reliance on persistency of market shares can be traced directly back to ordoliberalism with its ideal of complete competition, i.e., a market where all participants are price takers, and its concern even about short-term market power. From an ordoliberal perspective, a market share with 10 suppliers of 10 percent each seems to be inherently better (both from an economic and political perspective) than one where one player has 40 percent of the market and 9 players have between 6 and 7 percent. This view seems to be reflected in the *British Airways/Virgin* case, where the Commission held:

> "Despite the exclusionary commission schemes, competitors of BA have been able to gain market share from BA since the liberalisation of the United Kingdom air transport markets. This cannot indicate that these schemes have had no effect. It can only be assumed that competitors would have had more success in the absence of these abusive commission schemes."[23]

Assuming that "effect" refers to anticompetitive effect (rather than just any effect), then the Commission seems to suggest in this paragraph that a counterfactual in which "competitors [of a dominant firm] would have had more success" is per se a more competitive outcome.

20 Section 19(3) of the German Act against Restraints on Competition.

21 Case C-62/86, AKZO v. Commission, 1991 E.C.R. I-3359, at para. 60.

22 Phillip E. Areeda & Herbert Hovenkamp, Fundamentals of Antitrust Law 801a, 319 (2003).

23 Case IV/D-2/34.780, Virgin/British Airways, 2000 O.J. (L 30) 1, at para. 107.

B. THE "SPECIAL RESPONSIBILITY" OF DOMINANT FIRMS

Under EC competition law, the statement that it is not an offense for a firm to have a dominant position comes invariably with the qualification that by imposing such firms "a special responsibility not to allow its conduct to impair undistorted competition on the common market."[24]

The scope of this special responsibility is not entirely clear. In a narrow sense, it can be interpreted as saying no more than that Article 82 EC imposes obligations on dominant firms which are not imposed on non-dominant firms. A wider interpretation would suggest that it must refrain from any action which would increase its market power and harm competitors even where the behavior is efficiency-based.

The origin of the special responsibility can be traced back to the ordoliberal "as if" principle according to which firms which are not price takers, i.e., which posses (significant) market power, do not only have a negative obligation (i.e., not to commit certain harmful acts), but also a positive obligation (i.e., to behave as if they did not have any market power).

C. FORMALISTIC (I.E., PER SE) APPROACH TOWARDS ABUSE

The assessment of abuse under German and EC competition law has equally been shaped by ordoliberalism.

As described above, for Eucken and other ordoliberals, certain types of unilateral behavior, such as price discrimination, loyalty rebates and tying, were inherently abusive, i.e., clear examples of impediment competition with no redeeming features.

The view that certain types of unilateral behavior are per se harmful and therefore do not require any effects analysis has been re-iterated by the Court at regular intervals, most recently in *Michelin II* and *British Airways v. Commission*.[25] The Court took the view that it was sufficient for an abuse that the conduct "tends to restrict competition," i.e., "is capable of having, or likely to have, such an effect." According to the Court in *British Airways v. Commission*, "where an undertaking in a dominant position actually puts into operation a practice generating the effect of ousting its competitors, the fact that this hoped-for result is not achieved is not sufficient to prevent a finding of abuse."

24 Case 322/81, Michelin v. Commission, 1983 E.C.R. 3461, at para. 57.

25 Case T-203/01, Michelin v. Commission and Case T-219/99, British Airways v. Commission, appeal pending before the European Court of Justice, Case C-95/04.

D. THE CONTROL OF EXPLOITATIVE ABUSES

The control of exploitative abuses reflects the ordoliberal principle of forcing dominant firms to behave "as if" they were subject to complete competition. Indeed, the wording of Article 82 ("Such an abuse may, in particular, consist in: . . . (e) directly or indirectly impose unfair purchase or selling prices or other unfair trading conditions . . .") echoes Eucken's list of prohibited monopoly practices in his chapter on the "monopoly problem in the competitive order,"[26] general terms and conditions which should alter general contract law to the disadvantage of the trading partner of the dominant firm,[27] and prices which are in excess of the equilibrium price, i.e., where prices equal marginal costs.[28]

In an attempt to make Eucken's "as if" principal operational, methodologies for determining excessive prices have been developed in the early years of German and EC competition policy, for example, in *United Brands* which compared actual costs and prices and the prices of the dominant firm with that of its competitors.

The practical problems of price control which Eucken recognized (but arguably nevertheless underestimated) had the effect that only a small number of exploitative abuses were pursued under German and EC law.[29]

EC and German law contrasts with the U.S. policy under Section 2 of the Sherman Act whereby firms which have lawfully acquired monopoly power are entitled to exploit it, in other words where the concept of exploitative abuses does not exist.

IV. Ordoliberalism: A Consumer Welfare Perspective

A. THE CONSUMER WELFARE PERSPECTIVE

Contrary to ordoliberalism, the consumer welfare approach does not represent a coherent school of thought, but merely an agreed view on a number of general principles of antitrust enforcement.

26 Eucken, *supra n.* 5, at 68 et seq (at 240).

27 *Id.*

28 *Id.* at 69 (at 243).

29 The most prominent cases under EC law are Case 27/76, United Brands, 1978 E.C.R. 207, Case 26/75, General Motors, 1976 E.C.R. 1367, Case 226/84, British Leyland, 1986 E.C.R. 3263, and Case COMP/36.568, Scandlines Sverige v. Port of Helsingborg, *available at* http://ec.europa.eu/comm/competition/antitrust/cases/decisions/36568/reject_en.pdf.

There is, in particular, the general agreement that the only goal of competition policy should be consumer welfare,[30] to the exclusion of other goals such as fairness or the limitation of the power of large firms. This, in turn, means that government intervention is only justified in the case of actual consumer harm or, where the analysis is prospective, likely consumer harm, i.e., the consumer welfare approach is effects-based. This is not to say that proponents of a consumer welfare approach are necessarily in favor of an assessment of the direct effect on consumer harm in the form of an unstructured rule of reason; arguably the contrary is the case, namely that the majority of advocates of a consumer welfare approach would accept that the antitrust tool box is not sufficiently precise to allow for a direct assessment of consumer harm and that an indirect screen (such as the "no economic sense" test advocated by Greg Werden)[31] is required. However, the link between the indirect screen and consumer harm has to be established rather than merely assumed.

B. COMPARING THE INCOMPARABLE?

Is it appropriate to assess ordoliberalism from a consumer welfare perspective? It has been argued that ordoliberalism and the consumer welfare approach pursue fundamentally different policy goals: ordoliberalism has a political as well as an economic dimension with the preservation of economic freedom and complete competition as primary goals while the perspective of the consumer welfare approach is purely economic with the primary goal of consumer welfare maximization.

Clearly one should proceed with caution in comparing ordoliberalism and the consumer welfare approach, and in particular viewing one through the lens of the other. Nevertheless, a comparison seems feasible, despite the underlying differences in policy objectives:

First, while there is both a political and an economic motivation for the preservation of complete competition and economic freedom under ordoliberalism, at least from an ordoliberal perspective, there does not seem to be a conflict between the political and economic goals (in other words the optimal outcome from a political perspective happens to be the optimal outcome from an economic perspective); hence, ignoring the political aspect of ordoliberalism (at least according to ordoliberalism's own logic) should ultimately not affect the comparison.

Second, while economic freedom is the guiding star for ordoliberalism, there is the underlying assumption that the preservation of economic freedom will indirectly lead to technological progress and allocative efficiency.

---

30 As regards the dispute whether consumer surplus or total surplus standard should be pursued, *see* Massimo Motta, Competition Policy 19 et seq (2004).

31 Gregory Werden, *Identifying Exclusionary Conduct under Section 2: The "No Economic Sense" Test*, 73 Antitrust L.J. 413 (2006).

To the extent that ordoliberalism and the consumer welfare approach result in fundamentally different policy decisions, this raises questions about the validity of the two underlying assumptions, namely:

(i) whether there is actually no conflict between political and economic goals, and

(ii) whether the preservation of economic freedom does lead to consumer welfare maximization, at least in the long run.

A second possible objection to viewing ordoliberalism through a consumer welfare perspective is one of chronology. Ordoliberalism was developed by Eucken and others in the 1930s and 1940s while the consumer welfare approach reflects today's state of the art antitrust thinking, benefiting from the insights of the Chicago School and advances in game theory. This objection would indeed be valid were it not for the fact that ordoliberalism drives competition policy still today, in relatively undiluted form by the Federal Cartel Office in Germany and, at least partially and indirectly, by the European Commission in Brussels. Therefore, rather than contrasting economic thinking of the 1940s with economic thinking of 2006, we are contrasting two forms of intellectual underpinnings of current competition policy.

C. COMPETITION POLICY AT THE MACRO LEVEL

It is difficult to appreciate fully the towering intellectual achievements of ordoliberalism when many of their key findings are nowadays so universally accepted that they sound commonplace. The view that "transaction markets" are far superior to centrally administered markets, both from an economic and political perspective and that the price mechanism in competitive markets should determine how scarce resources are allocated has become self-evident in a world with few supporters for socialism outside North Korea and Cuba. This was not the case at the end of World War II when "[t]hroughout Europe, liberalism had been so thoroughly discredited that few wished to be associated with it [and] socialist solutions had captured the public imagination and/or the fancy of intellectuals."[32]

The same is true for the idea that competition may need protection through the legal framework, both in terms of antitrust policy and through a coherent set of economic policies more generally. At the end of World War II, only the United States had an antitrust policy. The number of countries with competition policy enforcement now exceeds eighty and it is a fairly recent development that certain areas of the legal framework, such as intellectual property rights, are closely assessed according to their impact on the competitive process.

The ordoliberal proposal of an "independent competition authority subject only to the law, in order to make it immune against the dangerous influences of

32 Gerber, *supra* n. 3, at 257.

third parties" is still gaining acceptance. Not long ago, the Office of Fair Trading in the United Kingdom gained greater political independent and other countries, such as Spain, still follow this path.

Finally, the ordoliberals "idee fixe," namely that the right legal framework is a key element for a stable and functioning market economy which in turn guarantees a stable and functioning democracy, has lost much of the urgency it had in 1945. This does not, however, invalidate the point.

D. COMPETITION POLICY AT THE MICRO LEVEL

1. Overview

The views of ordoliberalism regarding competition policy in the narrow antitrust sense are less universally accepted today. Developments in economic theory over the last forty years have raised doubts about the concept of complete competition, the "as if" principle of intervention, the goal of economic freedom, and the distinction between performance competition and impediment competition.

THE VIEWS OF ORDOLIBERALISM REGARDING COMPETITION POLICY IN THE NARROW ANTITRUST SENSE ARE LESS UNIVERSALLY ACCEPTED TODAY. DEVELOPMENTS IN ECONOMIC THEORY OVER THE LAST FORTY YEARS HAVE RAISED DOUBTS ABOUT THE CONCEPT OF COMPLETE COMPETITION, THE "AS IF" PRINCIPLE OF INTERVENTION, THE GOAL OF ECONOMIC FREEDOM, AND THE DISTINCTION BETWEEN PERFORMANCE COMPETITION AND IMPEDIMENT COMPETITION.

2. Complete Competition and Innovation

For ordoliberalism, the ideal form of competition is competition in a market in which none of the players has any (significant) market power or, as a second-best solution, in which firms with market power are forced to behave as if they did not have any market power, in terms of pricing ("[t]he price is to be fixed in such a way that offer and demand are in equilibrium at this price, and, at the same time, the marginal costs are just covered")[33] and other parameters.

This approach overlooks the dynamic nature of competition and the role of (temporary) market power as a key driver for innovation. As the U.S. Supreme Court put it in *Trinko*: "[Striving for monopoly is] an important element of the free-market system. The opportunity to charge monopoly prices—at least for a short period—is what . . . induces risk taking that produces innovation and economic growth.[34]

33 Eucken, *supra* n. 5, at 69 (at 243).

34 Verizon Communications Inc v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398, 407, 414 (2004).

This tension between static competition on the one hand and the exploitation of market power as an incentive for dynamic competition was also described by Attorney General Jacobs in *Bronner* where he cautioned against the frequent use of "as if" intervention such as imposing an obligation to deal on a dominant firm:

> "The justification in terms of competition policy for interfering with a dominant undertaking's freedom to contract often requires a careful balancing of conflicting considerations. . . . For example, if access to a production purchasing or distribution facility was allowed too easily there would be no incentive for a competitor to develop competing facilities. Thus while competition was increased in the short term it would be reduced in the long term."[35]

Ordoliberalism does not seem to acknowledge this inherent tension between short-term (static) competition and long-term dynamic competition or, in any event, to decide clearly in favor of maximizing short-term static competition.

In light of the fact that innovation and dynamic competition are the key drivers behind increases in consumer welfare, such a policy choice is at least questionable.

3. Economic Freedom, Complete Competition, and Performance Competition

Eucken and other ordoliberals assumed a virtuous circle between deconcentrated markets (i.e., complete competition) and economic freedom on the one hand and performance competition on the other: performance competition would safeguard complete competition and economic freedom and, conversely, complete competition and economic freedom (re-enforced by Ordnungspolitik) would, ensure performance competition. Consequently, under the competitive order, complete competition would be the pre-dominant market form, with the odd unavoidable monopoly to regulate.

Reality does not seem to confirm either part of this virtuous circle. There is no indication that performance competition leads to or preserves deconcentrated market structures. Indeed many competitive industries, in particular innovative industries with high investment costs and considerable risks, such as the pharmaceutical or software sectors, tend to gravitate towards oligopolistic market structures, while in other sectors, performance competition may result in a wide range of possible market structures.

35 A.G. Jacob's Opinion in Case C-7/97, Oskar Bronner GmbH & Co KG v. Mediaprint Zeitungs-und Zeitschriftenvertrag GmbH & Co KG and Others, at para. 57.

The second part of the virtuous circle is known under the name of Structure-Conduct-Performance paradigm (SCP paradigm)[36] advocated by the Harvard School. The SCP paradigm implies that performance (e.g., profit margins, production levels, consumer benefits) in certain industries is dependent on the conduct of buyers and sellers (e.g., advertising, R&D, investment), which in turn is dependent on the structure of the market (e.g., number of sellers, concentration ratio, barriers to entry). The paradigm was largely based on empirical work which suggested that profitability of an industry increases with increasing levels of concentration and higher barriers to entry.[37] During the heydays of the Harvard School in the 1950s and 1960s, U.S. antitrust policy had some resemblance with ordoliberalism, in particular the focus on keeping markets de-concentrated and a hostility towards large firms.

Since then, the view of a simple causal relationship between market structure, behavior, and performance has been rejected (both in relation to static and dynamic competition) in favor of a more complex interaction between structure and behavior; in particular, it has been recognized that behavior (e.g., raising barriers to entry) may affect market structure and that the size of firms in an industry may reflect superior efficiency.

This has important implications: to the extent the virtuous circle between complete competition and economic freedom on the one hand and performance competition on the other hand does not hold, conflicts between political and economic goals of competition policy may well occur. In other words, deconcentrated markets with a relatively large number of players which are likely to be optimal from a political perspective may not lead to an optimal economic outcome and vice versa.

To the extent performance competition could lead away from complete competition and economic freedom, a policy which protects economic freedom risks treating efficiency as an offense. Economic freedom will therefore not necessarily provide a good proxy for consumer welfare; in other words, a competition policy which protects economic freedom may in certain circumstances lead to consumer harm.

4. Performance Competition and Impediment Competition

Eucken and other ordoliberalism draw a clear line between performance competition and impediment competition: performance competition takes place in markets with complete competition; by contrast, impediment competition, according to Eucken, is a clear indication that the market is not characterized by

36 *See* E. S. Mason, *Price and Production Policies of Large Scale Enterprises*, 29 AMER. ECON. REV. 61 (1939).

37 J. Bain, *Relation of Profit Rate to Industry Concentration: American Manufacturing 1936-1940*, 65 Q.J. ECON. 293 (1951).

complete competition. Eucken lists price discrimination, tying, loyalty rebates, and exclusive agreements as examples of impediment competition.

In the last forty years, economic theory has made significant progress and many of the practices which were previously held to be primarily or exclusively anticompetitive, have turned out to have many pro-competitive explanations. Indeed, closer inspection has shown that most of the practices previously considered exclusively or predominantly anticompetitive occur in deconcentrated markets in which none of the market participants has any market power (i.e., real-world examples of complete competition).

Price discrimination

Eucken's claim that price discrimination is a clear indicator of market power and a sign of a defective competitive process has not been confirmed by economic research.[38] While it is a truism that price discrimination requires some control over prices (i.e., a downward sloping demand curve), it is observed in many competitive markets which come close to the ordoliberal ideal of complete competition, such restaurants. Rather than being an unequivocal sign of abusive behavior, price discrimination may frequently be evidence of a welfare-enhancing allocation of common costs (i.e., some form of Ramsey pricing).

Tying

Tying is equally a practice frequently observed in competitive markets.[39] Drugs companies tie cough and cold remedies, suppliers of electrical appliances bundle different foreign electrical adapters, and car manufacturers sell cars with radios (and CD players) as standard equipment. Again there are valid efficiency reasons, in particular product-specific scale economies.

Loyalty rebates

Loyalty discounts are also persuasive in many industry sectors, including sectors in which there are no dominant firms: "Nobody would claim that the coffee shop on the street corner offering a free espresso for every ten Euro of sales is doing so for sinister exclusionary motives."[40]

Indeed, several pro-competitive motives for loyalty rebates have been advanced, ranging from aligning the interests of manufacturers and distributors to risk allocation and solution to hold up problems. This means that the clear

38 *See, e.g.,* Michael E. Levine, Price Discrimination without Market Powers (Harvard Law School, Law-Econ Discussion Paper No. 276, 2000).

39 David S. Evans & Michael Salinger, *Why Do Firms Bundle and Tie? Evidence from Competitive Markets and Implications for Tying Law*, 22 Yale J. Reg. 37 (2005).

40 David Spector, *Loyalty Rebates: An Assessment of Competition Concerns and a Proposed Structured Rule of Reason*, 1(2) Competition Pol'y Int'l 93 (2005).

dividing line between performance competition and impediment competition, which Eucken and other ordoliberals saw, is not confirmed by modern economic theory. So-called impediment competition, rather than being a clear indicator of a defective competitive process, turns out to be behavior which is frequently efficiency motivated, as evidenced by such practices by firms without market power and where it is not easily apparent whether such behavior would be impediment competition even if carried out by a dominant firm.

If one accepts that types of behavior cannot simply be classified as either pro- or anticompetitive and that such behavior occurs in competitive as well as noncompetitive markets, then it becomes clear that the ordoliberal approach does not provide an operational set of criteria which would allow it to distinguish between (permitted) performance competition and (prohibited) impediment competition. Eucken's definition of performance competition as competition which occurs in markets with complete competition (and of impediment competition as competition which does not occur in markets with complete competition) would leave the concept of impediment competition as an empty shell as most types of unilateral behavior which could be harmful if carried out by a dominant firm in certain circumstances would frequently also have efficiency justifications in certain circumstances and, therefore, would be regularly observed in markets without dominant firms. Conversely, a definition which marks as impediment competition any behavior which in the case of a dominant firm could in certain circumstances be harmful would be vastly over-inclusive.

5. Direct Intervention and Error Costs

Finally, while Eucken and the ordoliberals generally favor an approach of indirect regulation whereby the state determines the legal and political framework but does not intervene directly in the competitive process, this principle is abandoned in relation to the "as if" principle of intervention, particularly in relation to exploitative abuses.

Even in Eucken's relatively clear-cut world with many markets of complete competition and a few markets with (unavoidable) monopolies as well as simple pricing rules (i.e., price equals marginal costs), Eucken recognizes that price regulation is a difficult exercise.

In a world with many more shades of grey, with market structures which cover the full range from atomistic to monopolistic, where it is not obvious which market structures are "unavoidable" and where there are no clear forcing rules (e.g., because in many markets a price equal to marginal costs would not allow firms to recover their fixed costs) these difficulties multiply.

The question that arises, particularly in relation to exploitative abuses, is in which circumstances government intervention is superior to no intervention at all.

V. Conclusion

The ordoliberalist school of thought as spearheaded by Walter Eucken was a critical contribution to setting Germany on a successful course (economic miracle) against all political odds after World War II. In many ways, Eucken and his colleagues were visionaries, laying the intellectual foundations for concepts that we today may take for granted, but that are in no way less relevant as they were at the time:

- the economic constitution as a nation's (and national economy's) fundamental choices for the prevalent economic system;
- the need to protect competition, both against certain state measures as well as against enterprises that command economic and eventually also political power to an extent that allows them to abolish competition; and
- political independence of competition enforcers, which were to be subject only to the law, in order to shield them from pressure groups.

In other areas, however, things have moved on and a true to original application of ordoliberalist competition policy recommendations seems not warranted any longer (e.g., in assuming that certain unilateral conduct such as tying is always anticompetitive or a form of impediment competition, which we today know it is not). This shall not be interpreted as a criticism to Eucken and his counterparts, whose contribution to the economic policy debate must be seen before the correct historical background. It is, however, a criticism to the guardians of ordoliberals of today which have prevented that ordoliberalism reflects recent (and not so recent) economic developments.

In other areas, however, things have moved on and a true to original application of ordoliberalist competition policy recommendations seems not warranted any longer.

VOLUME 2 | NUMBER 2 | AUTUMN 2006

Competition Policy International

The Competitive Order and Its Implementation (English Translation)

Originally published in *ORDO - Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft*

Walter Eucken

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2006, Vol. 2, No. 2.
Competition Policy International is a free publication. For additional copies, please visit cpi.esapience.org or call +1.617.844.1800.

The Competitive Order and Its Implementation

Walter Eucken

Originally published as Walter Eucken, *Die Wettbewerbsordnung und ihre Verwirklichung (The Competitive Order and Its Implementation)*, 2 ORDO, JAHRBUCH FÜR DIE ORDNUNG VON WIRTSCHAFT UND GESELLSCHAFT 1-99 (1949). English translation by Christian Ahlborn and Carsten Grave of Linklaters. Due to the article's length, some sections have been omitted (indicated by * * *). The complete translation is available online at http://cpi.esapience.org.

Part One: Economic Policies of the Past

How can a functioning and humane order be given to the modern, industrialized economy? This question is asked from a totally different position today than it was by those men asking the same question at the beginning and in the middle of the 19th century[1] This was the time when industrialization began. The liberals living in the middle of the century or Sismondi or the St. Simonists or Marx or Proudhon—in short all thinkers of those times—lived in a different economic and social world than we do today. The issue of an industrial workforce was already becoming heated, but this issue was a totally different issue then to what it is now. At that time the only experience was of the pre-industrial economy, and the onset of the great revolution. "Society today is subject to totally new conditions of existence, of which we have no experience," Sismondi wrote in 1827.

At that time corporate groups, cartels, credit banks and trade unions either did not exist at all or were only just starting to develop. There was no experience (which we have today) of a central direction of the industrial processes of the economy. However, it was in these past times that doctrines were developed which determine society today, and in which definitions such as that of socialism and capitalism were created, which most people still use today. In the meantime, however, industrialization in the old industrial states has passed through manifold stages and spread massively throughout the world. We could and should

1 See my essay "Das ordnungspolitische Problem" in ORDO, volume 1.

leave the stage of speculation in order to enter the stage of experience-based economic policy. We can draw upon considerable experiences in the areas of monetary policies, crisis policies, agricultural policies, cartel policies, trading policies, tax policies and the like. Indeed these experiences ought to be exhausted; selective descriptions are not sufficient.

In the 1870s, Hippolyte Taine abandoned his studies of literature, art and philosophy to devote himself to the current problems of economy, administration and law as he recognized their vital importance. For years, he locked himself away and concentrated passionately on his new task. He did not arrive at any definite result, however; he found no simple solution and no general principle. In these matters, he advised—as his dependents wrote—that one should "proceed tentatively, accepting the irregular and the incomplete, initially making do with partial solutions and continuing to work on ascertaining the laws and establishing the general provisions, that render this or that result possible or impossible." With his sharp eye for reality, Taine correctly identified the situation at that time. In those days, people did not know enough about the industrial and technical economy in order to provide it with a sufficient constitution. Today, this is different. The last half-century, in particular, with its rapid changes in economic policy, with interventionism, full-employment policy, experiments involving the centrally administered economy and the fast pace of industrialization, has taught us a great deal. Today, Taine could obtain much more specific results. The fundamental questions of the world and of mankind are not endemic to any particular period: the sages of yore tell us as much as Kant or Goethe. But new problems of economic policy were created with mechanization, industrialization, de-individualization and urbanization and it is only now with the experience of history that people are equipped to deal with these problems. Now is the time, however, as otherwise man will be crushed by the industrial machine.

We should thus familiarize ourselves with the economic experiences of the last one and a half centuries, in order to find a basis for solving the large order-based, policy-related problems of the industrial economy. Perhaps it is expedient to divide the economic policy from former industrial ages into two eras: the long era of "laissez-faire politics" and the subsequent shorter era of "economic experimentation." It may then become necessary to enter a new, third era.

CHAPTER 1: THE POLICY OF LAISSEZ-FAIRE

1. What was the basis of this economic policy? The answer is usually: It was the time of a "state-free economy."

A short glance at the historical reality would have shown that this answer is incorrect. It was precisely during this period that the state created a strict law of ownership, contract law, company law, patent law, etc. Every business and every household was subject to these state-made laws on a daily basis, whether it wished to buy or sell, take out a loan, or carry out any other commercial activity. In this situation, how can one speak of a "state-free economy"?

In this case, what was, in fact, the policy of laissez-faire? In this period states created *governmental* orders or constitutions for themselves, in order to introduce a functioning state apparatus and to protect the freedom of the individual. In the same way, they created *legal* systems by comprehensive codifications. However, the *economic* order and its development were not regarded as a particular task for the state. There was a conviction that a sufficient economic order would automatically develop within the scope of the law.

2. We know that at all times and in all places, everyday processes of the economy take place within the framework of particular forms. *The economic order is the sum of the forms that are realized in a particular country and at a particular time*. Because in industrial nations, which have an economy based on the intensive division of labor, the entire everyday economic process takes place in millions of separate households and businesses, but forms a single unit, in the Industrial Age it is necessary for the economic order to have a uniform controlling instrument to reliably control the overall process.

The economic policy of laissez-faire was originally based on the conviction that reasonable forms, i.e. a practicable general economic order, will develop of their own volition from the spontaneous forces of society if freedom exists and the principle of the rule of law is observed. Only in certain contexts—such as in note-issuing banks or the system of trade treaties—was the attempt made to shape parts of the economic order, but in general, the state left the forms of business to the private sector.

3. The policy of laissez-faire essentially dominated for over a hundred years. Broadly speaking, it was the economic policy of the 19th century. Its successes were significant. Industrialization established itself in Europe during this period and the supply of goods to a population which had more than doubled reached a level that had never been attained before. At the same time, an international economic order developed which functioned in a relatively rational way.

Yet, at the same time, significant damage was done: one need only think of the crises and the social tension. The freedom that the law-based state intended to guarantee was in practice threatened by the formation of factions of economic power. Employees became dependent upon employers, consumers upon monopolists, dealers upon groups of companies and cartels. The problem of distribution was not sufficiently solved and the kind of solution applied stirred up resistance among broad groups within society.

Criticism of the economic policy of laissez-faire is often oversimplified. Some fad word, such as "capitalism" or "Manchester capitalism," is presented and it is then reported what shameful acts were committed by this creature. Magical thought has supplanted observation and analysis in this area. Marx's criticism is often reiterated, yet it is antiquated, uses observations from the early period of industrialization and fails to take account of developments thereafter; not even the differences between market forms and monetary systems, without the application of which criticism is useless. Even Keynes' criticism is global and oversimplifies the task.[2] More incisive criticism is possible, and is necessary. It is necessary because it is particularly the economic policy of laissez-faire—that formed the basis for the further development of economic policy—which offers an abundance of detailed economic experience.

4. The economic policy of laissez-faire was originally based on the conviction that competition would arise ubiquitously where it was applied, and that in a competition situation the workforce and the flow of goods could be expediently distributed among the businesses and households in order to satisfy demand in the best possible way. Now, however, the following could be seen:

Firstly: The supplier and the customer always—wherever possible—seek to avoid competition and to acquire or assert monopolistic positions. There is an omnipresent, strong and irrepressible urge to eliminate competition and to acquire a monopolistic position. Everyone espies possibilities of becoming a monopolist. Why should three bakers in a 13th century town compete with one another? They could simply come to an agreement and create a monopoly. This was the situation earlier and the same applies today. The employers on the labor market, the suppliers on the goods markets and the workers aim for monopolistic positions. As soon as this aspiration succeeds, and monopolies, partial monopolies or even oligopolies become established in a market, the control mechanism of the prices, in an industrialized economy in particular, no longer functions sufficiently. The prices do not accurately express the shortage of goods. Monopolies make the shortage of goods appear greater than it actually is, as can be seen from the destruction of stored goods or from a reduced use of an existing production apparatus. At the same time, however, an aspiration for power and for monopolistic positions can also give rise to markets without equilibrium: two-sided monopolies or two-sided partial monopolies or even oligopolies. Strikes or lockouts are clear signs of such a lack of equilibrium in the labor markets.

Secondly: The prices partly failed to direct the economic process because the monetary systems were inadequate. In a modern economy, money is predominantly created through the granting of loans by banks, and disappears through the repayment of such loans. This fact is extremely important for the development of the economy in the 19th and the start of the 20th centuries. It has been decisive in promoting industrialization. However, the connection between credit volumes and the amount of money has led to an instability of money, to expansions and to contractions, which prevented the price mechanism from working reliably.

2 For a review of the economic policy of laissez-faire and its theoretical decline: A. Rüstow, "Das Versagen des Wirtschaftsliberalismus als religionsgeschichtliches Problem," 1945.

These two factors had a combined effect: the creation of pressure groups in the form of cartels, trusts, corporate groups, employer associations, trade unions and the instability of money. The extremely difficult task of providing millions of businesses and households with the correct production factors or goods in the correct proportions in due time, and of directing the entire interconnected economic process of the people towards the meeting of demands, and of selecting the correct investment possibilities in balanced proportions from the infinitely large number of possibilities available, is not possible if the prices arise within the scope of a power struggle between concentrated groups of industry, agriculture and workforce.

5. This was the mistake made in the principles and policies of the laissez-faire society or the free economy in the old style: It left both the battle for establishing the rules of the game, the framework and the forms of the economy, as well as the daily struggle for quantities and prices, up to the individual. It allowed this free battle for the forms of order, as long as certain legal principles were adhered to. Little regard was paid to the fact that this had the effect of creating pressure groups and allowing the establishment of monetary systems which did not fully function. It was hoped that the "invisible hand" mentioned by Adam Smith would create a successful system, and that the daily economic process would proceed in this system without friction. One did not distinguish between the different forms of economy and the daily economic process. *Perhaps* it is possible that in certain forms of a transaction economy the daily process may function without friction. However, the possibility that viable systems would arise by themselves was overestimated.

As soon as pressure groups came into existence, a circulus vitiosus became apparent. The pressure groups gained influence in economic and legal policy. They implemented, for example, trade policy measures which consolidated their positions. Their "general terms and conditions" excluded a large part of the law enacted by the state.[3] The economic policy of laissez-faire thus slid into an economic policy of interventionism. Interventionism is a continuation of and increase in the politics of laissez-faire. Because the pressure groups were supported by the state, they acquired a new power—which could also extend to the legal decision-making within the state. In a system of interventionism, the state supported the individual interested parties in their aspiration to secure certain positions of power or markets. However, it did not aspire to produce the conditions for a sufficient economic order.

6. Experiences with the economic policy of laissez-faire thus show:

Firstly: Prices and price ratios have not proved suitable to direct the daily process of the industrialized economy in all market forms. It is true that this finding cannot be grossly exaggerated. If the price system in the age of laissez-faire did not fully function, this does not mean that, for example, the price mechanism is

3 See note 39.

generally incapable of directing the economic process. The time of laissez-faire has merely shown that within the scope of certain market forms and monetary systems, the price mechanism does not sufficiently solve the problem of direction.

These difficulties have also become evident within the scope of distributing the social product.

Secondly: Based on the economic experiences of this era, the creation of the economic *orders* cannot be left to its own devices.

On the other hand, it would be rash to conclude from the experiences of this era that it is necessary to transfer the direction of daily *processes* of the economy to central bodies. It is the creation of the forms, of the framework and of the economic order as a whole, which clearly cannot be left entirely to the discretion of private entities.

In contrast to the policy of the laissez-faire, the central task of the economic policy is the creation of conditions within an industrialized economy which allow the development of functioning and humane economic orders.

In contrast to the policy of the laissez-faire, the central task of the economic policy is the creation of conditions within an industrialized economy which allow the development of functioning and humane economic orders.

* * *

Part Two: The Decision

CHAPTER 1: THE POSSIBILITIES

How can there even be an order for the economic process of the industrial world? What are the possibilities?

Their number is low. This was already evident from the historical overview. The deeper the analysis of reality, the clearer it becomes: there are economic orders dominated by a central administration of the economic process, and orders in which the plans and decisions of many businesses and households are decisive for directing the economic process. This means that either the economic system of a central administration, the "centrally administered economy," dominates; or the directional methods of the transaction economy are decisive for the order of the industrial economic process.[4] It is true that these methods of a transaction economy are totally different depending on whether or not the individual actors are combined into monopolistic, partly monopolistic or similar groups. One can—roughly speaking—distinguish between three methods which come into

4 For more information, see my "Grundlagen der Nationalökonomie," in particular part 3.

question for directing the industrial economic process: control by state central bodies, control by groups, or control by competition.

1. Those living in the 20th century have come to know the extraordinary disadvantages associated with a centrally administered economy. A form of order which the 19th century regarded as ideological has become reality. We must now draw the consequences from this experience. The concentration of economic power, its association with political power, the uncertainty and the insufficient supply of consumer goods, the increase in social dependence, the threat to the law-based state and freedom—we do not need to read about this in books; we experienced and experience it on an everyday basis. And there is another factor as well, namely the failure of the methods of a centrally administered economy in structuring international trade. Necessarily associated herewith are the transfer of exports and imports to central planning offices, which—in the absence of a precise costs calculation—introduce a strong element of arbitrariness and uncertainty into international trade and are not in a position to sufficiently control the finely differentiated world trade of the processing industries. The industrial economy encourages larger markets and the international division of labor; but economic orders of the central administration type were and are not in a position to provide these global economic relations with a firm foundation. In reality a conflict is breaking out. This also demonstrates a strong historical bias against the realization of a centrally administered economy.

2. Keynes said:[5] "I believe that in many cases the ideal size for the unit of control and organisation lies somewhere between the individual and the modern State. I suggest, therefore, that progress lies in the growth and the recognition of semi-autonomous bodies within the State-bodies whose criterion of action within their own field is solely the public good as they understand it, and from whose deliberations motives of private advantage are excluded, though some place it may still be necessary to leave, until the ambit of men's altruism grows wider, to the separate advantage of particular groups, classes, or faculties—bodies which in the ordinary course of affairs are mainly autonomous within their prescribed limitations, but are subject in the last resort to the sovereignty of the democracy expressed through Parliament."

This statement puts forward a specific form of the widespread theory that the direction of the economic process should be transferred to autonomous associations, professions and similar mandatory corporations. We do not want to discuss here whether such proposals revive medieval forms or not. However, it is surprising that Keynes even dared suggest such forms of order, which science has long recognized as only achieving a fragile equilibrium of the economy, tending towards disequilibrium. Experience has confirmed this on numerous occasions. If in the coal the mining, the iron, the cement or the potash industries, the trade or the workforce are combined in autonomous groups, group anarchy arises—

5 See Keynes, "The end of the Laissez-faire," 1926, p. 31 et seq.

with struggles between these groups, embargos, lockouts and strikes. Neither are the individual interests muted in such corporations; rather, group egoism tends to proliferate, as the groups possess power. These groups—even if they call themselves professions—have no intention of serving the common good. Their functionaries merely support the interests of their own groups, or that which they deem to be their interests, in the battle with other groups and with the state: whether American trade unions, international raw material cartels or national syndicates, and even if employee representatives are involved in the management of these groups. The direction of the processes of the economy by "professions" is not capable of reconciling own interests and common interests, it means group anarchy. The unbalance of market forms comprising two-sided monopolies or partial monopolies or oligopolies triggers a tendency towards centrally administered interventions. Think of the development of the German labor markets between the wars, where, in the battle between the groups of employers and employees, state mediation was increasingly required to establish labor conditions. A permanent solution to the problem of the economic order is not offered by coexisting or competing pressure groups. It is only possible to have sympathy for professions being the guiding force in the economy in a situation where the difficulty of the direction of the modern economic process and the character of economic power are not recognized.

3. This leaves the third type of economic order in which the market form of complete competition dominates.

This is a market form which was often partly realized in the industrialized economy; however, it was not universally realized and it lacked an adequate monetary order. Classical and—far more precisely—modern economic theory have shown[6] how a strict control of the economic process is fulfilled in a complete competition situation, and how the consumers direct the process.

This is how the principle of the competitive order arose. It emerged from daily life and academic experience.

4. The diversity of commercial phenomena and the magnitude of economic tendencies and cults shows that there are only very few possibilities for an economic order in which the modern economic process *can* be directed. In reality, this fact is fundamental. Every decision in economic policy should proceed from this starting point. This is an either-or situation. As the group-anarchistic, corporative or professions-based solutions to the problem of order-based policy can only exist temporarily, ultimately there is only a choice between a centrally administered control of essential parts of the economic process and the competitive order. It is time to examine this latter alternative.

6 The modern analysis of the process of control (Lenkungsprozess) in a completely competitive form of market has been considerably refined since the works of Menger, Walras, Jevons, Böhm-Bawerk, Marshall. The current status of research: G. J. Stigler, "The Theory of Price," 1947, p. 63-196. H. v. Stackelberg, "Grundlagen der theoretischen Volkswirtschaftslehre," 1948.

CHAPTER 2: WHAT IS THE COMPETITIVE ORDER?

1. If, for example, the cotton spinning works of a particular country have formed cartels, or if the market is dominated by a few independent yarn affiliates, or if numerous spinning works compete with each other for the sale of yarn to many dealers and weaving mills, this is important for price formation and for the direction of the economic process in the supply of yarn and beyond. Further questions that are obviously important for the general economic process are how banknotes or the deposit money of private banks are created, whether and to what extent credit is granted for this purpose, whether in a monopoly or in competition. Also: whether trade unions or employers' associations exist and how powerful these are. Production and distribution differ depending on how and in which forms supply and demand meet on the market and how prices and wages develop.

We take these everyday occurrences for granted. The economic policy of the competitive order aims to bring an order to all markets such that the overall economic process functions in an expedient way. The individual farmer, industrialist, craftsman and laborer, thus the individual business and household, should be able to both plan and act freely. They do not take orders but rather seek to apply their own labor force, their productive means and their money where they regard it most beneficial. Thus, households and businesses are not subordinate but rather coordinate among themselves. However, the framework of the economic process is not autonomously decided by businesses and households. The policy of competitive order does not leave the choice of market forms and monetary systems to the economy itself because the experience of the era of laissez-faire policy speaks for itself. The development of the framework in which businesses and households can plan and act freely is governed by the economic policy under which the framework is supervised. Businesses are free to choose what they produce, what technology they use, what raw materials they purchase and what markets they wish to sell on. Laborers are not obliged to serve in a particular form of employment either. They enjoy freedom of movement and the right to a free contract of employment. Freedom of the consumer exists, but not the freedom to choose how to define the rules of the game or the forms which the economic process takes. This particularly falls within the field of *Ordnungspolitik* (order-based policy).[7]

2. The market form that dominates in the competitive order is the market form of "complete competition." This coordinates the plans and decisions of the

7 E. F. Heckscher: "Der Merkantilismus" 1932 volume 1, p. 448 et seq. writes about the economic liberalism in England in the early 19th century as opposed to mercantilism:

> The old method would have been an attempt to create a barrier to fundamental changes (Umwälzungen). The new victorious method allowed them to take a free course. Therefore, they were enforced with power unparalleled in mankind's ancient economic history. The third alternative would have been neither to intervene in the course of events nor to regulate its course but to direct it in ordered ways. - This concept has never been tried.

individual businesses and households with one another. If this is not possible, particular measures of economic policy are required. In individual countries, the competitive order would be realized in a particular way depending on the actual or historical situation. So, for example, the competitive order in Germany will look entirely different to that in Belgium or in the United States—but this will be discussed later. What all "competitive orders" have in common is the fact that they are dominated by complete competition.

In addition to complete competition, the "personal economy" (simple, centrally administered economy) is supposed to, and will be, a widespread form of order; this might be the farmer who not only purchases seed, fertilizer, machinery, etc. on competitive markets and sells potatoes, pigs and vegetables on competitive markets, but at the same time also operates a personal economy himself by being a consumer of potatoes and meat. On the farm, the two forms of order merge together. They merge in a different way in the household of the metalworker who also owns an allotment where he grows potatoes, vegetables and fruit for his family. In view of the exceptional difficulties in bringing sufficient order to the modern economy based on the division of labor, it is important that economic policy should take care of these elements of personal economy. This ensures that people become less dependent on the market and have something to fall back on when times are hard. Overall, however, the personal economy can only be an ancillary form of order as it is not suited to direct the economic process based on the industrial division of labor. It is complete competition which characterizes the economic order.

3. What is complete competition? It is a particular, precisely definable market form and is not to be confused with laissez-faire. We know that laissez-faire has very often and increasingly led to monopolies, etc. —in short: to market forms outside complete competition. One need only think of the history of coalmining in Germany and elsewhere with its comprehensive formation of syndicates or of the labor markets of the 19th century when competition on both sides of the market seldom existed.

4. Complete competition is also entirely different to the "battle for a monopoly." A semi-monopolistic cement syndicate competing with outsiders living in its shadow, for example, is not complete competition, and neither is the competition between two oligopolistic shipping companies, two railway companies or two petroleum groups. In such cases, the aim is to achieve a monopoly. Unlike in complete competition, the costs are not a regulative factor, but instead prices are usually fixed much lower than costs in order to inflict damage and impose one's will upon one's opponents. In semi-monopolistic or oligopolistic opposition, restraints are often placed on the opponent's suppliers or purchasers. In complete competition, however, there can be no restraints. Oligopolists or monopolists of supply or demand apply market strategies that do not exist in complete competition. This difference is also essential for the evaluation of social questions: if a semi-monopolistic business on the demand side faces a large disordered workforce, then the remuneration structure will be completely different from that in complete competition. Complete competition has rightly been compared to a race. It is not a battle man-to-man but a race run in parallel. It is not impediment or injurious competition, but rather performance competition.[8]

5. Older criticism leveled at competition—by Sismondi, Marx, the St. Simonists and many others—has been largely devalued by the fact that they muddled the various market forms and referred to them jointly as competition. The description of the economy at that time tells that competitors waged personal vendettas against each other, sending each other into financial ruin in the process; that the competition of "rival capital rages" (Marx); that bigger capitalists beat smaller capitalists to death; that these wars are a senseless waste of assets; that workers are dependent upon an employer and that competition is ultimately a state of anarchy. The situations described actually often existed. Yet they were interpreted quite incorrectly and were described as effects of competition. What was described were battles for a monopoly and the dependence on monopolies and partial monopolies. More recent economic policy-related debate largely draws on such old misapprehensions. It is in this way that the destruction of wheat or coffee stocks comes to be attributed to competition whilst, in reality, it can only occur in monopolistic market forms.

Science has since developed an economic morphology. Insofar as this morphology aims to understand the real economy, and describing the forms found therein, it is also able to describe precisely what competition is. (Yet a science that establishes certain models axiomatically, failing to seek and find forms that exist in the real world is unsuitable for this purpose.)[9]

6. A farmer's economic planning does not take account of how his wheat sales affect the wheat market since his supply is too small in relative terms. His plans are therefore based on wheat or pig or vegetable prices that he regards as fixed values—as data—and these are prices formed on the market. He may expect certain market prices, estimate the prices he believes he will be able to charge or assume that these prices will move within particular parameters. This is competition.

Competition exists between the housewives of a town who buy food, textiles and shoes, or between lessors offering residential premises for rent. Competition has often been realized in agriculture, trade or industry, primarily in the process-

8 Regarding the problems of the oligopolistic and partly monopolistic battle or the battle of monopolies against each other: Franz Böhm, "Wettbewerb und Monopolkampf," 1933; Kestner-Lehnich, "Der Organisationszwang," 2nd edition 1927; Hans Möller, "Kalkulation, Absatzpolitik und Preisbildung," 1941.

9 About the newer market form: H. v. Stackelberg, l.c. and L. Miksch, "Wettbewerb als Aufgabe," 2nd edition 1947; as well as my "Grundlagen."

ing industries, e.g. the paper processing industry, in many branches of machine construction and the textile industry, to mention but a few.

If there is competition on the supply side, as well as on the demand side and if the economic planning of both sides is based on such competition, then the market form of complete competition is achieved. This is true, therefore, where there is competition not only for the supply of wheat, but also for the demand of wheat, or when a town's grocers are competing in the same way as its housewives. Anyone with industrial experience knows that this market form was frequently realized in the 19th and the early 20th centuries, not just in trade and agriculture but also in industry.

7. Yet this description of the market form of complete competition is not quite sufficient for economic policy. Are there methods that can be used to enable administrative practice to recognize complete competition and other market forms? Economic policy needs indications, *symptoms*, by which to implement an economic policy of competitive order. It needs a rule of thumb. Does such a thing exist? The answer is yes. Two methods exist.

> IF THERE IS COMPETITION ON THE SUPPLY SIDE, AS WELL AS ON THE DEMAND SIDE AND IF THE ECONOMIC PLANNING OF BOTH SIDES IS BASED ON SUCH COMPETITION, THEN THE MARKET FORM OF COMPLETE COMPETITION IS ACHIEVED.

The most direct method is to find out from companies themselves whether or not plans developed under competition. For example, if a company planning and building particular machines expects these machines to sell on the market for 500 marks each because they trade on the market at this price, then it can be assumed that competition exists. The price is not forced upon the market by way of a market strategy, but is taken from the market. Such assumptions are supported by the fact that there are no cartel agreements on the market or by information regarding competitors and the size of the market.

A second, less direct, method would be as follows: certain measures evident from outside, for example from the opposite side of the market, clearly indicate that complete competition does not exist because these measures cannot be implemented under complete competition: for example, restraints to purchasers or suppliers that have dealings with competitors, or loyalty rebates or predatory pricing or dumping or destruction of stocks. A further example: the price of raw silk has fallen sharply, yet the company does not reduce the price of its silk products. Here, the company cannot be under complete competition as the price mechanics of complete competition are such that the price of the product will fall if the price of the raw material falls. The company probably has a partial monopoly. Such indirect methods of determining the market form from outside are relatively simple to apply and powerful.

8. In the competitive order, complete competition serves not only to increase performance: it is the very form of the market whose prices direct the economic process. Competition is used to increase performance even in the various types of centrally administered economic order. Here, contests are held between individual businesses which are presented with awards; there are also contests held between the employees who are awarded for special performance. Competition is used as a means to increase performance whilst the economic process is directed by central planning offices.

In the competitive order, however, the economic process is directed by the prices of complete competition and by the plans of the many households and businesses complying with these prices. In the competitive order, the direction of the economic process *and* the increase in performance together should be ensured by complete competition.

9. One could also attempt to determine the competitive order from an entirely different perspective: by comparison with a law-based state. In the same way as a law-based state, so the competitive order should also create a framework in which the individual's freedom to act is limited by the freedom of others, thereby ensuring a balance between every unit of human freedom. Indeed, the desire for competitive order is closely related to the desire for freedom.

But the desire to see these realized is not enough—in the same way as affirming and simply hoping for the creation of a law-based state will not summon it into existence. A house has to be built and its plan has to be designed.

Part Three: The Realization

* * *

CHAPTER 1: THE ESTABLISHMENT OF THE COMPETITIVE ORDER

I. Basic Principle

1. We know that during the era of laissez-faire and the subsequent epochs of experimentation, economic policy underestimated or failed to see the importance and difficulty of the problem of exerting adequate direction on the economic process, a problem which entered an entirely new stage with the coming of industrialization. Above all, the fact that all economic circumstances are related was and still is neglected, i.e. that the directing mechanism must be indivisible if it is to function. Because of their general interdependence, every single economic policy intervention affects the economic process as a whole. Should interest be manipulated to a lower rate, for example, then this alters the entire price system and thus the entire direction of the economic process, as has often already been proven the case.

The lively economic debate on the influencing of the economy shrouds the fact that the actual problem of exerting this influence was given less and less attention over the decades. It is true that market forms and monetary systems within the framework of which prices were created were already permitted during the era of laissez-faire, only inadequately fulfilling the function of exerting a guiding influence. Later, however, during the era of economic experimentation, the lack of concern was even greater. Thinking in terms of price relations waned rapidly. Irrespective of whether the full-employment policy partially crippled the price system through the fixation of prices, the low-interest rate policy, the rationing of foreign currency, or whether the central planning agencies attempted to direct the economic process, which was only possible on the basis of global valuations, the incredibly difficult influential direction of the large-scale industrialized economic process and its overall context were always underestimated.

2. At last, this should change, and the crucial issue of modern economic policy should be treated as a crucial issue. This is to be achieved by making the establishment of a functioning price system of complete competition the essential criterion of every economic measure. *This is the basic principle of the economic constitution.*

The aim is not to pursue a policy to stabilize the business cycle which hinders or cripples the functionality of the price system under the impression of a momentary state of emergency, as with foreign currency control, credit expansion and the like. Likewise, tax policy may not promote the concentration process and thus encourage the advance of monopolies, e.g. through turnover tax or the structuring of corporate income tax. Here, and in all branches of the economic policy, the basic principle of the economic constitution should be present in all measures undertaken. There can be no exception to this approach. The modern economy is a large, interrelated system. All economic-political acts thus influence the overall process and must therefore all be coordinated with each other.

3. The basic principle does not only demand the avoidance of certain acts of economic policy: such as state subsidies, the creation of mandatory state monopolies, a general price stop, import bans, etc. Nor does it suffice to ban cartels, for example. The principle is not primarily a negative principle. Rather, a positive economic constitutional policy is required and its aim must be to further the development of the market form of complete competition and thus comply with the basic principle. Also in this respect, the policy of the competitive order differs entirely from the policy of laissez-faire, the fundamental concept of which did not include a positive commercial order-based policy.

Thus, what now needs to be developed are the individual positive principles which constitute the competitive order and which have their common focus in the aforesaid basic principle.[10]

10 Regarding this huge problem there exist three works which are particularly significant: F. Böhm, "Wettbewerb und Monopolkampf," 1933; L. Miksch, "Wettbewerb als Aufgabe," 2nd edition 1947 and Henry C. Simons, "Economic Policy for a Free Society," 1948.

* * *

III. Open Markets

1. The closure of supply and demand is a method that was and is used most to break or hinder the strong tendency to compete, which is especially effective in the industrial economy. The modern state as well as private and semi-public pressure groups have applied an exceptionally broad range of instruments to close supply or demand. Import bans or prohibitive customs duties or foreign trade monopolies isolate the suppliers of a country from foreign competitors, i.e. procure a local closure of supply. Investment bans, restrictions on cultivation, construction bans and the like have a similar effect. Entry bans and the prohibition of migration, the hindrance of a free career choice, license systems with demand tests for commerce, trade, industry and the creation of a numerus clausus can be included amongst this; likewise the prohibition on simultaneously pursuing various trades.

Thus far, we have talked of closure measures undertaken by the state. However, private pressure groups and monopolistic enterprises have also developed a system to hinder competitors or to prevent their emergence. The stoppage of materials, workers, supply and sales channels, the binding of customers through exclusive agreements and loyalty rebates, credit freezes and predatory pricing all prevent the emergence of competitors and close supply.

In addition to this are the unique closure measures involved when the state and private powers cooperate, i.e. where the state provides assistance to enable private parties to create a dam to close supply. Examples of this were patent law and resale price maintenance for branded products.

During the Middle Ages and mercantilism, very effective methods were applied to prevent the influx of people and capital into an industry. During the era of industrialization, similar methods were developed with no less success.[11]

2. If one enterprise is granted the exclusive privilege to conduct a specific trade, such as the post office which is authorized to convey letters or the central bank which is authorized to issue notes, then the direct consequence of the closure is the creation of a monopoly. Thus, from the closure automatically follows the elimination of all non-monopolistic market forms. However, it is frequently the case that supply is closed in favor of a multitude of suppliers and not purely for a single entrepreneur. This is the case, for example, with general restrictions of admission to retail trade or with investment prohibitions for entire industries, or with cultivation restrictions for sugar beet and other agricultural products. Is the closure of the markets perhaps reconcilable with the competitive order in this case? Is it not possible for competition to arise within the framework of

11 See particularly Böhm l.c., p. 75; Kestner-Lehnich l.c., p. 53 et seq.

closed markets, as was indeed the case, for example, with the supply of tobacco in Germany during the 30s? Is the closure of markets really irreconcilable with complete competition and does the implementation of complete competition really presuppose the need to keep all markets open to the extent possible?

In response, one can indeed say that the competition mechanism can be effective within the framework of closed markets. However, economic policy must nevertheless apply the principle of opening the markets because their closure would entail the acute risk of obstructing complete competition. There are two aspects which bring about this situation.

Firstly, the closure of supply and demand makes it considerably easier for monopolies to establish themselves. If an investment prohibition exists for cement works or if admission to a trade is closed, the likelihood of monopolies developing is extremely high. If import bans or investment prohibitions define the supply of steel, then steel cartels will have a firm foundation. Vice versa: where there is no market closure, where import bans, general restrictions of admission, investment bans and the like have been lifted, it is often difficult to maintain monopolies, and oligopolies revert to competition. For this reason alone, the opening of supply and demand is a precondition for the constitution of the competitive order in the industrial era.

No less important is a *second* aspect: even if complete competition is established on individual closed markets, the connection between the markets is disrupted as a result of the closure and the entire system of complete competition may not fully function. For example, if a state imposes an investment prohibition on machine tool manufacturing factories, then although it is possible for competition to exist between these factories, the investment prohibition prevents the volume of capital, and thus labor, iron and other materials, from being brought into the machine tool manufacturing sector that would have been possible on the basis of the price relations. Consequently, machine tools are more scarce than would have been the case had the investment prohibition been lifted. The general equilibrium, i.e. the coordination of the many markets and industries with each other, cannot be fully achieved if investment prohibitions and other closure measures are in force. The functions of control and selection which, in the case of open and complete competition, are assumed by prices on the one hand and—through prices—by consumers on the other, are partially assumed by the offices which impose the closure. In addition to this, the income generated in closed trades is akin to rents, which would be washed away with the opening of the markets.

Thus, the following principle applies: The constitution of the competitive order requires the opening of supply and demand. Very few exceptions hereto exist, such as the exclusive right granted to the central bank to issue notes.

3. The state may not confine itself to allowing anyone to carry on a trade, to lifting investment bans, general restrictions on admission, privileges, compulsory rights and privileged protection rights, to creating freedom of trade and freedom of movement, to avoiding state import bans; thus to generally refraining from undertaking state closure measures, so that the price system of complete competition can determine the selection process. Rather, it is necessary that the markets are not closed by private pressure groups either. What is the point of state-protected freedom of trade if this is actually lifted by the policy of the pressure groups? What is freedom of trade if an aluminum rolling mill cannot be established because the existing syndicate takes active measures to prevent this? Any form of "impediment competition," i.e. blockages of any and all kinds, loyalty rebates, exclusivity agreements and predatory pricing against competitors which aim to destroy or deter competitors must be prohibited.

The opening of markets has an economic constitutional purpose. For this reason, private pressure groups cannot be granted the right to eliminate them. They are a part of *Ordnungspolitik* (order-based policy) which may not be left in the hands of private individuals. Here as well, the difference to the policy of laissez-faire is evident, where private pressure groups were not only entitled to form, but also to close their markets by undertaking active measures.

4. In the individual areas of the economy, the enforcement of the principle raises significant, difficult questions in each case. For a more precise description of the situation, may I give two examples, namely customs and patent policy.

How are *protective customs duties* to be viewed pursuant to this principle, i.e. duties which do not act prohibitively and thus do not equate to import bans? Such customs duties do not close supply. This applies to the majority of duties that existed prior to 1914 within the framework of the so-called central European trade treaty system.

Such duties do not directly destroy the competitive order. Their effect is similar to that of an increase in the distances between countries. They shift the price relations, but do not make it impossible for the price system of complete competition to control the economy. Thus far, duties are reconcilable with the competitive order. And, moreover, the transition from the system of import bans or import licenses to the system of customs duties is a step in the direction of the competitive order.

Indirectly, however, customs duties can still be dangerous to the competitive order, namely where they facilitate the establishment of monopolies. It is a known fact that they can increase the propensity of the customs-protected industry of a country to form cartels, by making it possible to segregate the country as a sales territory from the global market and to control it monopolistically. And the incentive to establish monopolies especially exists in industries that are able to supply the internal market by themselves, where customs duties thus only gain any effect whatsoever through the establishment of cartels. The history of the German iron industry offers several examples of this.[12] In these cases, the phase-out of customs duties can be a means to eliminate the trend towards monopolies, thus to procure the establishment of the competitive order.

5. Forms of the economy which are irreconcilable with the competitive order, i.e. forms of the economy that are alien to the system, often arose in connection with modern *patent law*. Patent law also belongs to the multitude of more recent legal institutions which did not have the effects desired by the legislator. Patent law was intended to promote technical development as well as to protect and reward the inventor. The extent to which these goals have been achieved can remain undecided.

Contrary to expectations, despite certain statutory precautionary measures, patent law has triggered a strong trend towards the establishment of monopolies and concentrations in the industry. This is due to the fact that patents create an exclusive right to manufacture an object, to bring it onto the market, use it and sell it. Although many of the patents do not close supply, these are patents which only cover a minor part of the production process of a commodity and which can be circumvented by using other production methods, substitute products and the like. However, a different type of patent group exists, namely master patents, which close the supply of goods, such as inter alia the well-known Telefunken patents for the production of radio tubes, or the benzopurpurin patent of 1884, which became an important feature for the organization of the chemical industry.

Forms of the economy which are irreconcilable with the competitive order, i.e. forms of the economy that are alien to the system, often arose in connection with modern *patent law*.

The closure of supply through the application of patents has encouraged concentrations in two ways. A patent can grant individual firms an individual monopoly, as for example is often the case in the fine mechanics industry. Secondly, patents have also triggered or reinforced the establishment of cartels or groups. And this effect was more important. And this was not only the case for the actual patent cartels, patent trusts or patent pools. The exchange of licenses facilitates the establishment of cartels; the risk run by a member of a cartel in the event of his withdrawal, namely that he loses his right to certain patents, is what cements many cartels together. Patents also gained what can only be described as a decisive role in the establishment of modern-day corporate groups, namely in terms of their expansion and the struggle to keep out competitors. “Rather, the driving force of patents also has to be sought and appreciated where concentration forms develop which allow none of the patent power struggles or underlying license agreements to reach the public, and nevertheless

12 See, e.g., Wiedenfeld “Gewerbepolitik,” 1928, p. 148; Haberler, “Der internationale Handel” passim.

have these elements to thank for their creation, form and direction to a very considerable degree. They cannot be quantified, but only manifest themselves as inherent tendencies and latent possibilities" (G. Gather). On the whole, the establishment of the modern economy with monopolies or oligopolies has been decisively defined by patents, trademark protection, the resale price maintenance associated therewith and by leading advertising. Take, for example, the establishment of corporate groups in the chemical, detergent and cigarette industries. Through jurisprudence, the conditions for the implementation and acceleration of the concentration process were considerably encouraged and, simultaneously, the forces which pushed for complete competition were suppressed or invalidated.

6. A patent policy which draws the conclusions from these experiences is—in conjunction with the analogue treatment of trademark protection, resale price maintenance and leading advertising—an important element of economic policy aimed at realizing the competitive order. Its aim is to restrict or eliminate the closure of supply occurring as a result of the granting of the patent. How is this possible? Can one adhere to the fundamental concept of today's patent law, i.e. to the granting of the exclusive right to exploit the invention? In this case one would have to substantially relax the closure of supply arising from the present legislation and jurisprudence of the civilized countries. There are numerous proposals to achieve this: for example a shortening of the protection period and the expansion of compulsory licenses.

Whether or not it suffices to relax the exclusivity rights can remain undecided. Perhaps the granting of such rights and the closure of supply should be dispensed with entirely, introducing in their stead a system pursuant to which the patent owner is obliged to grant the use of the invention to all seriously interested parties in return for a reasonable license fee. As with all other monopolies, an obligation to contract would also exist in patent monopolies and the contractual conditions would have to be stipulated by the patent office in the event the parties could reach no agreement. Numerous senior patent policy proposals could continue to develop in this direction.[13]

* * *

13 On the patent issue, I owe numerous suggestions to two works by G. Gather, "Patente-Monopole-Machtpositionen," 1943 (unprinted) and "Reform der Patentgesetzgebung?" in this *ORDO* volume and also to the mentioned works of Franz Böhm, L. Miksch, and H.C. Simons, "Modern American literature," in particular about patent cartels, in F. Haussmann, "Der Wandel des internationalen Kartellbegriffes," 1947, p. 33; essays by H. Rasch and Tschierschky in the Kartell-Rundschau; W. Röpke, "Die Gesellschaftskrisis der Gegenwart," 1942, p. 362 and 389 et seq. and the literature mentioned therein; E. Liefmann-Keil, "Monopolbekämpfung durch spezielle Marktstrategie," Schmollers Jahrbuch, 67th volume, 1943.

CHAPTER 2: THE POLITICS OF THE ECONOMIC ORDER WITHIN THE COMPETITIVE ORDER—THE REGULATING PRINCIPLES

Strict compliance with the constitutive principles cannot prevent specific competitive orders containing certain forms of order that are alien to the system (I). Furthermore, and no less important: even if complete competition is realized, it contains weaknesses and defects which require correction (II - V).[14]

Therefore, certain "regulatory" principles are necessary, which must be applied in order to keep the competitive order functioning.

I. The Problem of Monopolies in the Competitive Order

1. Economic power should only exist in a competitive order to the extent necessary to maintain the competitive order. The management of households and businesses requires economic power in order to implement the envisaged economic plans. It is admittedly the case that in the competitive order, such economic power is subjected to necessary, strict, daily control by the price mechanism. However, a central bank which has the exclusive privilege to issue bank notes also exerts economic power, which gives rise to the difficult problem of its control. But also this power is created for the purpose of enabling the competitive order, by creating a sufficient monetary order.

2. Even in the competitive order, however, there will be monopolies which do not serve the maintenance of this order, but serve as a disrupting and threatening element. Certain positions of power arise even if these principles are completely applied. Thus, for example, a gas supplier has a monopoly in the surrounding city, i.e. on its market. Or a railway has a position of monopoly in its transport area. Or a factory producing precision scales or medicinal equipment or slide rules dominates its market monopolistically or partly monopolistically. Such monopolistic positions are established on the basis of genuine cost advantages. In these cases the optimum operating size is so important that the quantity produced by only one business is sufficient to supply the market. Several businesses would only be able to sell at prices insufficient to cover their costs.

Here the question arises: What is to happen with these monopolies? All ways of directly and indirectly preventing the creation of a monopoly have been exhausted. Despite this there are powerful entities in existence which endanger the entire order.

3. The question is not identical to the familiar question of monopoly (or cartel) supervision, which many industrial states e.g. Germany with its Cartel Regulation of 1923, have attempted and failed to solve.[15] Trade policies, compa-

14 See p. 32 above.

15 Refer to F. Hausmann's "Die wirtschaftliche Konzentration an ihrer Schicksalswende," 1930, p. 231 et seq. and H. Kronstein's "Cartel control," A. Record of Failure, Yale Law Journal, February 1946.

ny law, tax policies, cartel law and cartel case law, as well as all commercial law, there and then facilitated and promoted the creation of monopolies. To some extent there were even mandatory syndicates. The state then attempted to prune the strong wild shoots sprouting from the trees whose growth they had promoted, but the trees merely continued to flourish together with their wild shoots.

Experience has shown that a modern state is unable to establish an effective supervision of monopolies in an economic order in which large parts of the industry are monopolized. Here the political influence of the interest groups is too strong and the monopoly problems too manifest. Even if individual civil servants of the ministries attempt positive initiatives, the support they receive within the structure of the state is too weak, as precisely shown by the extensive German experience. One should not be under any illusions with respect to the efficiency of monopoly control in economic orders in which monopolies proliferate in the industry or in agriculture or among the workers. The United States and its monopoly policies also provide an example.

The situation is totally different in the competitive order. Here the main impact leads to another direction. The creation of monopolistic power entities is prevented. Not only by prohibitions of cartels, but also—and far more importantly—by an economic and legal policy which breaks through the strong forces of competition, as exist in a modern economy, by applying the constitutive principles. In this way the state largely escapes the influence of private pressure groups. Its ability to supervise monopolies is disproportionately greater if the leaders of the coal, potash, iron syndicates, etc., of the large groups and trusts and trade unions, do not have a right to take part in the decision-making process. At the same time the task is far more moderate. Only for the aforementioned unavoidable monopolies is the problem of monopoly supervision topical. The chance of its success is thus far greater.

4. Nevertheless, it is necessary and also possible to draw conclusions from the experiences of industrial states with their monopoly supervision. The first conclusion is that two methods of monopoly control often fail.

Firstly: the nationalization of monopolies does not solve the monopoly problem. State monopolies e.g. the railway or electric utilities often practice just as monopolist policies as private monopolies. In the same way as private enterprises, they seek to achieve the point at which the highest profit can be generated, which in a monopoly will usually substantially deviate from the point of optimum satisfaction of demand. In many cases, even the tendency to fully exploit the position of monopoly is greater among state monopoly administration bodies than among private enterprises. The state monopoly administration considers itself justified in such behavior as the income flows to the state or the town, i.e. represent an indirect tax and are not used for private purposes. Apart from this, the state feels far more secure with respect to potential competition; it can, e.g., restrict the supply of any emerging substitute product by legislative means—as occurred in Germany with the emergence of the automobile sector as competition for the state railway.

Nationalization has the effect of merging economic and political power. It concentrates the two spheres of economics and politics.

However, the problem of economic power and abuse of power has never, anywhere, been solved by concentration. A concentration of power on the one side increases dependency on the other side: of workers, purchasers and suppliers. Nationalization of the large monopolies, e.g. of the heavy industry, does not mean that the power of the interested parties is effectively subjected to supervision, but that the supervising parties become interested parties.

Secondly: Such risks and other influences have given rise to the ideal of mobilizing the functionaries of the workforce for purposes of monopoly supervision. What is interesting in this connection is, e.g., the German Coal Economy Law of 1919, which appeared at the same time as the Socialization Act.[16] By delegating functionaries of the coal workers and the employees to the management of the mandatory coal mining syndicate as well as the coal association of the German Reich and the coal council of the German Reich, the State wanted to emphasize the public interest in the monopolistic bodies.

The attempt failed, for a reason which is of fundamental importance. If the employees participate in the profits of a monopoly, they have just as strong an interest in the monopoly and in the monopolistic policies as the entrepreneur. The workers of a coal mine often agreed to demands for higher prices for the monopoly products if wage increases were promised in the event of price increases. Increases in railway prices are often approved by the railway unions. The recipients of wages and salaries have far greater interests than the fragmented interests of consumers.

This is where the often friendly, effective attitude of the unions vis-à-vis the cartels has its roots. If the functionaries of the workforce participate in the leadership of monopolies, this gives the monopolies a broader foundation. As a result, the workers unite with the entrepreneurs to form *one* monopoly group. It was illusory to hope that this unification would support the overall interests. Now entrepreneurs and worker functionaries collaborated to support the monopolistic politics of their industry. This is not a case of bringing a counterweight to oppose the weightiness of the monopoly leaders, but a strengthening of the weightiness of the monopoly.

Neither the nationalization of monopolies nor the control by the workforce can solve the monopoly problem in the competitive order.

16 Hereunto: report by the Commission for Socialisation and report on its action, Berlin 1920.

5. The monopoly supervision should thus be transferred to a state monopoly office. In order to withdraw it from the persistently dangerous influences exerted by interested parties (although these are weakened in the competitive order), this should be an independent office which is only subject to the law. It should not be, for example, a department of the economic ministry, which is far more subject to the pressure of interested parties.

This monopoly office is exclusively responsible for all questions of monopoly supervision. It is also responsible for deciding whether or not the precondition of a dominant economic position is fulfilled in the specific case.[17] What is needed is a new central office, something which is currently lacking, and the establishment of which is both necessary and accomplishable. Such a large monopoly office would be a central figure in the modern, industrialized state. Without such an office the competitive order and with it the modern law-based state is threatened. The monopoly office is just as indispensable as the Supreme Court.

THE AIM OF MONOPOLY LEGISLATION AND MONOPOLY SUPERVISION IS TO ENSURE THAT THE BEARERS OF ECONOMIC POWER BEHAVE AS IF COMPLETE COMPETITION PREVAILED. THE BEHAVIOR OF THE MONOPOLISTS SHOULD BE "ANALOGOUS TO COMPETITION." THIS IS THE PRINCIPLE WHICH ARISES FROM THE BASIC PRINCIPLE OF THE COMPETITIVE ORDER.

A monopoly office has the task of dissolving avoidable monopolies and supervising unavoidable ones. Admittedly, the number of such monopolies will be relatively low in the competitive order. Part monopolies and two-sided monopolies would be just as liable to its supervision as supply and demand monopolies. If, therefore, a large factory dominates the supply of the country with spiral springs, as part monopolist, and if there are many further smaller businesses supplying the same product, the large supplier cannot escape monopoly supervision by reference to smaller competitors.

6. The aim of monopoly legislation and monopoly supervision is to ensure that the bearers of economic power behave as if complete competition prevailed. The behavior of the monopolists should be "analogous to competition." This is the principle which arises from the basic principle of the competitive order.[18] Its implementation is to extend to the following aspects:

a) As is generally known, the general terms and conditions of associations, the industry, banks and insurance companies, as well as individual companies such as gas and electric utilities, railways, etc., often exclude state law. The law created

17 See p. 26 et seq. above.

18 See p. 32 et seq. above.

by the economy itself have ousted state law, especially from the monopolized areas of the economy.[19]

In contrast, the aim is to create the situation which would arise in the event of complete competition. General business usances which arise to supplement statutory provisions on the markets are admissible, but not general terms and conditions which deviate from the statutory rules to the detriment of the contractual partner. This already eliminates a serious defect, which is seriously neglected in legal policies and which was precisely perpetrated by the monopolies.

b) Every form of impediment competition by embargos, loyalty rebates, predatory pricing, etc., is prohibited. The monopoly office must also supervise these aspects. This creates a condition which would automatically arise in a complete competition situation, where impediment competition would be pointless.[20]

19 Franz Böhm, "Ordnung der Wirtschaft," p. 157, writes:

> Thus the consequence of the development of terms and conditions of associations was, inter alia, an unprecedented legal disintegration in the areas of civil law and commercial law, a legal disintegration which today is much more significant than at the time of the German Federation or after the foundation of the Reich before the enactment of the German Civil Law Code. For purchase agreements which a single trader concludes with suppliers of different lines of business today, state law only applies to a diminishing extent, but the most colourful 'laws' conceivable of countless associations and influential one-man businesses therefore do apply. The most precarious aspect however is the contents of such 'laws'. Whilst at the time of political disintegration, the provisions of the purchasing and credit law pursuant to the German Allgemeines Landrecht or General State Law, the Napoleonic Code and the other state laws were dictated throughout by the endeavour to grant a fair balance of interests (transfer of perils, liabilities for warranties, consequences of default, the impossibility of performance, etc.), among writers of terms and conditions of market associations, etc., the tendency prevails to amend the rights and liabilities unilaterally in favour of the one market party. Thus not right but wrong is created. The agreements regarding an arbitration tribunal were often also made in order to avoid an unwanted interpretation of the law of standard forms (Formularrecht) by national courts. Furthermore, the extraordinary spread of arbitration has the effect that the state judiciary, to an ever greater extent, has been pushed away from facts relating to cartel and market law. Moreover, the implementation of arbitrational procedures for breaches of cartel obligations frequently took the form of criminal law processes: under the form of damages and contractual penalties processes, in reality, a private criminal justice process takes place.

See also: L. Raiser, "Das Recht der Allgemeinen Geschäftsbedingungen," 1935.

20 The task which exists here has been described by Großmann-Doerth as follows:

> The present: On the one hand, this is the national legal law of obligation contracts, substantially the result of the 19th century jurisprudence based on ancient literary tradition, today more and more separated from economic life, therefore meagre, often mummy-like, a running on the spot. - And on the other hand it is the General Terms and Conditions: it dominates, rather than the state law and often in contrast thereto, the life of the obligation contracts. . . . It is necessary that the General Terms and Conditions are finally recognised as almost the most important civil law-related political task given to us. Finally, the conclusion has to be drawn therefore that for some

Admittedly, in order to achieve a result analogous to competition, it is necessary to introduce an obligation to contract, as here coercion is necessary to achieve the same result as would automatically arise under complete competition.

c) As is generally known, under complete competition the same prices will become established for the same goods and services. Supply monopolists, for example, whilst striving for the highest profit, have a tendency to demand differentiated prices for the same goods or services from individual segments of demand. This price differentiation should be prohibited in the competitive order.

d) What is most difficult is to implement the fundamental principle within the scope of determining price levels. The price is to be fixed in such a way that offer and demand are in equilibrium at this price, and, at the same time, the marginal costs are just covered. A chemical plant with a monopoly for a particular medicine must sell its product at a price with two characteristics: the price must be such that offer and demand are in equilibrium; consequently no rationing is necessary. At the same time, the price—e.g. 3 Mark per unit—must equal the costs of the last unit produced. Difficulties arise within the scope of determining the costs. This is because experience has shown that the information provided by the management of the enterprise concerning the costs of production is imprecise and requires stringent all-round examination.[21]

e) This is not the only aspect to the monopoly office's price controls. The office not only has to ensure that the most favorable point for supplying the market is achieved with the *given production apparatus*. Under complete competition, an ongoing, long-term pressure to rationalize the production apparatus arises. It is necessary to carry out a price control of the monopolies which also expresses a long-term pressure analogous to a competition situation.

The monopolistic chemical plant which is subject to the control of the monopoly office will not only have to adjust its general terms and conditions to complete competition, and not only have to refrain from embargos and price differentiation, and not only fix its prices at a level at which they are in equilibrium and cover marginal costs, but it is also subject to a long-term pressure to reduce the costs and prices of the products by rationalization. Otherwise the production apparatus will become antiquated, as often occurs with monopolies, and the supply of goods will not be optimal. Therefore, the plant must expect its

footnote 20 cont'd

> time it has no longer reflected our feeling for right and wrong when economy and state establish right against each other. . . . It should at last be naturally understood that the one who establishes right is subject to the legislator's responsibility vis-à-vis the entirety, whether right is established by government bodies or by the economy itself.

("Die Rechtsfolgen vertragswidriger Andienung," 1934, p. 201 et seq.).

21 S. L. Miksch l.c., p. 91 et seq.

prices to be revised by the monopoly office from time to time. Under certain circumstances, if possible improvements have been omitted, it must expect a reduction to below the point discussed in d), and thus losses. The monopoly office should refrain from regarding the existing production apparatus as a given variable on a long-term basis.

The questions of monopoly control are manifold and difficult. However, they can be solved *if* the other constitutive and regulating principles of the competitive order are followed, i.e. if the creation of monopolies is restricted to a minimum and if the monopoly control is dealt with under the simple and effective principle of competition analogy.

7. Combating and supervising monopolies also has a prophylactic effect. The otherwise abundant objective of attaining a monopolistic position, which—as has been seen—is a central aspect of economic history, is considerably weakened or eliminated if such decisive monopoly supervision becomes effective.

A further problem in this connection is the issue of how oligopolies are to be dealt with. This includes cases such as the following: Three companies of the electricity industry supply certain electrical machines, or two petrol groups dominate one market, or five aluminum smelters supply the market—without creating a cartel. Or—which is a very common occurrence—part oligopolies arise. For example, two large oven factories dominate the market, but there are also many smaller businesses in this sector which adjust their pricing policies to the larger companies. This oligopoly—or part oligopoly—situation often passes by rapidly, and soon leads to the creation of a cartel, i.e. to a collective monopoly or an individual monopoly, by overpowering the opponent. Sometimes, however, the unstable condition of the oligopoly or part oligopoly exists for many years or decades. How should this situation be dealt with from the perspective of economic policy? It is true that with the general policies accompanying the competitive order—e.g. its trade policies, patents, protection of registered designs, tax policies, etc.—the number of oligopolies becomes far smaller with strongly expanded markets. However, it is still important to ask how one should handle a cartel which only has a few members and after the dissolution of which the market will become oligopolistic?

There are two views: According to the first opinion, as has been put forward particularly impressively by Miksch, a special regulation is necessary for oligopolies and part oligopolies: namely the “tied competition” which takes place under state supervision. According to the other view, this is too great a burden for the state. According to this view, an active monopoly supervision is indeed sufficient, and it also offers something better for such cases. With a decisive monopoly supervision, the oligopolists have no reason to destroy the others by aggressive means or to attain a position of monopoly of their own. This is because it comes up against a rigorous monopoly control. Furthermore, the oligopolists themselves will attempt to behave as if complete competition prevailed, as they will otherwise come to the individual attention of the monopoly office. An example: a cement cartel is dissolved. As a result, the seven members become oligopolists. That one company now seeks to overpower the others is unlikely. This is because all measures of impediment competition—predatory pricing, blockades, loyalty rebates etc.—are prohibited and punishable. If, however, it becomes a monopolist by using competitive means, it would be subject to the comprehensive, deterring supervision of the monopoly office. What about if the seven companies remain oligopolists? In this event they will behave in almost the same way as in the competition situation. In some respects they will have to behave in this way: They too are subject to the legal provisions on general terms and conditions, on impediment competition and price differentiation. And if they do not approach the prices of the competition, they will have to expect the intervention of the monopoly office on a daily basis.[22]

* * *

For the complete translation, visit cpi.esapience.org.

22 On monopoly control: the cited works by Böhm, Miksch, Haussmann, Röpke, Kestner and others. Latterly: B. Pfister, "Leistungswettbewerb und Monopolkontrolle," Hochland, volume 40, p. 558 et seq.