VOLUME 4 | NUMBER 2 | AUTUMN 2008

Competition Policy International

Edited by David Evans

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

From the Editor

David Evans

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

From the Editor

Over 90 percent of antitrust litigation in the United States is filed by private plaintiffs, sometimes as class actions, and always seeking treble damages. As Judge Douglas Ginsburg and Leah Brannon have observed in these pages, such cases have been the source of most of the fodder that the courts have used to develop the precedents that constitute antitrust U.S. law. The situation has been far different in the European Community where the rights of private action have been limited, class actions largely unknown, and multiple damages uncommon. In the last few years several of the EC countries have embraced or considered adopting some aspects of private rights of action for violations of the competition laws, especially for the recovery of damages. Most recently the European Commission issued its White Paper on antitrust damage actions. The move in the EC towards private litigation for violation of the antitrust laws, and the issues and challenges this presents, begins the Autumn 2008 issue of CPI. The symposium consists of articles by authors from several different parts of the competition policy community: Christopher Cook, Vincent Smith, Assimakis Komninos, and Renato Nazzini & Ali Nikpay.

Thus far, private actions in Europe have mainly followed from the findings of a competition authority that a company participated in a cartel. It is thus fitting that we turn to two articles concerning collusion. The first, by Malcolm Coate, reports the result of an empirical study of the role of collusion in the FTC's merger reviews. The second, by Stephen Davies & Matthew Olczak, considers the conditions for overt and tacit collusion and uses both empirical and experimental evidence to address whether one theory fits all.

The issue concludes with a collection of papers that, roughly speaking, debate how the relative roles of static and dynamic competition in the economy affect antitrust rules for firms with significant market power. The colloquy begins with a paper by Keith Hylton and me which examines the implication of the fact that the antitrust laws generally do not condemn firms for having or acquiring significant market power, or enjoying the fruits of that power, as such. We conclude that the antitrust laws, like the intellectual property laws, are based on a tradeoff between static and dynamic monopoly power. Richard Schmalensee, the Chairman of our Editorial Board, has solicited comments from Jonathan Baker, Christian Ewald, Richard Gilbert, and Herbert Hovenkamp—all of whom disagree with the certain aspects of the article—

sometimes quite strongly. We are planning to continue this debate in the pages of *GCP—The Online Magazine* early next year. This grouping is followed by an article by Dennis Carlton and Ken Heyer which examines antitrust policy towards monopolies from a different perspective. They distinguish between the extension of monopoly power, which should be the subject of antitrust prohibitions, and the extraction of rents from a monopoly, which they argue should not be, in part because of its role in stimulating innovation.

The previous papers mention "Schumpeter" over 40 times. It is therefore fitting that the economist who coined the second most popular two words in economics should be the subject of our classic writings on antitrust this month. Thomas McCraw, the author of *Prophet of Innovation: Joseph Schumpeter and Creative Destruction*, has written an essay reviewing the Viennese Harvard professor's musings regarding antitrust. It contains excerpts from the small portions of Schumpeter's writings that actually dealt with antitrust as well as from a classic review of Schumpeter's views from a piece written a half-century ago by Professor Edward Mason.

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

David S. Evans
University College London and University of Chicago

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Private Enforcement of EU Competition Law in Member State Courts: Experience to Date and the Path Ahead

Christopher J. Cook

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Private Enforcement of EU Competition Law in Member State Courts: Experience to Date and the Path Ahead

*Christopher J. Cook**

A much cited 2004 study described private competition law litigation in Europe as being in a state of "total underdevelopment." Ever since, the European Commission, along with several member state governments and competition authorities, has made facilitating private enforcement of competition laws a cornerstone competition law policy. The situation is evolving—across Europe courts are becoming more open to antitrust plaintiffs and litigation is active. But with the main procedural and policy issues that govern civil litigation under the responsibility of 27 countries with rich and diverse legal traditions, progress has been uneven. The consequence is a patchwork under which European litigants on both sides of the table face difficult questions of standing, jurisdiction, access to evidence, and entitlement to and calculation of damages—the answers to which vary widely by country.

This article provides an overview of the current state of private enforcement of EU competition law. In doing so, it summarizes and assesses the central recommendations of the European Commission's 2008 White Paper on damages actions for breaches of EU competition law, considering the context under which the White Paper was issued as represented by the policy options laid out initially in the Commission's 2005 Green Paper, existing member state legal systems, and experience in national antitrust damages litigation.

*Cleary Gottlieb Steen & Hamilton LLP, Brussels. It would be impossible for one person to stay current regarding developments in national litigation across the European Union, and this article accordingly reflects input from many experts who reflect the great international breadth of Cleary Gottlieb's lawyers in Brussels and elsewhere. I am particularly grateful for substantive input from Shaun Goodman, Neil Rigby, and Andrew Ward, for coordinating assistance provided by Tanya Dunne and Colin Raftery, and for valuable comments from John Temple Lang.

I. Overview of EU Competition Law Application by National Courts

A. INTRODUCTION

The EC Treaty places individuals at the heart of EU competition policy, since it includes competition law among the tools to be used to achieve purposes such as economic development and an improved standard of living. In practice, however, private lawsuits have not played a prominent role in the enforcement of EU competition law. In contrast with the situation in the United States—where the commonly held view is that private actions make up 90 percent of antitrust enforcement—EU competition law enforcement has been driven almost exclusively by public authorities. Notwithstanding the formal recognition in Regulation 1/2003 of member state courts' power and obligation to apply EU competition law directly, member state law and practice concerning private actions for breach of competition laws remain relatively undeveloped.[1]

The European Commission is doing what it can to encourage member states to change this situation. Successive Commissioners for Competition have advocated private enforcement in national courts as an important prong of EU competition policy.[2] Commission press releases accompanying Article 81 cartel decisions now openly invite follow-on private actions for damages.[3] In late 2005, the Commission released a Green Paper for public comment, raising for discussion

1. A 2004 report for the European Commission characterized the situation as a state of "total underdevelopment." *See* D. Waelbroeck et al., Study on the Conditions of Claims for Damages in Case of Infringement of EC Competition Rules: Comparative Report (Ashurst report) (2004) [hereinafter "Comparative Report"], *available at* http://ec.europa.eu/comm/competition/antitrust/others/actions_for_damages/study.html.

2. *See, e.g.*, M. Monti, European Commission, Speech No. 04/403, Private litigation as a key complement to public enforcement of competition rules and the first conclusions on the implementation of the new Merger Regulation, Speech at the IBA - 8th Annual Competition Conference, Fiesole (Sep. 17, 2004); N. Kroes, European Commission, Speech No. 05/533, Enhancing Actions for Damages for Breach of Competition Rules in Europe, Dinner Speech at the Harvard Club, New York (Sep. 22, 2005); and N. Kroes, European Commission, Speech No. 05/613, Damages Actions for Breaches of EU Competition Rules: Realities and Potentials, speech at the conference 'La reparation du prejudice cause par une pratique anti-concurentielle en France et à l'étranger: bilan et perspectives', Paris (Oct. 17, 2005).

3. Commission press releases accompanying cartel decisions routinely conclude with the following statement:

> Any person or firm affected by anti-competitive behaviour as described in this case may bring the matter before the courts of the Member States and seek damages, submitting elements of the published decision as evidence that the behaviour took place and was illegal. Even though the Commission has fined the companies concerned, damages may be awarded without these being reduced on account of the Commission fine.

policy options that would facilitate private damages actions for breach of EU competition rules.[4] The Commission followed in April 2008 with a White Paper setting forth concrete proposals for action.[5]

The White Paper is a modest 10-page document that summarizes the Commission's proposals to address perceived obstacles to the development of private antitrust damages litigation in Europe. It is accompanied by detailed supporting documents, including a Staff Working Paper[6] which summarizes much of the reasoning underlying the White Paper's recommendations and an Impact Assessment[7] which analyzes the benefits and costs of various policy options that were considered in developing the White Paper. The White Paper introduces proposals for member state action on a range of issues central to the development of private antitrust enforcement in the European Union, including standing to bring claims, collective actions, disclosure/discovery rules, and the quantification of damages. The development of private enforcement in the European Union over the coming years will be driven in large part by the responses in legislation and in practice to the White Paper's recommendations.

The development of private enforcement in the European Union over the coming years will be driven in large part by the responses in legislation and in practice to the White Paper's recommendations.

This article provides an overview of the current state of private enforcement of EU competition law. In doing so, it summarizes and assesses the White Paper's central recommendations in light of the context in which they were issued, as represented by the policy options laid out initially in the Green Paper, the existing member state legal systems, and the experience to date in national antitrust damages litigation.

4. Commission Green Paper on Damages Actions for Breach of the EC Antitrust Rules, COM(2005) 672 final (Dec. 19, 2005) [hereinafter "Green Paper"].

5. Commission White Paper on Damages Actions for Breach of the EC Antitrust Rules, COM(2008) 165 final (Apr. 2, 2008) [hereinafter "White Paper"].

6. Commission Staff Working Paper Accompanying the White Paper on Damages Actions for Breach of the EC Antitrust Rules, SEC(2008) 404 (Apr. 2, 2008) [hereinafter "Working Paper"].

7. Commission Staff Working Document, Accompanying document to the White Paper on Damages actions for breach of the EC antitrust rules, Impact Assessment, COM(2008) 165 final, SEC(2008) 406 (Apr. 2, 2008) [hereinafter "Impact Assessment"]. The Impact Assessment is based in significant part on findings set forth in a 671-page Impact Study prepared by a team of external consultants. *See* CEPS, EUR & LUISS, Making Antitrust Damages Actions More Effective in the EU: Welfare Impact and Potential Scenarios, Final Report (Dec. 21, 2007) [hereinafter "Impact Study"].

B. GOALS OF PRIVATE ENFORCEMENT AND RELATION TO PUBLIC ENFORCEMENT

Civil litigation in member state courts is the primary means by which private parties may enforce the rights afforded them by competition law directly, through actions for injunctive relief or damages against another private party or the state. From the Commission's perspective, private litigation is not aimed at protecting the public interest but at protecting individuals' "subjective rights under Community law."[8] Such subjective rights include victims' rights to compensation for losses sustained as a result of competition law violations. The Commission regards improving citizens' awareness of and ability to directly enforce their rights under EU competition law as important in part because the possibility to be awarded damages "makes the competition rules instantly relevant for citizens."[9] Putting competition law to work for individual consumers has emerged as a central plank of Commissioner Neelie Kroes' tenure.

From the Commission's perspective, an important indirect benefit of private litigation is that it adds to the amount of competition law enforcement, thereby contributing to deterrence and consequently to compliance with competition rules. Private litigants may supplement public enforcement by taking action against infringements that the competition authorities are unwilling or unable to pursue due to lack of resources.[10] Increased private action may also improve the detection rate of competition infringements, as private parties who are victims of anticompetitive conduct may be better placed than public enforcers to identify violations.[11] The Commission thus regards private enforcement, particularly at the consumer level, as a complement to public enforcement and has made facilitation of private enforcement a clear policy goal. In the words of Commissioner Kroes:

---

8. Council Regulation (EC) No. 1/2003 of December 16, 2002 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty, 2003 O.J. (L 1) 1, at recital 7.

9. Kroes (Sep. 22, 2005) *supra* note 2.

10. M. Monti, European Commission, Speech No. 04/403, Private litigation as a key complement to public enforcement of competition rules and the first conclusions on the implementation of the new Merger Regulation, Speech at the IBA - 8th Annual Competition Conference, Fiesole (Sep. 17, 2004), at 2. The European Court of Justice relied on a similar argument over forty years ago in support of establishing the direct effect of EU law in Case 26/62, Van Gend & Loos, 1963 E.C.R. 1 ("The vigilance of individuals concerned to protect their rights amounts to an effective supervision in addition to the supervision entrusted by articles 169 and 170 to the diligence of the commission and of the member states.").

11. *See* D. Woods et al., *Private enforcement of Community competition law: modernization and the road ahead*, 2 COMPETITION POL'Y NEWSLETTER 31, 33 (2004).

> "No matter how closely public intervention mirrors the concerns of consumers, no matter how effectively the fines that we impose punish and deter unlawful behaviour, the victims of illegal behaviour will still not be compensated for their losses. [Public enforcement] cannot make amends for the damage and suffering caused to consumers. Therefore, consumers should be empowered to enforce their rights themselves."[12]

As the discussion above indicates, the effort to promote private antitrust actions has been gathering momentum for several years, but it has not been without its critics, even from within the Commission. For example, it has been argued that public enforcement is inherently superior to private enforcement, partly since public enforcement benefits from more effective investigative and sanctioning powers than private actions, which are driven purely by profit motives and are globally more costly for society. These commentators also argue that deterrence should be achieved through tougher public sanctions (including jail sentences) and increased resources for competition authorities, rather than the threat of private damages actions.[13] In response, the Commission has repeated the need to "strike the right balance" between effective private enforcement and excessive litigation[14] and clarified that its intent is to "foster a competition culture, not a litigation culture."[15] In accordance with these statements, the White Paper is careful to emphasize the need to preserve a "genuinely European approach" to the issue of damages actions that is "rooted in European legal culture and traditions."[16]

C. THE LEGAL BASIS FOR PRIVATE ENFORCEMENT

Because there is no European court competent to hear damages actions brought by private plaintiffs for breach of EU competition law, private enforcement requires individuals to invoke European competition law before national courts, which must apply EC law directly. The direct applicability of Articles 81 and 82

12. N. Kroes, European Commission, Speech 07/698, Making consumers' right to damages a reality: the case for collective redress mechanisms in antitrust claims, Speech at the Conference on collective redress for European consumers, Lisbon (Nov. 9, 2007).

13. *See, e.g.*, W.P.J. Wils, *Should Private Enforcement Be Encouraged in Europe?*, 26 World Competition 473 (2003).

14. Kroes (2007), *supra* note 12; Kroes (Sep. 22, 2005), *supra* note 2.

15. Kroes (Sep. 22, 2005), *supra* note 2.

16. White Paper, *supra* note 5, at § 1.2.

EC has long been recognized in the case law,[17] and the power and obligation of national courts to apply EU competition law is now formalized in Article 6 of Regulation 1/2003.[18]

Although litigation in national courts is based on national procedural rules, the principle of EC law primacy and the duty of loyal cooperation under Article 10 EC impose constraints on national courts on how they handle private damage actions based on alleged EU competition law violations. The first consequence of these principles is that national courts are obliged to construe national law in light of European law[19] and even to disapply any provision of national law that would be contrary to European law.[20] National courts will not be able to apply national laws that frustrate damage actions under EU competition law.

Some additional constraints on national courts posed by the EC law primacy principle are now embodied in Regulation 1/2003. Article 3 of Regulation 1/2003 compels national courts to apply European competition law if the conduct in question may affect trade among member states. The courts may apply national competition law alongside the EC provisions, but only if the outcome under the national law does not differ from that under EC law (the only exception being that national laws on unilateral conduct may be stricter than Article 82 EC[21]). In addition, Article 16 of Regulation 1/2003 promotes the uniform application of Community law by prohibiting national courts from issuing judgments running counter to a previous Commission decision relating to the same agreement or practice.[22]

The procedural autonomy enjoyed by member states in the absence of harmonization[23] is also limited by the two cornerstone EC law principles of equivalence and effectiveness, derived from Article 10 EC. The equivalence principle

17. *See, e.g.*, Case 127/73, BRT v. SABAM, 1974 E.C.R. 51, at para. 16 ("[A]s the prohibitions of Articles [81] (1) and [82] tend by their very nature to produce direct effects in relations between individuals, these articles create direct rights in respect of the individuals concerned which the national courts must safeguard.").

18. Article 6 of Regulation No. 1/2003 provides that: "National courts shall have the power to apply Articles 81 and 82 of the Treaty." *See* Reg. No. 1/2003, *supra* note 8, at art. 6.

19. Case C-106/89, Marleasing, 1990 E.C.R. I-4135.

20. Case 106/77, Simmenthal, 1978 E.C.R. 629.

21. Reg. No. 1/2003, *supra* note 8, at art. 3(2). Pursuant to Article 3(3), the application of national legislation, the primary objective of which is different from Articles 81 and 82, may lead to a different outcome.

22. This provision formalizes the *Masterfoods* doctrine from Case C-344/98, Masterfoods, 2000 E.C.R. I-11369 [hereinafter *Masterfoods*], at paras. 49-52.

23. Case 33/76, Rewe, 1976 E.C.R. 1989, at para. 5; and Case 45/76, Comet, 1976 E.C.R. 2043, at para. 12.

requires national courts not to treat claims founded on EC law less favorably than claims under national law. Consequently, all the mechanisms available to individuals to enforce their rights under national competition law are extended to European competition law. The effectiveness principle goes further, providing that member states may not render enforcement of European law impossible or extremely difficult. National courts may even be forced to invent remedies that do not exist in national law if their absence puts at risk the effectiveness ("effet utile") of European competition law.[24]

This principle has been central in the development of private enforcement of EU competition law. In the European Court of Justice's *Banks* case from the early 1990s, Advocate General Van Gerven argued that the effectiveness principle supports allowing damages actions for losses sustained by EU competition law violations, and invited the Court to develop a case law to this effect.[25] The Court declined to set such a precedent in that case, but moved in this direction several years later in the landmark *Crehan* judgment, holding that:

> "The full effectiveness of Article [81 EC] . . . would be put at risk if it were not open to any individual to claim damages for loss caused to him by a contract or by conduct liable to restrict or distort competition. [. . .] The existence of such a right strengthens the working of the Community competition rules and discourages agreements or practices, which are frequently covert, which are liable to restrict or distort competition. From that point of view, actions for damages before the national courts can make a significant contribution to the maintenance of effective competition in the Community."[26]

Crehan opened the way to private enforcement of EU competition law. Although the judgment concerned Article 81 EC, the Court's holdings clearly go beyond the facts of the case to apply also to non-contractual relationships, including breaches of Article 82 EC. Nevertheless, *Crehan* does not create a new European remedy. Rather, "the consequences in civil law attaching to an infringement of [Articles 81 and 82 EC] . . . are to be determined under nation-

24. Case C-213/89, Factortame I, 1990 E.C.R. I-2433.

25. Opinion of Advocate General Van Gerven, Case C-128/92, Banks v. British Coal Corporation, 1994 E.C.R. I-1209.

26. Case C-453/99, Courage v. Crehan & Crehan v. Courage, 2001 E.C.R. I-6297 [hereinafter *Crehan*], at paras. 26-27.

al law . . . , subject, however, to not undermining the effectiveness of the Treaty."[27] Thus, the principle of a private right to damages for competition law violations has been recognized in European law, but it must be implemented by national law.

Today, in the absence of Europe-wide rules, each member state has its own rules governing civil litigation and the systems differ widely. While some member states maintain rules that act as obstacles to effective private competition law enforcement, others are more receptive to private actions, either due to features of their civil procedural law or because they have adopted specific legislation relating to competition law claims. Section II below outlines several key procedural and substantive issues relating to the development of private competition law enforcement and summarizes how they have been approached in various member states and in the White Paper.

WHILE SOME MEMBER STATES MAINTAIN RULES THAT ACT AS OBSTACLES TO EFFECTIVE PRIVATE COMPETITION LAW ENFORCEMENT, OTHERS ARE MORE RECEPTIVE TO PRIVATE ACTIONS, EITHER DUE TO FEATURES OF THEIR CIVIL PROCEDURAL LAW OR BECAUSE THEY HAVE ADOPTED SPECIFIC LEGISLATION RELATING TO COMPETITION LAW CLAIMS.

II. Key Issues in the Development of Private Enforcement

A. STANDING: WHO CAN BRING A CLAIM?

The Court of Justice in *Crehan* articulated a broad standard of who has the right to bring a private action for EU competition law violations, holding that the full effectiveness of EU competition law "would be put at risk if it were not open to *any individual* to claim damages for loss caused to him by a contract or by conduct liable to restrict or distort competition."[28]

Not all member states have always clearly supported such a broad right of standing. Member states require a potential claimant to demonstrate an "interest" in bringing an action, but some had imposed additional limitations that set

27. Case T-395/94, Atlantic Container Line, 2002 E.C.R. II-875, at para. 414.

28. *Crehan*, *supra* note 26, at para. 26 (emphasis added). *See also* Joined Cases C-295/04 to C-298/04, Vincenzo Manfredi et al. v. Lloyd Adriatico Assicurazioni SpA et al., 2006 E.C.R. I-6619 [hereinafter *Manfredi*], at para. 61 ("any individual can claim compensation for the harm suffered where there is a causal relationship between that harm and an agreement or practice prohibited under Article 81 EC").

the bar higher for plaintiffs to get into court.[29] In recent years, however, several countries with apparently diverging rules (e.g., Germany) have amended laws to make their courts more open to claimants, and today there are few if any practical limitations on the right of any natural or legal person who has suffered harm as a result of violations of Article 81 or 82 EC to bring a claim in an appropriate member state court (jurisdictional issues are addressed below). The most pressing issues concerning potential antitrust plaintiffs' access to the courts now relate to the availability of representative or class action proceedings and the funding of claims.

1. Representative and Collective Actions

Damages suffered individually by final consumers and low-volume purchasers will often be too small to make litigation worthwhile even if the aggregate harm caused by anticompetitive behavior is large. The Commission has for several years advocated the view that effective private enforcement requires some form of collective action to consolidate small claims and spread the costs and risks of litigation. While perhaps implicitly recognizing the rarity of collective actions in Europe to date, in the period leading up to issuance of the White Paper Commissioner Kroes stated that "representative action—empowering groups that truly represent the interests of consumers—is [close] to the heart of European traditions."[30]

In the United States, collective antitrust claims are brought primarily through "opt-out" class actions, in which a single plaintiff is able to commence an action on behalf of an entire class of unnamed plaintiffs (thus requiring those who do not wish to participate in the action to opt out). Opt-out class actions are generally regarded as efficient vehicles for aggregating small claims that would be unlikely to be individually litigated as they lower the cost to individual litigants, bring superior legal resources to bear, and strengthen the plaintiffs' negotiating position. On the other hand, opt-out class actions are procedurally complex and often slow to proceed (particularly since the initial class certification stage of the litigation is often lengthy and highly contentious), are expensive to pursue and defend, and significantly diminish individual control by plaintiffs.

29. A much-discussed example from Germany is perhaps most notable. In 2003, the Landgericht Berlin (district court) placed a tight restriction on standing by requiring that the claimant (who was seeking damages from members of a cement cartel) show that the law on which the claim was based had the purpose of protecting the claimant, and thus that "purchasers of cement at cartel prices could not claim damages unless they had been individually targeted by a market-sharing cartel." Max Boegl Bauunternehmung et al. v. Hanson Germany (Ger. Dst. Ct. Jun. 27, 2003).

30. N. Kroes, European Commission, Speech 07/128, Reinforcing the fight against cartels and developing private antitrust damage actions: two tools for a more competitive Europe, Speech at the Commission/IBA Joint Conference on EC Competition Policy, Brussels (Mar. 8, 2007).

Along with contingent legal fees, burdensome discovery processes, and treble damages, the opt-out class action process has been widely regarded in Europe as one of the principal "excesses" of the U.S. system. Today, no EU member state has a procedural device directly parallel to the opt-out class action, although several countries do allow certain types of collective actions to be brought and the laws in this area are evolving quickly. In 2007 Commissioner Kroes stated expressly that she "would not support the introduction" of opt-out class actions in Europe.[31] Unsurprisingly, then, the White Paper does not advocate the introduction of opt-out class actions, but instead proposes the adoption of two alternative measures to facilitate collective redress:

- representative actions brought by qualified entities, such as consumer associations or trade associations, on behalf of identified or, in some limited cases, identifiable victims. These entities would be either (i) officially designated in advance or (ii) certified on an *ad hoc* basis by a member state to bring an action on behalf of some or all of their members for a particular antitrust infringement. Damages would be awarded to the representative entity and used preferably to compensate the harm suffered by all those represented in the action; and

- opt-in collective actions, in which victims decide to combine their individual claims for damages suffered into one single action.

As explained below, each of these types of actions has been applied at the member state level, and neither Commission proposal is likely to be controversial. Below are summaries of some of the approaches currently being taken.

a. Representative Actions and Consumer Claims

Several member states permit collective damages actions to be brought by consumer associations. Consumer actions are a useful mechanism to bring some claims that would otherwise not be brought due to their small individual value and the large number of claimants.

A few such actions have already been initiated in the antitrust context. In England, under the Competition Act 1998, certain specified bodies are entitled to bring collective consumer claims for damages on behalf of, and with the consent of, named individual consumers. These are follow-on actions before the Competition Appeals Tribunal ("CAT") all relating to the same infringement, with damages paid to the individual consumers. The first consumer action was commenced in February 2007 by the consumer association Which? on behalf of consumers who purchased replica football shirts from retailers during 2000 and 2001, following on from the previous decision of the U.K. Office of Fair Trading

31. *Id.*

("OFT") that retailers had fixed the prices of replica football jerseys.[32] On January 9, 2008, Which? announced that it had reached a settlement with the defendant, retailer JJB Sports, according to which JJB would pay cash refunds of £20 to consumers who had joined the Which? action and £10 to consumers who had purchased shirts during the infringement period but had not opted into the collective action.[33]

A similar case is ongoing in France. In 2006, UFC-Que Choisir, a major French consumer association, launched an unprecedented follow-on damages action based on a decision by the French Competition Council finding that the three main mobile phone operators (Orange, SFR, and Bouygues) had entered into a market sharing cartel and fined the operators a total of €534 million.[34] UFC-Que Choisir created a website[35] where customers of the mobile phone operators could (i) use a "damage calculator" to determine how much damages they could claim and (ii) give the association the power to bring a claim on their behalf. In October 2006, UFC-Que Choisir filed a complaint before the Commercial Court of Paris seeking damages on behalf of over 12,500 consumers.[36] In December 2007, the Court ruled that the action against Bouygues was inadmissible.[37] UFC-Que Choisir has since suspended its parallel actions against Orange and SFR and appealed the Bouygues judgment to the Paris Court of Appeal, which is expected to rule on the case in early 2009.

In Spain, in late 2007 the Association of Banking Services Users (*Ausbanc Consumo*) launched a claim against Telefónica seeking €458 million in damages,

---

32. Case No. CA98/06/2003, Price Fixing of Replica Football Kit (OFT, Aug. 1, 2003).

33. *See* Press Release, Which?, JJB to pay fans over football shirt rip-off (Jan. 9, 2008), *available at* http://www.which.co.uk/reports_and_campaigns/consumer_rights/reports/Ripoffs,%20scams%20and %20fraud/JJB_agree_shirts_deal_news_article_557_128985.jsp.

34. *See* Press Release, Conseil de la concurrence, The Conseil de la concurrence imposes fines totalling 534 million Euros on Orange France, SFR and Bouygues Télécom (Dec. 1, 2005), *available at* http://www.conseil-concurrence.fr/user/standard.php?id_rub=160&id_article=502.

35. Cartelmobile.org, Home, *at* http://www.cartelmobile.org (last visited Sep. 5, 2008).

36. *See* Cartelmobile.org, Soutenez-nous!, *at* http://www.cartelmobile.org/soutien.php (last visited Sep. 5, 2008).

37. French law provides two forms of collective action that may be used by consumer associations: a proceeding in the "collective interest" of consumers ("procédure dans l'interêt collectif des consommateurs") (French Consumer Code, at art. L421-1 to L422-8) and a "joint representation" action ("action en représentation conjointe") (French Consumer Code, at art. L422-1 to L422-3). UFC-Que Choisir had brought a collective interest action, which the Court held was inadmissible since such actions cannot be initiated (but only joined) by a consumer association. The association should have used the joint representation procedure, but by soliciting plaintiffs on its website and via e-mail, UFC-Que Choisir violated the rules governing joint representation actions, which state that a mandate to bring action "may not be solicited by means of a public appeal on radio or television, nor by means of posting of information, by tract or personalised letter. Authorisation must be given in writing by each consumer."

following on from the European Commission's decision finding Telefónica guilty of abusing its dominant position in the Spanish broadband internet market and imposing a fine of €151 million.[38] Damages are sought on behalf of all holders of an ADSL line in Spain during the relevant period. Using data from the Spanish telecom regulator and the European Commission, Ausbanc arrived at its damages figure by subtracting the Commission's fine from the €600 million it says Telefónica earned during the five years it abused its dominant position. In December 2007, the Madrid Commercial Court granted leave for the action to proceed and stayed the action for two months to give aggrieved parties time to appear, having published notice of the action in the leading Spanish newspapers.[39] This claim is the first of its type in Spain by a consumer group. It benefits from the new Spanish competition law, which allows individuals to claim damages before the courts without a prior declaration from the Spanish competition authority that a practice is prohibited.

Similar procedural devices are available in some other member states, such as Austria,[40] Portugal,[41] and Sweden,[42] although they have apparently not yet been used to seek damages for competition law infringements.

Despite the precedents that exist, the consumer claim procedure is limited in scope, as it is an opt-in procedure, is generally only available to consumers, and proceedings must usually be brought as a follow-on action. In May 2008 Which?, the U.K. consumer group that led the litigation against JJB Sports, stated that

38. Commission Decision of 4 July 2007, Case COMP/38.784 - Wanadoo España v. Telefónica, 2008 O.J. (C 83) 5.

39. *See* Press Release, Ausbanc Consumo, Action Against Telefonica Is Granted Leave to Proceed (Dec. 12, 2007), *available at* http://www.ausbanc.es/medios/notasprensa/detalle/2007/np20071212telefonica_ing.pdf.

40. The Austrian Consumer Protection Act also provides for the possibility of collective actions brought by certain enumerated associations, including the Verein für Konsumenteninformation (a consumer rights association), on behalf of the general public to obtain cease and desist orders. *See* Consumer Protection Act (Konsumentenschutzgesetz - "KSchG"), at § 29 and Act against Unfair Competition (Gesetz gegen den unlauteren Wettbewerb - "UWG"), at § 14.

41. Portugal allows for a limited opt-out procedural device in representative or collective action cases brought by an association or public prosecutor (so-called popular actions), although these actions seem to serve more as actions in the collective interest than as means of securing individual compensation and have not yet been applied in the antitrust context. *See* Impact Study, *supra* note 7, at 310 & 316.

42. In Sweden, private damages actions may be initiated by class actions under the Class Action Act. Lag (2002:599) om grupprättegång. *See also* Swedish Competition Act, Konkurrenslag (1993:20), at art. 33, *available at* http://www.kkv.se/t/Page_905.aspx (English translation). Class actions may be initiated by: (i) private individuals or legal entities; (ii) consumer or labor organizations; or (iii) any public authority designated by the government. The Swedish class action system is based on the "opt-in" model, which means that only class members who have given written notice to the court may ultimately participate as passive members in the proceedings.

further representative damage claims based on the opt-in model used in the JJB action are "highly unlikely" because they are not financially viable. (JJB Sports reportedly set aside £100,000 to cover payments under the settlement—far less than the £6.7 million fine that it paid to the OFT.) Which? claims that only an opt-out model would allow the group to assemble a sufficient number of claimants to make actions worthwhile.[43] This position is at odds with the White Paper's focus on promoting representative claims based on an opt-in regime (i.e., precisely the model that Which? has applied but now considers unviable).

The OFT had already signaled its desire for the law to move in the direction proposed by Which?. In late 2007 the OFT issued recommendations regarding the promotion of private actions, at the heart of which is a series of ways to expand the use of the representative action device. Among the most important OFT recommendations are: (i) to allow stand-alone claims to be brought through representative actions (as opposed to the current U.K. system under which representative actions are limited to follow-on actions before the CAT); (ii) to make available a representative action procedure for small and medium-sized businesses (as opposed to consumers only)—in line with the White Paper's proposal; and (iii) to introduce the possibility of opt-out representative actions for damages on behalf of consumers/businesses at large (as opposed to the current opt-in procedure on behalf of named consumers).[44] The OFT's proposals gained further momentum in July 2008, when the U.K. Civil Justice Council ("CJC"), an advisory public body established under the Civil Procedure Act 1997 with responsibility for overseeing and coordinating the modernization of the civil justice system, recommended a series of measures to encourage better enforcement of consumer rights. Among the CJC's recommendations to the Lord Chancellor is establishing a court-supervised opt-out procedure for collective redress.[45]

The issue of collective redress is also under active consideration in other jurisdictions. For example, the French Parliament is scheduled to discuss the law on the "class action à la française" by the end of 2008, after three years of postponements. The European Commission is also currently studying collective redress more generally (in relation to areas besides just competition law) and intends to issue another communication on the issue in December 2008.

43. Deborah Prince, Which? Head of Legal Affairs, Presentation at the British Institute of International and Comparative Law, 8th Annual Trans-Atlantic Antitrust Dialogue, London (May 15, 2008).

44. U.K. Office of Fair Trading, Private actions in competition law: effective redress for consumers and business (OFT Discussion paper, No. 916, Apr. 2007) [hereinafter "OFT Discussion Paper"], at §§ 5-7.

45. *See* http://www.civiljusticecouncil.gov.uk/files/Improving_Access_to_Justice_through_Collective_Actions.pdf.

b) Acquisition of Claims

In Germany, representative damages claims are not permitted, as damages may only be awarded to parties who have suffered loss. (Actions by professional associations to recover profits illegally obtained through breach of competition rules are unlikely to be pursued, as recovered profits must be paid to the state.) However, multiple claims may be assigned to a single entity to enable a single collective case to proceed. This procedure is currently being used in a case before the Düsseldorf Regional Court, in which claims have been brought on behalf of 29 companies by the Belgium-based company Cartel Damages Claims ("CDC")[46] against six cement producers, following on from a 2003 finding of infringement and imposition of €660 million in fines against the cement cartel.[47] In this action, CDC reportedly acquired the damages claims of 29 companies (direct purchasers) affected by the activities of the cartel for EUR 1,000, and agreed to return approximately 75 to 85 percent of the EUR 151 million claim sought to these customers in the event of success.[48] Costs of the proceedings (reportedly well over EUR 2 million) were covered by initial lump sum contributions from the assignors.

The Regional Court admitted the claims in an interim judgment on February 21, 2007. In addition to confirming its jurisdiction over the case, the court confirmed CDC's right to bring the claim, characterizing the assignment of the claims to it as "a full transfer of rights," which granted CDC standing under German law.[49] On May 14, 2008, the Higher Regional Court of Düsseldorf affirmed the lower court's interim judgment on appeal, confirming that CDC's claims are admissible. The Higher Court's judgment can also be appealed, so it may be some time before the case is finally resolved. However, if this type of action is ultimately upheld in Germany or elsewhere,[50] it may pave the way for opportunistic companies to acquire claims in order to investigate and pursue them.

c) Mass settlement

In the Netherlands, collective actions brought by representative associations may seek declaratory relief, on the basis of which individual actions for damages may be brought. Since July 2005, it is also possible for a representative associa-

46. *See* CDC, Cartel Damage Claims, *at* http://www.carteldamageclaims.com/english/introduction_engl.htm (last visited Sep. 5, 2008).

47. Case O (Kart) 147/05 – CDC (Regional Court of Düsseldorf, 2005) [hereinafter *CDC*].

48. *Germany: A Private Affair,* Focus Europe (The American Lawyer, Summer 2007).

49. However, the Court confirmed that the defendant's arguments to the effect that the transfer was invalid were to be considered separately in assessing the merits of the claims.

50. CDC is also reportedly pursuing potential cases in the Netherlands and the United Kingdom.

tion to settle a mass damages claim. The new Act on Collective Settlement of Mass Damages[51] includes a process by which a legal entity may be created to represent the interests of a group of injured individuals and given the authority to reach a settlement agreement with the party that caused the damage. The parties must jointly submit the settlement agreement to the Amsterdam Court of Appeal for approval. The settlement agreement must contain: (i) the criteria to allow injured individuals to be eligible for payment; (ii) an accurate estimate of the total number of injured individuals; (iii) the total amount of damages to be awarded; and (iv) the formula used to calculate the damages for each injured individual. Injured individuals satisfying the criteria under the settlement agreement will have the option to "opt out" before a specified date set by the court (usually within 3 months), after which the settlement becomes binding upon all class members who have not opted out. Thus far, this procedure has been used only in the product liability and financial sectors, and it remains to be seen whether it will be expanded to the competition law sector.

2. Costs and Funding Claims

Litigation is expensive. Civil litigation in most EU member states operates on a "loser pays" principle designed to discourage unmeritorious claims (and defenses) by requiring the losing party to pay its own costs as well as a proportion of the costs of its adversary (in the United Kingdom, normally around 60 percent). This works to deflate legal costs for successful private plaintiffs. However, it also creates risk. Given the inherent uncertainty in bringing a lawsuit, a plaintiff will always face the risk of losing the case and having to pay not only its own legal costs but also the defendant's. In addition, the "loser pays" rule can work against bringing claims involving small amounts of damages, since even successful claimants often do not recover 100 percent of their costs, and must make up the shortfall from the damages award. In complex cases, the amount of the cost award shortfall may even exceed the damages awarded, thereby deterring meritorious damage claims from being brought.

The Commission's 2005 Green Paper cited Article 6 of the European Convention on Human Rights (which guarantees everyone fair access to courts) as providing support for the notion that member states are under a legal obligation to design their cost rules (including the level of court fees) so that damages actions can "effectively" be brought before the competent national courts. The Green Paper had posed a number of options that would potentially have alleviated cost burdens on plaintiffs. However, the White Paper takes a conservative approach and does not suggest any specific changes to national cost regimes in favor of claimants. It merely encourages member states "to reflect on their cost

51. Wet collectieve afwikkeling massaschade, Dutch Code of Civil Procedure, at §§ 7:907-10 & 1013-18.

rules so as to facilitate meritorious litigation, taking into consideration existing practices."[52] Appropriate measures could include:

- encouraging settlements as a way to reduce costs;
- setting court fees at a level where they do not become a disproportionate disincentive to bringing antitrust damages claims; and
- giving national courts discretionary "cost-capping" authority (i.e., the possibility of issuing cost orders derogating, in justified cases, from the normal "loser pays" cost rules).[53]

The measures to reduce fee risk and burden on potential antitrust complainants that have been enacted or are being considered at the member state level include the following.

a) Cost capping

The Green Paper asked whether it would be appropriate to exempt unsuccessful plaintiffs from paying the defendant's legal costs, save where actions have been introduced in a "manifestly unreasonable manner." Such a rule would put claimants in a strong position, given the low likelihood of a finding of "manifest unreasonableness." The White Paper does not take up this proposal from the Green Paper, but the possibility of such "cost-capping" orders already exists in some member states. In the United Kingdom, for example, while cost-capping orders have been used infrequently in the commercial context, they might—if granted in the context of consumer claims—prove a useful tool in facilitating claims by consumers or businesses of modest means.[54] The OFT has also recently suggested that in representative actions it may be appropriate to cap the claimant's liability for the defendant's costs at zero.[55] In Germany,[56] claimants of limited means may be granted a reduction in the costs payable (i.e., the fees of the defendants and of their own lawyers) if they lose their case.

b) Contingency fees

In the United States, a common solution to the costs difficulty for plaintiffs is to allow for contingency fees (i.e., arrangements in which no fee is charged by the plaintiff's lawyers unless the claim is successful), in which case the legal fee is

---

52. Working Paper, *supra* note 6, at para. 245.

53. White Paper, *supra* note 5, at § 2.8.

54. In damages actions before the Competition Appeals Tribunal ("CAT"), CAT Rule 55 provides that costs are not based on the "loser pays" principle but are within the discretion of the CAT.

55. OFT Discussion Paper, *supra* note 44, at paras. 8.11-8.18.

56. GWB, at § 89a.

often expressed as a percentage of the damages awarded. Contingency fees are another aspect of the U.S. system that has drawn criticism from some in Europe, generally on grounds that they promote excessive litigation—either because bringing a claim carries little or no risk for the plaintiff (encouraging meritless litigation) or because they can result in exorbitant fees for plaintiffs' lawyers.

Contingency fee arrangements are not permitted in most EU member states, although some partial exceptions have recently developed. In Italy, outcome-based fee structures that would likely permit "no win, no fee" arrangements have been allowed since 2006. In Germany, third-party funders may acquire claims with the condition that the third party will pay a "success fee" to the assignor of the claim if the third-party funder wins the case. This structure represents something much like a contingency fee: the third party buys a claim for a reduced amount and, in the event of winning the case, pays a share of the proceeds to the assignor (i.e., the injured party). As explained above, this approach was recently endorsed by the Regional Court of Düsseldorf (Landgericht) in the *CDC* judgment.[57] However, because this structure effectively permits circumventing the prohibition in German law on contingency fees, the judgment has been controversial and is currently under appeal.

c) Conditional fees

In England, conditional fees—lawyers' fees that may be increased if the claim is successful—are permitted. The maximum increase is currently set at 100 percent. However, the OFT has recently questioned whether, in view of the risk and complexity, this cap provides sufficient incentive for law firms to offer conditional fees in antitrust cases. The OFT recommends this cap be reconsidered for certain competition law damages cases, subject to judicial supervision of the funding arrangement. The OFT also suggests that fee increases in excess of 100 percent might more appropriately be deducted from the damages award rather than paid by the defendant.[58]

d) Legal aid

While legal aid is not likely to be available or sufficient to fund most claims (particularly since in most member states damages claims are ineligible for legal aid), on occasion it can represent an additional or even a primary source of funding. In a notable example in England, pub owner Bernie Crehan was granted "exceptional funding" to pursue his case against Inntrepreneur (including, in particular, defending Inntrepreneur's appeal at the House of Lords of the Court of Appeal's 2004 damage award). The English Lord Chancellor can grant exceptional fund-

57. CDC, *supra* note 47.

58. OFT Discussion Paper, *supra* note 44, at paras. 8.1-8.10.

ing where a case has "significant wider public interest" or the lack of public funding would lead to obvious unfairness in the proceedings.

e) Third-party funding

Another approach is to permit third parties to fund litigation. There are several firms specializing in litigation funding in Germany and England. In England, third-party commercial funders offer loans to claimants that are repayable only if the claimant is successful. If the claimant loses, the commercial funder's liability for the defendant's costs will normally be capped at an amount equivalent to the amount of funding provided to the claimant, so long as the claimant (rather than the funder) remains in control of the litigation as the primarily interested party.[59] Recently, UK plaintiffs' lawyers and insurance providers have come together to offer a "risk-free" funding model to antitrust damages claimants under which the law firm undertakes cases on a conditional fee basis while the insurer provides after-the-event insurance to protect the claimant against adverse costs orders.[60] The OFT has recently expressed the view that third-party funding is an important potential source of funding for competition law damages claims, and has encouraged the creation of a merits-based litigation fund for competition claims where commercial basis funding is not available.[61]

However, the question of jurisdiction—which country to litigate in—is a complex issue for potential claimants, defendants, and courts hearing damages actions under EU competition law. When one or more of the parties to the dispute operates on a cross-border basis and/or is located in different member states, where should the claim be heard?

As explained above, in Germany third parties have even been permitted to acquire claims, rather than just provide funding. In the *CDC* judgment, the Regional Court of Düsseldorf held that a claim may be assigned to a third party (in this case the firm Cartel Damages Claims ("CDC")), which will fund and pursue the claim and may agree to share any proceeds of the case with the injured parties. As noted above, this approach largely circumvents the German legal prohibition on contingency fees, but the judgment was upheld on appeal in May 2008. Additional avenues for appeal remain open to the defendants, however, so final case resolution may be some time away. Whatever the outcome, whether such an approach would be accepted outside Germany is not clear, particularly since many other member states also prohibit or restrict contingency fees or third parties acquiring and directing litigation.

59. Arkin v. Borchard Lines, 2005 E.W.C.A. Civ. 655.

60. *Cartel Case Approaches*, Law Society Gazette, August 21, 2008, at 3.

61. OFT Discussion Paper, *supra* note 44, at 8.19-8.26.

B. JURISDICTION: WHERE CAN THE CLAIM BE BROUGHT?

1. Choice of Forum

Article 6 of Regulation 1/2003 gives national courts the express power to apply Articles 81 and 82 EC. However, the question of jurisdiction—which country to litigate in—is a complex issue for potential claimants, defendants, and courts hearing damages actions under EU competition law. When one or more of the parties to the dispute operates on a cross-border basis and/or is located in different member states, where should the claim be heard?

a) EU defendants

For defendants domiciled within the European Union, this question is answered by Regulation 44/2001 EC, the so-called "Brussels Regulation."[62] The basic rule is that defendants are to be sued in the member state in which they are "domiciled," which may be the location of its statutory seat, central administration, or principal place of business. That member state court will have jurisdiction to rule on all the harm suffered by the claimant, including harm suffered outside the jurisdiction.

There are several alternatives to this straightforward rule that are likely to apply in competition law damages actions, potentially creating additional jurisdictional possibilities.

- First, in matters relating to contract, a claim may be brought in the member state where the contractual obligation was or should have been performed (Article 5(1)).
- Second, in matters of tort (e.g., breach of statutory duty, which is the heading under which many member states place competition law damage claims), a defendant may be sued in the member state where the harmful event occurred (Article 5(3)). According to the European Court of Justice's judgment in *Bier*,[63] this provision gives claimants two alternatives: (i) the place where the harmful event that is the origin of the damage occurred; or (ii) the place where the damage was suffered (although jurisdiction under this second alternative will be limited to damage suffered in that member state[64]).
- Third, where there are multiple defendants, a claim may be brought in the member state in which any one of them is domiciled, provided that the claims are so closely connected that it is expedient to hear

62. Council Regulation (EC) 44/2001 of 22 December 2001 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters, 2001 O.J. (L 12) 1.

63. Case 21/76, Handelskwekerij G. J. Bier BV v. Mines de potasse d'Alsace SA, 1976 E.C.R. 1735, at paras. 24-25.

64. Case C-68/93, Shevill v. Presse Alliance SA, 1995 E.C.R. 1, 415, at para. 33.

> and determine them together to avoid the risk of irreconcilable judgments (Article 6).[65]

The European Courts have held that these exceptions to the general "defendant's domicile" rule should be interpreted narrowly.[66] The Article 5 exceptions are only to be applied based on the existence of a "particularly close connecting factor" between the dispute and a member state other than the defendant's domicile.[67] In light of the *Bier* judgment, Article 5(3) is susceptible of broad interpretation, since, in a sense, any claimant will suffer loss at its place of business, regardless of where the harmful act was committed. To prevent the court of the claimant's place of business assuming jurisdiction in every such case and so weakening or entirely subverting the basic jurisdictional principle of Article 2, it is settled that the "place of damage" means the place where the physical damage is done or the recoverable economic loss is actually suffered.[68]

In addition to the rules under the Brussels Regulation, if the parties have contractually agreed that any disputes that might arise should be determined in the court of a particular member state, the contractual jurisdiction clause, rather

65. *See also* Provimi v. Roche and Ors., 2003 E.W.H.C. 961 (Comm. Ct.) [hereinafter *Provimi*], discussed further *infra*.

66. *See* Case 189/87, Kalfelis v. Schröder, 1988 E.C.R. 5565, at para. 19 ("the 'special jurisdictions' enumerated in Articles 5 and 6 of the Convention constitute derogations from the principle that jurisdiction is vested in the courts of the State where the defendant is domiciled and as such must be interpreted restrictively").

67. Case 220/88, Dumez France v. Hessische Landesbank, 1990 E.C.R. I-49, at para. 17:

> [T]hose cases of special jurisdiction, the choice of which is a matter for the plaintiff, are based on the existence of a particularly close connecting factor between the dispute and courts other than those of the defendant's domicile, which justifies the attribution of jurisdiction to those courts for reasons relating to the sound administration of justice and the efficacious conduct of proceedings.

68. *Id*. at para. 20:

> It follows from the foregoing considerations that although, by virtue of a previous judgment of the court (in [*Bier*] . . .), the expression "place where the harmful event occurred" contained in Article 5(3) of the [Brussels Regulation] may refer to the place where the damage occurred, the latter concept can be understood only as indicating the place where the event giving rise to the damage, and entailing tortious, delictual or quasi-delictual liability, directly produced its harmful effects upon the person who is the immediate victim of that event.

See also Case C-364/93, Marinari, 1995 E.C.R. I-2738, at para. 14:

> Whilst it has . . . been recognised that the term "place where the harmful event occurred" within the meaning of Article 5(3) of the [Brussels Regulation] may cover both the place where the damage occurred and the place of the event giving rise to it, that term cannot be construed so extensively as to encompass any place where the adverse consequences can be felt of an event which has already caused damage actually arising elsewhere.

than the Brussels Regulation, will determine jurisdiction so long as it is sufficiently clear that the clause was intended to apply to claims for breach of competition rules (Article 23). However, the English High Court interpreted this Article narrowly in the 2003 *Provimi* judgment, discussed further *infra*, and its practical value is at present unclear.

Finally, even where a court does not have substantive jurisdiction to try a claim, it may still have jurisdiction to grant interim relief pending trial in another member state where there is a "real connecting link" between the member state and the requested relief (Article 31). This could be, for example, where the persons or assets subject to the requested relief are located within the jurisdiction. In such cases, however, jurisdiction must be established under national law, as Article 31 merely permits jurisdiction, but is not in itself a basis for jurisdiction.

b) Non-EU defendants

For claims brought against defendants not domiciled in the European Union, jurisdiction is determined according to national law. Most member states have jurisdictional rules similar to those in the Brussels Regulation. The most common grounds for establishing jurisdiction over non-EU defendants in EU member states include: (1) the defendant has property in the jurisdiction; (2) the defendant has a business or branch in the jurisdiction; (3) the conduct in question affected the market in the jurisdiction; (4) the defendant has been duly served in the jurisdiction; and (5) there are several defendants in connected claims, one of which is domiciled in the jurisdiction.[69]

2. Which Court?

Article 35 of Regulation 1/2003 leaves to member states the decision as to which national tribunals should be given competence to apply EU competition law. Member states' court systems and their allocation of competence over competition law claims among their various courts differ substantially. Some member states have created or assigned jurisdiction over competition law damages claims to specialized tribunals, while others have no special institutions or procedures and simply treat competition law cases like other tort damage actions. In some member states, jurisdiction differs depending on whether the claims are based on national law or EC law. Below is a summary of the systems in several member states.

a) England

Antitrust disputes in England and Wales are usually heard in the Chancery Division of the High Court. Follow-on actions for damages based on a prior finding of infringement by the European Commission or the Office of Fair Trading

69. *See* Comparative Report, *supra* note 1, at 41-42.

may also be heard in the Competition Appeals Tribunal ("CAT"), a specialized panel that hears competition disputes and conducts proceedings with a more flexible and case-managed approach.[70]

b) France

In France, as of January 1, 2006, exclusive jurisdiction over claims (including damages actions) based on competition law has been granted to specialized courts.[71] It remains unclear, however, whether follow-on damages actions must also be heard in these specialized courts. Only the Paris Court of Appeal may hear appeals in cases related to anticompetitive practices.[72] The French Supreme Court (Cour de Cassation) has set up economic training for judges to deal with competition matters.

c) Germany

The regional courts ("Landgericht") have competence over competition damage claims, irrespective of the amount of damages claimed.[73] In most Federal States ("Bundesländer"), this competence is granted to specialized regional courts for antitrust disputes, which have exclusive jurisdiction.[74]

d) Italy

In Italy, damages claims based on Italian competition law are heard at first instance in the courts of appeal, with decisions subject to review by the Court of Cassation on questions of law only. Claims based on EU law are heard in the lower civil courts (giudici di pace and tribunali).[75]

70. Competition Act 1998, at §§ 47A & 47B.

71. *See* Décret No. 2005-1756 du 30 décembre 2005 fixant la liste et le ressort des juridictions spécialisées en matière de concurrence, de propriété industrielle et de difficultés des entreprises, en application de l'Article L. 420-7 du Code de commerce. This decree lists eight courts of first instance and eight commercial courts that have exclusive jurisdiction over claims relating to anticompetitive practices as of January 1, 2006. The specialized courts are located in Paris, Marseille, Bordeaux, Lille, Fort-de-France, Lyon, Nancy, and Rennes.

72. Criminal courts have been empowered to award damages for breach of specific competition law provisions, applying in such cases the same statutory provisions as civil judges, but it is unclear whether such powers continue since the establishment of the specialized civil courts.

73. GWB, at § 87.

74. *Id.* at § 89.

75. Italian Competition Act, at § 33.

e) Spain

The new Competition Act, which entered into force on September 1, 2007, equates the treatment of EC and national competition law claims before Spanish courts by extending the competence of the specialized commercial courts ("Juzgados de lo Mercantil") to hear claims for damages based on national competition law. Previously, these courts could only hear claims for damages based on EU competition law; claims based on national law had to be "follow-on" actions rather than stand-alone claims, and had to be filed before the ordinary civil courts. The new Act allows claimants to file damage actions on a stand-alone basis or as a follow-on action, even prior to the conclusion of appeals to decisions by the Spanish competition authorities.

f) Sweden

All Swedish courts have jurisdiction to hear damages claims for competition law violations, but as confirmed recently,[76] only the Swedish Competition Authority, the Stockholm District Court, and the Market Court (on appeal) have authority to issue decisions prohibiting competition law violations.

3. Scope for Forum Shopping?

The jurisdictional rules outlined above leave open the possibility that any of several member states may be an appropriate forum for private EU competition law litigation, which inevitably invites a degree of forum shopping. Different procedural factors such as the duration and cost of litigation, the availability of disclosure, the possibility of collective actions, and the likelihood of obtaining interim relief may militate in favor of one jurisdiction over another. This can leave defendants uncertain as to where claimants might try to establish jurisdiction, and can also require defendants to litigate in an effort to resist jurisdiction where claimants bring cases with a weak jurisdictional basis. Yet it also provides scope for pre-emptive forum selection by potential defendants.

Two recent judgments by courts in England are particularly instructive on these issues.

a) Consolidating EU-wide claims in one member State: The Provimi *case*

A 2003 judgment by the English High Court represents a milestone for plaintiffs (particularly in follow-on cartel litigation) as the court allowed the consolidation of Europe-wide damages claims in a single proceeding even though some of the plaintiffs and some of the defendants were foreign, some of the transactions in question took place outside England, and some of the claims related to injury suffered outside England.

76. Göta Hovrätt Ö 638-07, Skyways Express / Norrköpings kommun genom Marknadsföringsbolaget Nya Norrköping (Apr. 18, 2007).

The *Provimi* case involved claims brought by vitamin purchasers against Hoffmann-La Roche and Aventis in the wake of the European Commission's infringement decision in the vitamins cartel.[77] The claimants were two English companies and a German company, each of which was a direct purchaser from the defendants. The court first established the jurisdiction of the English courts in relation to claims involving the defendants domiciled in England on the basis of Articles 2(1) and 5(3) of the Brussels Regulation (as well as corresponding provisions of the Lugano Convention[78]). Article 2(1) established jurisdiction over the English defendants in the proceedings. Regarding Article 5(3), which provides that the defendant can be sued "where the harmful event occurred," the court was prepared to assume that where the claimant is domiciled in England, the harmful event occurred in England.[79]

The court's treatment of non-U.K. claims potentially has the most far-reaching effect. The German claimants had purchased from German subsidiaries of the defendants (Swiss and French companies), but had not made any purchases from U.K. subsidiaries. These claimants brought action in the United Kingdom against the U.K. subsidiaries, seeking damages for loss suffered outside the United Kingdom. The defendants argued that the U.K. courts lacked jurisdiction under the Brussels Regulation since neither the direct sellers nor their parent companies were U.K. domiciled (failing the Article 2(1) test) and the harmful event and loss had occurred in Germany where the purchases had been made (failing the Article 5(3) test).

On an application for strike out (where the legal threshold is whether the claimant has an arguable case), the court found against the defendants based on the following reasoning. The concept of an "undertaking" under EU competition rules refers to an economic unit that is wider than a corporate entity. Thus, where a parent undertaking commits an infringement of Article 81 EC and a subsidiary implements that infringement by charging a cartelized price—even if unknowingly (as here, where the parent companies set the prices charged by their U.K. subsidiaries)—the subsidiary arguably also commits an infringement and may be sued for all losses flowing from it. Furthermore, the court found it a triable issue whether the infringement by the U.K. subsidiary caused loss to the non-U.K. claimant—even though there had been no direct commercial relationship between the two—because, but for its participation in the cartel through charging cartelized prices, the U.K. subsidiary arguably would have offered products at

77. Commission decision of 21 November 2001, Case COMP/E-1/37.512 - Vitamins, 2003 O.J. (L 6) 1 [hereinafter *Vitamins*].

78. European Union, The Lugano Convention on Jurisdiction and the Enforcement of Judgments in Civil and Commercial Matters (Sep. 16, 1988).

79. *Provimi*, *supra* note 64, at para. 10.

lower prices that would have been available to the claimant.[80] Jurisdiction of the non-U.K. claimant to sue the U.K. subsidiary defendants was thus established.

Additionally, the court held it had jurisdiction over claims by the non-U.K. claimant against the non-U.K. co-defendants under Article 6(1) of the Brussels Regulation, finding that the claims at issue (all of which involved the same essential facts) were closely connected to the claims against the U.K. defendants and that, in view of the undeveloped state of the law in the area of private damages actions for competition law infringements, there was a risk of irreconcilable judgments if the non-U.K. claims were heard in separate proceedings in Germany or France.[81] This judgment potentially opens the doors to the English courts for foreign claimants against any defendant undertaking with a U.K. subsidiary that charged "cartelized" prices, even if unknowingly.

Also importantly, the court held that jurisdiction clauses in certain of the defendants' standard terms and conditions that would have given exclusive jurisdiction to courts in Switzerland, France, and Germany under Article 23 of the Brussels Regulation were not sufficiently broad to cover the claimants' tort claims for damages under Article 81 EC—notwithstanding the broad language of certain of the clauses, which applied to "all disputes arising out of the legal relationship." This effectively means that jurisdiction clauses that do not specifically mention potential competition law damages actions are unlikely to be controlling as regards jurisdiction, notwithstanding the apparent clarity of Article 23. Given the obvious commercial reality, such clauses are unlikely to be included in contracts. Accordingly, contractual jurisdiction clauses are unlikely to have much effect on cartel damages claims, at least in the United Kingdom.

b) Challenging jurisdiction: The SanDisk *case*

At the other end of the spectrum from *Provimi*, another recent judgment from England illustrates that the Brussels Regulation does not confer unlimited choice of jurisdiction on claimants. In *SanDisk v. Sisvel*,[82] a U.S. claimant brought an action in the English High Court against non-U.K. defendants for alleged anticompetitive conduct relating to the licensing and enforcement of certain patents. As the defendants were domiciled in Italy, France, Germany, and the Netherlands, the English courts lacked jurisdiction under Brussels Regulation Article 2(1). SanDisk argued that the court had jurisdiction under Article 5(3), maintaining that the defendants' conduct (a series of alleged abuses including tying essential and non-essential patents, misuse of the patent system, and employing sham legal actions including obtaining seizures in the Netherlands,

80. *Id*. at paras. 37-42.

81. *Id*. at paras. 43-49.

82. SanDisk Corporation v. Koninklijke Philips Electronics NV & Ors, 2007 E.W.H.C. 332 (Ch).

Germany, and Italy of SanDisk products that were allegedly in violation of the defendants' patents) amounted to a campaign of anticompetitive behavior pursued throughout Europe, including in the United Kingdom. SanDisk also requested interim relief (an order preventing the defendants from initiating patent enforcement actions against SanDisk that would prevent SanDisk from carrying on its business without giving prior warning to SanDisk) under Article 31 of the Brussels Regulation.

The court ruled against SanDisk on all counts, finding that it lacked both substantive and interim jurisdiction over the claims. The court emphasized that alternative forums (the defendants' domiciles) were clearly available to the claimant and cited repeatedly the obligation to narrowly interpret the exceptions to Article 2 narrowly.[83] Regarding Article 5(3), the judge concluded that because none of the decisive "first steps" of the alleged abuses took place in the United Kingdom and SanDisk did not suffer immediate damage in the United Kingdom by reason of any of these abuses, the Article 5(3) test was not satisfied. Regarding interim jurisdiction, because the substance of the complaint and the interim relief sought lay in the enforcement measures taken by the defendants against SanDisk in other countries, the United Kingdom lacked the necessary "real connecting link" to the alleged abuse. Courts in the member states where seizure orders against SanDisk products had been obtained were "beyond doubt in the best position to decide what, if any, measures of warning [of future infringement actions] it is appropriate for Sisvel to give SanDisk."[84]

This case offers several lessons. First, the appearance created by judgments like *Provimi* notwithstanding, choice of jurisdiction under the Brussels Regulation is not without limit and, particularly, the potentially broad Article 5(3) "place of harm" provision will not always be given an expansive reading. Second, potential claimants need to weigh the risk that by bringing actions only tenuously related to the United Kingdom in an effort to gain benefit of the advantageous U.K. civil procedure rules, entire claims may be struck out before even getting to the substance of the case. In assessing the jurisdictional arguments, the judge in *SanDisk* also gave preliminary consideration to the substantive merit of the alleged abuses, since in order to assume jurisdiction under Article 5(3), the court must first find that the claimant has demonstrated that it has a "good arguable case." The judgment concludes expressly that SanDisk's claims "cannot so be described,"[85] which will no doubt handicap any efforts by SanDisk to pursue the claims in another jurisdiction. Third, the case highlights the burden that even successful defendants will face in resisting jurisdiction by forum-shopping claimants. The defendants here will have incurred large legal bills, perhaps only

83. *See also supra* note 65.

84. *Id*. at para. 53.

85. *Id*. at para. 41.

60 percent of which may have been recoverable from the unsuccessful claimant under the U.K. costs rules.

c) Pre-empting choice of jurisdiction

The various jurisdictional possibilities afforded by the Brussels Regulation not only provide claimants with forum shopping opportunities, but also open the way for potential defendants to steer anticipated litigation away from claimant-friendly jurisdictions. For both claimants and defendants, there can be significant advantage in moving quickly to establish jurisdiction.

The various jurisdictional possibilities afforded by the Brussels Regulation not only provide claimants with forum shopping opportunities, but also open the way for potential defendants to steer anticipated litigation away from claimant-friendly jurisdictions. For both claimants and defendants, there can be significant advantage in moving quickly to establish jurisdiction.

Claimants will want to establish jurisdiction where the rules are favorable and it is convenient for them to litigate. Article 27 of the Brussels Regulation provides that once jurisdiction is established in one member state, courts in other member states must decline jurisdiction in any subsequent proceedings brought in the same action. Courts may also choose to stay subsequent proceedings in related cases (Article 28). Thus, in cases involving multiple claimants, the first claimant to litigate may have a choice of jurisdiction, while subsequent claimants may have none.

Similar considerations apply from the defendant's perspective. A company that anticipates being sued for damages (e.g., because it anticipates receiving a decision finding it guilty of cartel behavior) can seek a preemptive negative declaratory judgment against potential claimants in the jurisdiction of its choice. The company would bring an action against a potential claimant seeking a declaration that there has been no infringement or, more likely, no damages. Potential claimants from other member states could be brought into the action under Brussels Regulation Article 6(1), which, as noted above, provides that where there are multiple defendants, a claim may be brought in the member state in which any one of them is domiciled, provided that the claims are so closely connected that it is expedient to hear and determine them together to avoid the risk of irreconcilable judgments. Such an action could allow a potential defendant to have claims against it heard in a jurisdiction with restrictive discovery rules or where the passing-on defense is recognized. Faced with such an action, a claimant may find that it is unable to bring a subsequent damages claim in another member state, as under Articles 27 and 28 the courts decline jurisdiction or stay proceedings pending the outcome of the declaratory action. The main obstacle to defendants' use of this tactic may be more practical than legal, as it effectively requires a company to sue its customers—inevitably straining those relationships.[86]

4. What Law Applies?

Until recently, there had been some ambiguity with respect to which country or countries' laws should apply in damages actions involving harm in more than one member state. Regulation 864/2007, which comes into force on January 11, 2009, clarifies the issue by harmonising EU-wide conflict-of-law rules for non-contractual claims (including tort claims based on competition law violations).[87] Under the Regulation, the general rule is that the court should apply the law of the country where the damage occurred.[88] The Regulation also offers a specific rule with respect to multi-jurisdictional competition law claims: when the defendants' conduct allegedly affected more than one country, a claimant who sues in the court of the defendant's domicile (pursuant to the general rule of the Brussels Regulation) may choose to base the entire claim on the law of the court seized, provided that the market in that member state is among those directly and substantially affected by the competition law violation.[89] This clarifies that under the specified conditions, the court may apply its own law to the entirety of the claim, regardless of where the damage occurred—as opposed to applying its own law in respect of damage in its territory and the law of another member state in respect of damage incurred there.

5. Limitation Periods

Limitation periods can have a decisive effect on the availability of damages in many typical circumstances. For instance, a short limitation period could time-

86. Another possible adverse consequence is the effect on discovery in other pending litigation. Following the European Commission's 2006 decision in the Synthetic Rubber cartel (Commission decision of 29 November 2006, Case COMP/38.638 - Synthetic Rubber (not yet reported)), in which (among others) the Italian firm Eni was fined EUR 272 million, Eni launched an action for negative declaratory judgment in the Court of Milan, seeking to have the judge ascertain that the unlawful behavior, if any, had not had an impact on the market so that no damage claims could be brought against it. As part of that proceeding, Eni disclosed a non-confidential version of the Commission's statement of objections ("SO"). During the same time period, in litigation before the European Court of First Instance, Eni had: (1) sought to prevent Michelin (a customer) from using the SO, which it had received as a third party to the Commission's cartel investigation, against Eni in follow-on damages litigation; and (2) challenged the Commission's decision to send the SO to Michelin in the first place. In July 2008, the CFI ruled that, because Eni had since disclosed the SO in the Milan proceeding, Michelin was free to access and use the SO without restriction. The Court then declined to rule on the fundamental question whether the Commission had been right to send the SO to Michelin as there was no longer reason to adjudicate the issue. Order of the Court of First Instance of 2 July 2008, Case T-12/07, Polimeri Europa v. Commission, 2008 O.J. (C 223) 73.

87. Regulation (EC) No. 864/2007 of the European Parliament and of the Council of July 11, 2007 on the law applicable to non-contractual obligations (Rome II), 2007 O.J. L 199/40.

88. Id., Article 4.

89. Id., Article 6(3)(b).

bar follow-on actions in the wake of an infringement decision by a regulator or a successful test case. Member state rules on limitation periods vary widely as regards both the length of the relevant time period (ranging from one to 30 years) and, equally important, the triggering event. Limitation periods in some member states start to run from the date on which the infringement occurred, irrespective of the claimant's knowledge, while in other member states the clock starts ticking only once the damage was, or should have been, detected by the potential claimant. Some member states have rules combining both subjectively and objectively fixed limitation periods. This diversity of approaches means that liability will extinguish at different times across the European Union, even in respect of the same infringement. Liability must therefore be assessed on a state-by-state basis to determine potential exposure.

Below are some examples of different member states' rules concerning limitation periods, which illustrate the diversity of approaches among the member states.

a) Austria

Damages claims in Austria are time-barred three years from the date on which the claimant becomes aware of the damage and of the identity of the person responsible for such damage. If the claimant is not aware of the damage or of the person responsible for such damage, the applicable time period is 30 years from the date on which the damage was incurred.[90]

b) Belgium

The limitation period for actions based on tort law is 5 years from the day on which the injured party became aware of the damage, or, in any event, 20 years from the date on which the infringement occurred.[91]

c) England

In England, cases must be brought before the courts within six years from the date on which the cause of action accrued (the date on which it is reasonable to conclude that damage has been or will be suffered, having regard to any deliberate concealment by the defendant). In follow-on actions before the CAT, cases must be brought within two years following the later of the completion of any appeal (or the lapsing of any appeal period) and the date on which the cause of action accrued.[92]

90. Austrian General Civil Code (Allgemeines Bürgerliches Gesetzbuch), at § 1489.

91. BCC, at art. 2262bis.

92. The CAT recently ruled that this two-year period cannot be extended by mutual agreement of the parties, although the CAT has discretion to extend the time limit. Judgment of 17 October 2007, Emerson Electric Co. & Ors. v. Morgan Crucible Company plc, 2007 C.A.T. 28.

d) France

In France, the limitation period expires 10 years after the claimant became aware of, or should have become aware of, the damage.

e) Germany

The limitation period for private competition damages claims under German law is 3 years, beginning at the end of the calendar year when the plaintiff was injured and the injured party knew or should have known of the circumstances giving rise to the claim. The limitation period is tolled as of the date on which any European competition authority commences an investigation or proceedings for an infringement, and does not resume until six months after the competition authority's case has been decided or proceedings have otherwise been concluded.[93] In any event, the limitation period cannot be later than 10 years after the damage was incurred or 30 years after the date of the infringement.

f) Italy

In Italy, private damage claims based on competition law infringements are governed by both tort and contract law. The limitation period expires 5 years (in respect of tort claims) or 10 years (in respect of contract claims) after the plaintiff first knows or reasonably should have known of the injury, as well as of its unjust nature (i.e., that the harm was caused by of a breach of the competition rules). Thus, the limitation period might start running from the publication date of a decision by the Italian Competition Authority to open an investigation or a court judgment concerning an infringement.[94]

g) Spain

In Spain, proceedings must begin within 1 year from the date on which the injured party discovers the damage or an infringement decision is adopted.

The European Commission is concerned that limitation rules in some member states can act as a barrier to the recovery of damages, and the White Paper accordingly makes two suggestions towards harmonizing limitation periods:[95]

- first, in the case of a continuous or repeated infringement, the limitation period should not start to run before the day on which the infringement ceases or before the victim of the infringement can reasonably be expected to have become aware of the infringement and of the harm it caused him; and

93. GWB, at § 33(5).

94. Case No. 2305/2007, SAI v. Nigriello (Court of Cassation, 2007).

95. White Paper, *supra* note 5, at § 2.7.

- second, member states should remain free to set their own limitation periods with reference to stand-alone actions, but in case of follow-on actions, a *new limitation period* of at least two years should start once the infringement decision on which a follow-on claimant relies has become final.

The latter rule would give injured parties the advantage of preserving the right to claim damages until the initial public enforcement proceedings have run their course, thus eliminating the burden and risk of having to prove the existence of the infringement in a stand-alone action.

C. PROOF OF INFRINGEMENT

In all member states, as is normally the case in all civil damages actions, successful claimants in competition law cases must prove the existence of an infringement to the requisite legal standard. As a practical matter, the requirement to establish the infringement divides the universe of potential competition law damages claims into two categories: stand-alone actions and actions following on from a prior finding of infringement.

1. Proof of infringement in a stand-alone action

In a standalone action alleging a competition law infringement, the claimant is in the customary position of a tort plaintiff: having to prove the existence of the infringement itself before the question of damages will be addressed. In most member states, as under Article 2 of Regulation 1/2003, the burden lies with the claimant to prove that the conduct in question had an appreciable effect on competition and therefore infringes the competition rules. Such proof often requires detailed evidence of both the defendant's specific conduct and of competitive conditions in the relevant market, making access to evidence critical. Without benefit of the resources and investigatory powers of a competition authority, claimants in stand-alone actions must fall back on the possibilities offered by national discovery rules, many of which are highly restrictive and effectively require the claimant to have sufficient evidence to discharge the burden of proof even before launching an action. For this reason, the burden of proving an infringement can represent a significant obstacle for potential claimants in stand-alone actions.

a) Discovery/disclosure

The majority of member states follow the civil law tradition, which does not embrace the concept of disclosure of documents between the parties in civil litigation. While the courts retain powers to order production of documents, the parties' ability to compel production of documents is limited. Lack of access to evidence in these jurisdictions substantially impairs the claimant's ability to prove an infringement in a stand-alone action.

The European Commission has cited lack of disclosure as one of the major obstacles to private enforcement. The White Paper follows proposals first tabled in the Green Paper, suggesting that member states adopt special rules expanding the possibilities for claimants to obtain documentary evidence from third parties in EU competition law actions for damages. In particular, the Commission proposes granting national courts the power to order parties to proceedings (or third parties) to disclose precise categories of relevant evidence, provided that the plaintiff:

- has presented all the facts and provided evidence reasonably available to him and that these show plausible grounds to suspect that he has suffered harm from an antitrust infringement committed by the defendant;
- has shown that despite all efforts, without the discovery order he would not be able to produce or obtain the requested evidence;
- has specified sufficiently precise categories of evidence to be disclosed; and
- has satisfied the Court that the evidence requested is both relevant to the case and necessary and proportionate.

The White Paper states that such a "fact-pleading" disclosure regime, under strict judicial control, would assist in overcoming the inherent information asymmetry that disadvantages plaintiffs, while still preventing so-called "fishing expeditions"[96] and "discovery blackmail."[97]

The White Paper recommends further that national courts should be granted powers to impose sufficient sanctions to deter the destruction of relevant evidence or refusal to comply with a discovery order. It also highlights the importance of granting adequate protection from discovery to corporate statements by leniency applicants[98] and to the investigations of competition authorities.

96. *See, e.g.,* Working Paper, supra note 6, at note 39 ("[A] strategy to elicit in an unfocused manner, through very broad discovery requests, information from another party in the hope that some relevant evidence for a damages claim might be found.").

97. *See, e.g., id.* at note 40 ("[A] strategy to request very broad discovery measures entailing high costs with the intention to compel the other party to settle rather than to continue the litigation, although the claim or the defense may be rather weak or even unmeritorious.").

98. The White Paper recommends that all corporate statements submitted by a leniency applicant under Article 81, regardless of whether the application for leniency is accepted, is rejected, or leads to no decision by the competition authority, should be protected from disclosure. In a related point, the White Paper proposes for further consideration a rule whereby the immunity recipient's civil liability would be limited to claims by his direct and indirect contractual partners. White Paper, *supra* note 6, at § 2.9.

These ideas have been met with resistance from several member states. However, the Commission views this as a critical issue. Absent broader discovery rules, stand-alone actions will remain difficult and risky to bring (particularly in the civil law member states)—yet without stand-alone actions, it is difficult to see how the overarching goal of making private actions an important complement to public competition law enforcement will be achieved.

In common law jurisdictions such as England and Wales (as well as Ireland and Cyprus), the discovery issue is seen as less pressing since pre-trial disclosure obligations are an established part of civil litigation. These rules ensure that parties have access to any documents on which the other party intends to rely, as well as documents that adversely affect that party's case or support another party's case. In some circumstances, disclosure may even be obtained at the pre-action stage (in order to avoid proceedings) or against third parties. These rules remain less expansive than U.S.-style discovery, as there are no depositions, the scope for discovery against third parties is more restricted and the range of documents required to be disclosed is more limited. Nevertheless, in the absence of serious reform in the civil law jurisdictions, the possibility of obtaining important documentary evidence from (in particular) defendants will inevitably draw potential competition law claimants, particularly those with stand-alone claims, toward the common law jurisdictions.

Follow-on actions are likely to be less risky for claimants to bring and will doubtless represent a substantial proportion of all private competition law damages claims in Europe over the coming years.

2. Proof of infringement in a follow-on action

In a follow-on action, the burden of proving an infringement is substantially alleviated since the claimant can take advantage of the already-existing competition authority decision finding an infringement. When a claimant can gain benefit of a presumption of infringement based on a prior decision, the damages action will effectively be limited to the issues of causation and quantification of damages, significantly easing the claimant's burden. For this reason, follow-on actions are likely to be less risky for claimants to bring and will doubtless represent a substantial proportion of all private competition law damages claims in Europe over the coming years.

a) EU decisions

In actions following on from prior European Commission infringement decisions, the situation is clear: Commission decisions are binding on national courts as to the existence of an infringement. According to Article 16 of Regulation 1/2003, in order to avoid the risk of conflicting decisions, national courts "cannot take decisions running counter to the decision adopted by the Commission." In cases where the Commission has initiated proceedings and a decision is therefore contemplated, national courts must also avoid giving conflicting decisions, which may create an obligation to stay any national proceedings on the same matter.

The extent to which Commission decisions that are subject to appeal are binding on national courts as regards the existence of an infringement is slightly less clear. The general rule is that Commission decisions under appeal are not binding, but that in most cases it will be appropriate for the national courts to stay proceedings pending the final resolution of appeals against the Commission decision.[99] The rationale for this rule lies in the need to avoid conflicting decisions. It is at least arguable, however, that there is no such risk if the pending appeal of the Commission decision does not relate to the existence of an infringement—as may well be the case when, for example, a cartel leniency applicant does not contest the issue of liability but appeals the Commission's calculation of the fine. A claimant who sought to use the Commission decision merely as evidence of the existence of an infringement may have strong arguments that there is no risk of conflicting decisions and thus no need for the court to stay proceedings pending the potentially lengthy appeals process through the European courts.[100]

The U.K. Competition Appeals Tribunal recently addressed a number of related jurisdictional and evidentiary issues in the *Morgan Crucible* litigation. The case involved a follow-on action under s47A of the Competition Act brought by Emerson Electric and others against Morgan Crucible ("Morgan"), which had applied for and received immunity from fines under the European Commission's leniency program in connection with the carbon and graphite products cartel.[101] The Commission decision had been appealed to the European Courts by several defendants, but not Morgan (which had received immunity from fines). The CAT first ruled that the damages claim could only be brought upon the express permission of the CAT, regardless of the fact that Morgan was not party to the

99. *See Masterfoods*, *supra* note 22, at para. 57:

> When the outcome of the dispute before the national court depends on the validity of the Commission decision, it follows from the obligation of sincere cooperation that the national court should, in order to avoid reaching a decision that runs counter to that of the Commission, stay its proceedings pending final judgment in the action for annulment by the Community Courts, unless it considers that, in the circumstances of the case, a reference to the Court of Justice for a preliminary ruling on the validity of the Commission decision is warranted.

100. *See* Iberian UK Ltd. v. BPB Industries plc, 1996 C.M.L.R. 601 (English High Court, 1996), at para. 69 ("Except in *the clearest cases of breach or non-breach*, it will be a proper exercise of discretion to stay proceedings here to await the outcome of the European proceedings." (emphasis added)). The claimant would argue that when an appeal of a Commission decision does not contest the existence of an infringement, the decision represents a sufficiently clear precedent on that issue even pending the appeal.

101. Commission decision of 28 April 2004, Case COMP/38.359 - Electrical and mechanical carbon and graphite products, 2004 O.J. (C 102) 6.

appeal—confirming that the existence of any related appeal was sufficient to prevent a claimant from bringing a damages claim without the CAT's permission.[102] The claimants then sought the CAT's permission to proceed against Morgan and, in a subsequent judgment, obtained permission to proceed.[103] In granting permission, the CAT rejected Morgan's attempt to reject the claim as unfounded and focused on the claimants' concerns that, were their claim not to be allowed to proceed pending the outcome of appeals in the European Courts, there was a reasonable prospect that critical documents relating to Morgan's conduct would not be preserved, and would therefore be unavailable for discovery.[104]

Thus, although the claimants prevailed, the CAT's reasoning suggests that permission to proceed against a cartel immunity recipient pending appeals by other defendants will be granted only in exceptional circumstances. Defendants will probably be able to postpone such actions by giving suitable assurances to the claimants that evidence will not be destroyed pending the outcome of the appeals. The decisions thus provide a measure of reassurance that immunity applicants will not open themselves immediately to joint and several liability for damages caused by a cartel; a more lenient approach toward the claimants would have represented a serious blow to immunity programs throughout the European Union.

b) National decisions

Due in part to a lack of precedent, the situation in respect of actions following on from infringement decisions by national regulators is slightly less clear. In the United Kingdom, prior decisions of the Office of Fair Trading (OFT) and Competition Appeals Tribunal (CAT) are binding on national courts as proof of infringement once appeals have been exhausted.[105] In Germany, prior decisions of any EU antitrust authority, as well as judgments by national courts as to the validity of such decisions, are binding as proof of infringement.[106] Similar rules apply in Belgium.[107] Other member states do not make prior decisions of a com-

102. Judgment of 17 October 2007, Emerson Electric Co. & Ors. v. Morgan Crucible Company plc, 2007 C.A.T. 28.

103. Judgment of 16 November 2007, Emerson Electric Co. & Ors. v. Morgan Crucible Company plc, 2007 C.A.T. 30.

104. In support of this conclusion, the CAT referred specifically to "a previous history of destruction of documents by Morgan Crucible."

105. United Kingdom, Competition Act 1998, at §§ 47A & 58A.

106. GWB, at § 33(4).

107. Belgian courts are precluded from taking decisions contrary to Belgian Competition Council decisions. LPEC (at art, 11, § 1) expressly states that the Belgian Competition Council is an administrative court. Accordingly, under Belgian law, the Council's decisions have autorité de chose jugée (i.e., are binding) unless and until such decisions are annulled by the Brussels Court of Appeals—the only

footnote 107 cont'd on next page

petition authority binding, but national courts are likely to regard such decisions as persuasive evidence of the existence of an infringement.

The White Paper proposes to clarify this situation by extending the same precedential value that is accorded to Commission decisions to infringement decisions by national competition authorities ("NCAs") that are members of the European Competition Network. The Commission thus proposes that final infringement decisions taken by an NCA under Article 81 or 82, and final judgments by review courts upholding those decisions, should be accepted in every member state as irrebuttable proof of the infringement in subsequent actions for damages.[108] Such a rule is logical but will test the faith of member states in the competence and diligence of their fellow countries' competition authorities and courts.

The proposed rule would apply only to NCA decisions that are final (where the defendant has exhausted all appeal avenues) and that relate to the same practices and same undertaking(s) concerned in the follow-on litigation. The Working Paper clarifies that binding effects should only be granted to decisions relating to:

> "(i) the same agreements, decisions or practices that the NCA found to infringe Article 81 or Article 82 EC, and (ii) to the same individuals, companies or groups of companies which the NCA found to have committed this infringement (normally, the addressee(s) of the decision)."[109]

c) Precedential value of decisions in similar cases

The Regulation 1/2003 rule on the binding effect of Commission decisions applies only to decisions that address the same issues and the same parties as those before the national court. This raises the question whether decisions dealing with facts that are merely similar to those at issue before a court will also be binding. The U.K. House of Lords considered this issue in the *Crehan* case,[110]

footnote 107 cont'd
court that has jurisdiction to hear appeals on Council decisions (LPEC, at art. 75). This implies that the parties to the proceedings before the Belgian Competition Council are bound by the Council's infringement decision, including in subsequent proceedings involving the same parties and the same practice before the courts. This is the case even if the decision is appealed, since the appeal, in principle, is not suspensive (LPEC, at art. 76, § 4). Decisions by other national competition authorities do not have specific evidentiary value, but may carry some weight with the courts.

108. White Paper, *supra* note 6, at § 2.3.

109. Working Paper, *supra* note 6, at para. 154.

110. Inntrepreneur Pub Company and others v. Crehan, 2006 U.K.H.L. 38.

holding that claimants will not be entitled to "piggy back" a private damages action on a prior competition authority infringement decision where the prior decision deals with a different—even if very similar—situation.[111] The House of Lords explained that while it is clear that conflicting decisions must be avoided (citing *Delimitis*[112] and other precedents), there is no risk of such conflict where the legal and factual context of the case that was examined by the Commission is not completely identical to that before the national court (citing the Opinion of Advocate General Cosmas in the *Masterfoods* case[113]). This is consistent with the approach proposed in the White Paper.

Based on this judgment, while a prior Commission infringement decision in a similar case may be admitted as evidence, and may even be highly persuasive, it will not constitute binding proof of the infringement and the court will need to reach its own determination on that issue. In such instances, defendants will be entitled to argue that the similar case at issue should be decided differently—in effect challenging the findings of the Commission to the extent the cases are not distinguishable. Factors affecting the amount of weight to be accorded to the Commission decision would include: whether the decision was addressed to a party against whom the decision is relied on in the national court proceeding; if not, whether that party had an opportunity to participate in the Commission proceedings; in the case of findings of fact, whether these were essential or non-essential to the Commission's conclusions; and whether the findings of fact related to the same time period at issue in the national case or were otherwise on the same subject.

The effect of this judgment cuts both ways. On one hand, while courts may regard prior similar decisions as highly persuasive, the fact that the court will need to reach its own determination on the issue of infringement may nonetheless raise the evidentiary burden for claimants. In the knowledge that defendants are likely to argue that the prior decision should not apply to the new case, claimants will need to consider whether they will be able to adduce evidence in addition to the prior decision that will persuade the court that an infringement occurred. On the other hand, this approach also means that claimants are not necessarily barred from commencing proceedings where there is a prior similar decision finding non-infringement, as the claimant may seek to persuade the court that the regulator's decision is inapplicable, either because the facts are different or because the conclusion was wrong.

111. Specifically, the House of Lords held that the national judge was not bound to apply against the defendant Inntrepreneur prior to European Commission findings against other U.K. brewers that their similar pub lease agreements had the effect of hindering access to the market, in contravention of Article 81(1) EC.

112. Case C-234/89, Delimitis v. Henninger Bräu AG, 1991 E.C.R. I-935.

113. Opinion of Advocate General Cosmas in *Masterfoods*, *supra* note 22.

3. Potential Sources of Evidence

In addition to the regular discovery/disclosure avenue discussed above, claimants in follow-on actions may have access to other sources of evidence deriving from the regulator's investigation. Most straightforwardly, the findings of fact in Commission or national authority decisions will be of use to claimants. It is also possible for claimants to obtain access to the actual case files held by competition authorities.

a) Access to European Commission case files

European Commission case files will contain a wealth of evidence potentially of value to claimants in related damages actions, including documents provided voluntarily to the Commission (e.g., leniency applications and supporting materials), documents obtained by the Commission under compulsory process (e.g., documents seized in "dawn raids," responses to information requests), statements of objections, and internal Commission reports and analyses. By law, access to Commission case files, excluding business secrets and confidential information, is granted only to addressees of statements of objections[114] and to complainants whose arguments are rejected by the Commission (and, in the latter case, only in respect of the documents on the basis of which the complaint was rejected).[115] These rules do not as a general matter grant file access to any interested party (e.g., customers of firms found to have participated in a price-fixing cartel), even if they have participated in the Commission's proceedings as a complainant. Complainants do, however, have a right to receive a non-confidential version of the Commission's statement of objections, with a view to facilitating the complainant's participation in the ongoing proceeding.[116]

Special rules apply to access to the Commission's case file in respect of leniency applications. The Commission's 2006 Leniency Notice expressly recognizes that

> "normally public disclosure of documents and written or recorded statements received in the context of this notice would undermine certain public or private interests, for example the protection of the purpose of inspections and investigations … even after the decision has been taken."[117]

114. Commission Regulation (EC) No. 773/2004 of 7 April 2004 relating to the conduct of proceedings by the Commission pursuant to Articles 81 and 82 of the EC Treaty, 2004 O.J. (L 123) [hereinafter "Implementing Regulation"], at art. 15.

115. *Id.* at art. 8.

116. *Id.* at art. 6(1).

117. Commission Notice on Immunity from fines and reduction of fines in cartel cases, 2006 O.J. (C 298) 17, [hereinafter "Leniency Notice"], at para. 40.

Accordingly, third-party access to leniency applications is subject to additional restrictions intended to prevent public disclosure. First, only addressees of the statement of objections are granted access to other parties' leniency statements; complainants are not given access. Second, in practice leniency statements are often made orally, and third parties that get access to the file are not allowed to take copies of such statements (although they can make their own transcripts of oral statements). Third, companies that get access to leniency statements are only entitled to use the information contained therein "for the purpose of judicial or administrative proceedings for the application of the Community competition rules at issue in the related administrative proceedings," with sanctions provided for the misuse of such information.[118] The latter restriction appears stricter than the general restriction on use of information (other than leniency statements) obtained from the Commission file, which may be used only "for the purposes of judicial or administrative proceedings for the applications of Articles 81 and 82 of the Treaty."[119] It is not clear whether this restriction contemplates use of the materials only in the proceeding itself or an appeal thereof, or whether using the materials as evidence in related private enforcement actions is permitted. Moreover, the sanctions for misuse of information that apply with respect to disclosed leniency applications do not extend expressly to other information obtained from the Commission's file.

Several cases pending before the European Courts will test the Commission's power to withhold documents in its case file from prospective damages claimants. An interesting example involves the Dutch government, which is contemplating bringing a damages action against members of the bitumen cartel that was fined €267 million by the Commission in 2006. The Dutch government was an indirect purchaser from the cartel, as it ultimately pays for road construction. The Dutch government asked the Commission for a copy of the full confidential version of the infringement decision to use in its damages case. The Commission takes the view that the Dutch government should be treated like any other damages claimant as regards access to its case file, and has declined to turn the decision over. The Dutch government has responded by applying to the Court of First Instance for an order against the Commission compelling disclosure.[120]

118. *Id.* at para. 34.

119. Implementing Regulation, *supra* note 110, at arts. 15(4) & 8(2).

120. Case T-380/08, Netherlands v. Commission. *See also* Case T-437-08, CDC Hydrogene Peroxide v. Commission (prospective damages claimant seeking access to an index of information gathered during the Commission's investigation of the hydrogen peroxide cartel); Case T-344-08, EnBW Energie Baden-Württemberg v. Commission (prospective damages claimant seeking information from Commission case file in the insulated gas switchgear cartel); Case T-399-07, Basell Polyolefine v. Commission (prospective damages claimant challenging Commission decision refusing to turn over information from its case file in the organic peroxides cartel).

b) Access to national competition authority case files

It is sometimes possible for claimants to obtain access to files held by NCAs. The approach varies among member states and is not always clear. Some illustrative examples are below.

(i) Belgium

In Belgium, documents gathered by the Belgian competition authority during the course of its investigations may be produced in court, by order of the court, but only for purposes of applying the Belgian Competition Act or Articles 81 or 82 EC.[121] According to commentators, the Belgian competition authority (or even the European Commission) could also be summoned to produce specific documentary evidence within the framework of a pending case, subject to safeguards to guarantee the protection of legitimate business secrets. This interpretation has, however, not been tested in the context of damages litigation.

(ii) England

In England and Wales, claimants may petition courts to order third-party disclosure against the OFT,[122] which could result in OFT case files being turned over to private claimants. However, the OFT has stated that it will view such requests as exceptional, that it will take all measures to avoid disclosing leniency documents, and that it will oppose third party disclosure applications in the form of a "fishing expedition" (e.g., if an application asked for disclosure of all documents submitted to the OFT by a person or all documents submitted in support of a leniency application, without further particularization).[123] The OFT is particularly concerned about protecting against the disclosure of documents provided in the context of a leniency application, since "if undertakings are discouraged from applying for leniency due to the risk of private actions, it is likely that a smaller proportion of cartels will be uncovered."[124] The OFT has recommended that a power be conferred on the Secretary of State to provide that "leniency documents, appropriately defined, are excluded from use in litigation without the consent of the leniency applicant."[125] The OFT defines "leniency documents" as "documents that are created for the purpose of the leniency application," which

---

121. Belgian Competition Act, at art. 84.

122. CPR, at rule 31.17. The court may make a third-party disclosure order only where the documents of which disclosure is sought are likely to support the case of the applicant or adversely affect the case of one of the other parties and disclosure is necessary in order to dispose fairly of the claim or to save costs.

123. OFT Discussion Paper, *supra* note 44, at paras. 6.8-6.10.

124. *Id.* at para. 9.4.

125. *Id.* at para. 9.5.

would seem to exclude from protection preexisting internal company documents that are provided to the OFT as part of a leniency application.

(iii) Germany

In Germany, claimants as interested third parties may have a right of pre-trial access to files of the Federal Cartel Office ("FCO") provided that: (i) the plaintiff's reasonable interests outweigh the legitimate interests of the wrongdoer (or any third party) in non-disclosure; and (ii) granting access does not imperil the FCO's investigation.[126] In an atmosphere of increasing support for private damages litigation in Germany, the claimant's interest in evaluating the scope of possible competition damages claims may well prevail over the interests of the defendant (or third party). The balance of interests, therefore, may often allow access to the FCO's file, subject to the protection of legitimate business secrets. Like the OFT, the FCO has taken the view that leniency applications should be protected, and, as such, does not grant access to them.[127] If leniency applications were disclosed to potential claimants, the full effectiveness of the German leniency program would be at risk, as cartel members would likely be deterred from coming forward and cooperating with the authority. The FCO's position on not granting access to leniency applications has not yet been tested in the German courts.

(iv) Ireland

Discovery is available to parties to litigation and third parties, provided they are subject to the jurisdiction of the Court. The parties must request discovery (and specify the precise category of documents sought and the reasons why such documents are required) and discovery must be "necessary for disposing fairly of the matter or for saving costs."[128] Discovery is available with respect to the Irish Competition Authority's files, subject to privilege. Discovery is not available, however, with respect to foreign competition authorities or the European Commission unless these authorities are parties to the action.

(v) Italy

In Italy, individuals have a general right of access to documents held by the public administration, but it remains unclear whether this right may be exercised by claimants in civil actions with respect to confidential documents in the competition authority's file.

126. Administrative Offence Law (Gesetz über Ordnungswidrigkeiten, OWiG), at § 46 (in connection with § 406e of the German Code of Criminal Procedure (Strafprozessordnung, StPO)).

127. *See* Notice No. 9/2006, Guidelines Leniency Program (Federal Cartel Office, 2006), at para. 22.

128. R. v. Secretary of State for Home Affairs Ex-parte Harrison, 1997 J.R. 113, 119.

On this issue, the White Paper merely states that "adequate protection [from disclosure] should be given to corporate statements by leniency applicants and to the investigations of competition authorities."[129] The Staff Working Paper explains further the Commission's view that leniency statements should be protected from disclosure in civil litigation both before and after the competition authority's decision has been taken.[130] With respect to evidence other than leniency statements in the possession of a competition authority, the Commission notes that in some circumstances disclosure to private parties may interfere with the authority's ongoing investigations. The Staff Working Paper accordingly suggests that member states consider a rule whereby courts should temporarily refrain from ordering disclosure against the Commission or an NCA if it is shown that a disclosure order would jeopardize an ongoing investigation (as might be the case, for example, in respect of internal company documents submitted along with a corporate statement in the leniency context).[131] While the Commission's recognition of the need to protect leniency applications from disclosure is welcome, it would have been useful if the Commission had also made clear (as the OFT has done) that the competition authority's file should not become the de facto principal source of evidence for claimants on "fishing expeditions" and that disclosure requests against a competition authority should be the exception rather than the rule.

c) Use of economic evidence to prove infringement

Given the increasing emphasis on economic analysis in competition law, the need to prove or disprove the existence of an infringement (in cases where the court is not bound by a prior decision) will often require the parties to adduce detailed economic evidence, typically in the form of expert reports, on several key issues: (i) the definition of the relevant market; (ii) an explanation of the dynamics of competition in that market and an assessment of market power or dominance; and (iii) a determination of the extent to which the impugned conduct has had an effect on competition (as well as (iv) an estimate of the level of damages suffered, which is discussed in the next section).

Conducting a detailed economic analysis is challenging in itself, particularly when the requisite data may not be readily available; an additional hurdle arises in deciding how best to use the resulting evidence in the litigation context. The difficulty lies in conducting a sufficiently robust empirical analysis while also presenting the methods, assumptions, and results in a manner that supports the legal argument and is intelligible to the court. This may involve a trade-off between

129. White Paper, *supra* note 5, at § 2.2.

130. Working Paper, *supra* note 6, at para. 118.

131. *Id.* at para. 119.

(i) relatively simple techniques, where the results may be more intuitive and thus more appealing to the court, but which may have more limited probative value, and (ii) more complex methods which, while potentially more robust, will inevitably require more data and be more difficult for the court to understand.

There are significant procedural differences between the member states regarding the use of expert evidence. While most member states permit parties to appoint experts, evidence from court-appointed experts (common in competition law cases in countries such as France and Italy) may be given greater weight. In England, experts are chosen by the parties, affording a greater degree of control. At the same time, however, the procedural rules impose a number of limits on this control: (i) the court must first give permission (by finding that the issues on which the expert intends to give evidence are matters that require expert evidence); (ii) experts owe an over-riding duty to the court (rather than to the parties that have appointed them and paid for the evidence), and their independence will be tested closely; and (iii) the data and modeling used by an expert must normally be provided to the opponent in advance of trial to enable them to replicate and test the evidence in preparation for cross-examination.

D. PROOF AND QUANTIFICATION OF DAMAGES

Once the existence of an infringement has been established, the two additional elements generally needed to support a claim are proof of damages suffered and a causal link between the infringement and those damages.

1. Causation

In any damages claim, it is incumbent on the plaintiff to prove that the infringement in question caused the damages suffered. The defendant is only liable for those damages that were caused by the unlawful action. A typical approach to the causation issue, applied in England and Wales, is to apply the so-called "but for" test: but for the infringement, would the claimant have suffered the loss?

Notwithstanding the straightforward appearance of this test, proof of causation in antitrust cases can be highly complex. The financial loss suffered by the victims of anticompetitive behavior will often consist of paying a supra-competitive price. The claimant in such circumstances has to show that a price rise was the consequence of the defendant's unlawful action. The defendant might respond that any price rise was caused by something different, such as normal market functioning or third party actions. As discussed below, additional complications can arise in jurisdictions that recognize the passing-on defense (which includes most member states). Proving a causal link might require reconstruction of a hypothetical marketplace free of the anticompetitive conduct at issue, which can call for a highly complex economic analysis.

Two recent French judgments illustrate how claimants may have difficulty proving the existence of a causal link between the infringement and the alleged damages suffered. First, the *Arkopharma*[132] case was a follow-on damages claim by a vitamins purchaser against Roche (a member of the vitamins cartel that had previously been sanctioned by the European Commission).[133] Arkopharma claimed that, since it had bought vitamins from Roche at prices affected by the cartel, it had suffered damage resulting from Roche's wrongfully increased prices. The Court dismissed the claim on two grounds, one of which was the claimant's failure to prove a causal link between the infringement (as to the existence of which the Commission's decision was dispositive) and the alleged damages. In particular, the Court found that Arkopharma had failed to demonstrate that it was unable to pass the overcharge on to its customers. According to the Commission's findings, the vitamins cartel covered more than 80 percent of the worldwide market. Consequently, market conditions were the same for all vitamin purchasers, each of whom could only buy products from cartel members. Each could therefore pass any overcharge on to its own customers without incurring any loss of customer base. The fact that Arkopharma had chosen not to pass through the overcharge reflected a deliberate commercial decision, and the loss of margin suffered (as Arkopharma in most cases effectively absorbed the overcharge attributable to the cartel) was not causally linked to the infringement.

The second interesting French judgment on the issue of causation is *Doux*.[134] This was another follow-on action from a Commission infringement decision, this time against the company Ajinomoto in the wake of the lysine cartel.[135] Doux argued that it had suffered damage since the cartel in which Ajinomoto participated had raised the price of lysine, an essential component of animal feeds produced by Doux. The court dismissed Doux's claim, finding that Doux had failed to establish a causal link between the infringement and the alleged damage suffered, since Doux had not demonstrated that market conditions prevented it from passing on the overcharge to its customers. Similar to the reasoning in the *Arkopharma* case, the court noted that all lysine producers had participated in the cartel and therefore all producers of animal foodstuffs were in the same position (i.e., they all suffered the same overcharges). Doux could therefore have passed the overcharge on to its customers without any risk of losing clients to the competition. If Doux had suffered damage in the form of reduced margins, therefore, such loss was not attributable to the infringement but to Doux's own commercial strategy.

These judgments illustrate the close relationship between the issues of causation and the passing-on defense, which is considered further below. When the passing-on defense is allowed, direct purchaser claimants may have more difficul-

132. Arkopharma v. Roche (Nanterre Commercial Court, May 11, 2006).

133. *Vitamins*, *supra* note 77.

134. Doux Aliments v. Ajinomoto (Paris Commercial Court, May 29, 2007).

135. Commission decision of 7 June 2000, Case COMP/36.545 - Amino Acids, 2001 O.J. (L 152) 24.

ty proving that the infringement caused any damage to them. At the same time, courts that disallow or are more restrictive with respect to the passing-on defense may be more hesitant to reject claims on the ground that the causation element has not been established.

2. Quantification of Damages

As explained in Section I.C above, in *Crehan* the European Court of Justice established the right of individuals to obtain damages for breaches of EU competition law, but the types of damages available to claimants and the methods by which those damages are calculated remain largely matters of national law. According to the ECJ's more recent judgment in *Manfredi*,[136] each member state may choose how best to provide for compensation of damages, provided that (i) domestic rules do not discriminate against damage claims for breach of EU competition rules, as compared with claims under national rules (the principle of equivalence) and (ii) domestic rules do not render the exercise of the right to damages excessively difficult (the principle of effectiveness). The principle of effectiveness requires member states to allow claimants the potential to claim compensation for actual loss, lost profit, and interest caused by the infringement of EU competition law, but issues such as punitive damages and restitution are left to the member states (subject to the principle of equivalence). This has created a legal patchwork across the European Union in which defendants face the prospect of significantly different degrees of liability in various countries depending on permitted types of damages and methods of quantification. The incentives for forum shopping are clear.

a) Which damages are available?

(i) Compensatory damages

The primary basis on which damages are assessed in all member states is to award compensatory damages for loss actually suffered by the claimant as a result of the infringement. For example, in England and Wales, where the approach is broadly similar to that taken in the civil law jurisdictions, claimants must establish, on a balance of probabilities, that the infringement caused the loss and that the loss was not too remote, speculative, or inconsequential to be recovered.

In assessing compensatory damage claims, courts often apply a counterfactual analysis, comparing the claimant's actual position to the situation the claimant would have been in "but for" the illegal conduct. Such an approach most straightforwardly includes overcharges (e.g., higher prices resulting from cartel behavior) among recoverable compensatory damages (although the French *Arkopharma* and *Doux* judgments summarized above highlight that overcharges may not be recoverable if the claimant was able to pass them on to customers).

136. *Manfredi*, *supra* note 28.

Compensatory damages also typically encompass lost net profits, which can be in the form of opportunity cost (measured by reference to earnings, as in a case where a product reseller had to reduce its purchases because of cartelized pricing), and can also cover lost going concern value (normally measured by reference to market valuation).

To be recoverable, compensatory damage claims must not be too speculative or too remote from the conduct at issue. In applying the usual "but for" framework, some speculation in establishing what would have happened absent the illegal conduct is unavoidable, but courts set boundaries on the extent of permissible speculation. For example, a claimant might allege that the defendant's abusive rebate scheme caused damage in the form of both (i) loss due to customers terminating agreements with the claimant as a result of the defendant's unlawful pricing terms and (ii) lost enterprise value due to market share erosion following from an inability to compete against the defendant's scheme. The second claim would likely be regarded as more speculative than the first, and could well be disallowed. The issue of remoteness deals with how directly relevant an alleged harm is to the conduct at issue (i.e., how many causal links are needed to connect the damage to the infringement). For example, a claim in the example above for loss due to the cost of borrowing additional operating capital, which the claimant allegedly would not have needed but for the defendant's unlawful rebate scheme, might be seen by a court as too remote from the infringement.

Whether a given claim will pass or fail the speculation and remoteness requirements is highly fact specific, so it is difficult to identify clear trends as to how cases are likely to be decided. In England, the courts have in most cases held that future profits are too speculative to be recovered. By contrast, in Italy certain cases have awarded claimants damages on account of loss of future profits, calculated based on the average duration of contracts that were terminated as a result of the competition law infringement.

EXEMPLARY DAMAGES ARE INTENDED TO PUNISH THE DEFENDANT AND HAVE A DETERRENT EFFECT. IN CONTRAST TO THE U.S. SYSTEM, MOST EU MEMBER STATES REGARD EXEMPLARY DAMAGE AWARDS AS CONTRARY TO PUBLIC POLICY AND DO NOT PERMIT THEM.

(ii) Exemplary/punitive damages

Exemplary damages are intended to punish the defendant and have a deterrent effect. In contrast to the U.S. system, which grants successful plaintiffs treble damages under the Sherman Act of 1890, most EU member states regard exemplary damage awards as contrary to public policy and do not permit them. The only present exceptions to this rule are in Ireland,[137] where exemplary damages are permitted for conscious and deliberate competition law violations (although such awards have been rare), Cyprus, and England, where exemplary damages can be

137. Irish Competition Act, at § 14(5).

awarded if a defendant's unlawful conduct had been calculated to result in profit that would exceed any compensation that might have to be paid to claimants (which could apply in the antitrust context).[138]

An important judgment on the issue of punitive damages, in the *Devenish* litigation,[139] was issued by the English High Court in late 2007. In a follow-on claim for damages against the companies involved in the vitamins cartel, the court held that the claimants were not entitled to recover exemplary damages. The defendants had already been assessed record fines by the European Commission, which in the Court's view precluded an award of exemplary damages under Community law. First, the court reasoned that because exemplary damages and regulatory fines are both intended to punish and deter anticompetitive behavior, the non bis in idem (double jeopardy) principle precludes an award of exemplary damages in a case in which the defendants have already been fined (or had fines imposed and then reduced or commuted) by the European Commission. Second, the court found that an exemplary damage award would "run counter" to the Commission's decision, in violation of Article 16 of Regulation 1/2003 EC, since it would essentially amount to a conclusion that the Commission's fines were inadequate to punish and deter. Finally, the Court acknowledged that under domestic English law, exemplary damages are within the Court's discretion, but noted that the fact that a defendant had already been fined by the European Commission would be a strong argument against the award of exemplary damages. The claimant was therefore not entitled to recover exemplary damages under either EU or national law. The UK Court of Appeal upheld the trial court's judgment in October, 2008.[140] The persuasive effect of this precedent in other member states remains to be seen, but the ruling—that under EU law punitive damages cannot be awarded in follow-on damages claims—is clear.

Despite a general lack of support for exemplary damage awards at the member state level, the European Commission has at various times considered trying to press for change. The Green Paper raised the possibility of introducing mandatory double damages for cartel behavior across the European Union as a means of creating additional incentive for injured parties to bring damages claims.[141] In March 2007, Commissioner Neelie Kroes stated that although she would not support the introduction of U.S.-style automatic treble damages, "double damages for hard core cartels are worth considering, but only if it is proven that sin-

138. Rookes v. Barnard, 1964 A.C. 1129(1), at 37 (*per* Lord Devlin).

139. Devenish Nutrition Ltd. & Ors. v. Sanofi-Aventis (France) & Ors., 2007 E.W.H.C. 2394 (Ch.).

140. Devenish Nutrition Ltd. v. Sanofi-Aventis (France) & Ors., 2008 E.W.C.A. Civ 1086.

141. Green Paper, *supra* note 4, at § 2.3; and Working Paper, *supra* note 6, at § III.B.3.

gle damages are not enough to get the victims to court."[142] The tension between this idea and the *Devenish Nutrition* court's ruling that exemplary damages in follow-on litigation are prohibited by Community law and would likely not be awarded in England and Wales—the largest jurisdiction where exemplary damages are potentially available—is evident.

Perhaps in implicit recognition of this, the White Paper does not advocate the introduction of multiple or punitive damage awards. The Commission points out that Community law does not prevent member states from providing for punitive damage awards,[143] and does not exclude introducing them in future if private damages actions in Europe do not become more common over the coming years.[144] But the White Paper's recommendations are firmly rooted in the compensatory principle of damages as articulated by the Court of Justice in *Manfredi*, which is the same as or very similar to the basis already used in most member states. To increase transparency and awareness, the White Paper suggests that the rules set forth in *Manfredi* should be codified in a Community legislative instrument.[145]

(iii) Restitution

An alternative remedy available in many member states is restitution, which aims to prevent unjust enrichment of the defendant by permitting the claimant to recover the amount of illegal gain obtained by the defendant. Restitution may be of particular relevance in two circumstances likely to arise in the antitrust context: (i) where the claimant seeks to recover profits made by an infringing party, on the basis of a theory of unjust enrichment; and (ii) where the claimant seeks to recover sums paid that cannot otherwise be recovered because the parties are co-contractors to an agreement that is void for being in breach of Article 81 EC. In the view of some authorities, restitution may also be an appropriate basis for calculating damages in representative actions, where the calculation of compensatory damages on an individual basis may be too complex or inefficient.

Approaches to the issue of restitution differ among member states. For example, in Germany the Federal Cartel Office may order the defendant to pay an amount corresponding to the gain made from the antitrust infringement, even if this exceeds the amount of the claimant's loss.[146] In Italy, restitution is also avail-

142. Kroes (2007), *supra* note 30.

143. Working Paper, *supra* note 6, at paras. 188-92.

144. *Id.* at para. 195.

145. White Paper, *supra* note 5, at § 2.5.

146. GWB, at § 34.

able, although under Italian law restitution may not exceed the claimant's loss. Further, while the profit realized by the defendant as a result of the unlawful conduct is, in principle, irrelevant to the calculation of compensatory damages, it may in certain circumstances be taken into consideration to estimate the claimant's loss of income. For example, in refusal to deal cases where the incumbent keeps competitors out of a new market, the incumbent's actual profit may be a proxy for the profit lost by competitors.[147] In England and Wales, restitution is an equitable remedy within the discretion of the court. While not yet awarded in the antitrust context, there has been some suggestion by the Competition Appeals Tribunal that restitution may be an appropriate basis on which to quantify a monetary award for antitrust harm, and the OFT has recommended that courts be given discretion to award damages on a restitutionary basis in representative actions where calculation of compensatory damages on an individual basis may be evidentially too complex or inefficient.[148] However, in the recent *Devenish* judgment,[149] the English High Court ruled that an antitrust claimant (in this case, a purchaser in follow-on litigation based on the vitamins cartel) was not entitled to an account of the profits of the parties to the cartel or restitution of unjust enrichment, but could seek compensatory damages only. The court rejected the claimant's argument that compensatory damages would be insufficient due to difficulties in proving the exact amount of loss, and declined the claimant's invitation to recognize that restitutionary awards are available in all tort cases, including the breach of statutory duty claim at issue. The UK Court of Appeal upheld the trial court's judgment in October, 2008.[150]

(iv) Interest

The European Court of Justice has held that compensating a claimant for loss suffered must take into account the time value of money, which means that interest on the loss is an essential element of compensation.[151] Specific rules for the calculation of interest are left to the member states. The availability of pre-judgment interest can have a significant impact on the total value of a damages award. It is notable that, while the U.S. system provides for treble damages, pre-judgment interest accrues only from the date of bringing the claim, rather than the date of injury; in addition, such interest is awarded only on a showing that the defendants engaged in dilatory or bad faith conduct during the litigation.

---

147. *See* Wind v. Telecom (Rome Court of Appeals, Jan. 20, 2003).

148. *See* OFT Discussion Paper, *supra* note 44, at paras. 7.31 & 7.35.

149. Devenish Nutrition and Ors v. Sanofi-Aventis, 2007 E.W.H.C. 2394.

150. Devenish Nutrition Ltd. v. Sanofi-Aventis (France) & Ors., 2008 E.W.C.A. Civ 1086.

151. Case C-271/91, M. Helen Marshall v. Southampton and South-West Hampshire Area Health Authority, 1993 E.C.R. I-4367, at para. 31.

In Europe, by contrast, pre-judgment interest dated from the time of injury is available in several member states. For example, in England, damages are typically assessed as at the date of the infringement, with pre-judgment interest generally awarded (simple interest at the claimant's normal borrowing rate, or the Bank of England base rate plus one percent); post-judgment simple interest on the judgment debt, and on any costs award made, is also payable at a rate of 8 percent. In Germany, interest is due from the moment the damage occurred, with the interest rate fixed at the base rate of the central bank plus five percent. In Spain, courts may award interest from the date on which the damage occurred, but this rule is flexible and not always applied.[152] On the other hand, pre-judgment interest is generally not available in France, although judges retain discretion in awarding it. Relatively severe rules on interest such as those in Germany and England may result in damage awards that are as high as would be involved in a case of exemplary damages, although the principle behind awarding interest is compensatory rather than punitive.

The White Paper cites the *Marshall* and *Manfredi* judgments in support of its position that victims' rights to full compensation for the harm caused includes not only actual loss and loss of profit, but also interest from the time the damage occurred until the sum awarded is actually paid.[153]

b) How are damages calculated?

As noted above, compensatory damages in all member states aim to put the claimant in the same position it would have been in "but for" the infringement. There are no limits on the amount of damages that may be awarded and no obligation for the court to take into account any fines that may already have been imposed on the defendant by a competition authority.

Compensatory damages are usually calculated as the difference between the claimant's actual position and the hypothetical position that the claimant would have been in but for the unlawful conduct. This measure therefore includes both actual losses sustained by the claimant and profits missed as a result of the infringement. Quantifying this can be difficult, since reconstructing the counterfactual "but for" scenario typically requires making key economic assumptions, small changes to which can have significant effects on the outcome.

Claimants will most commonly seek compensation for two types of antitrust damages: overcharges (e.g., artificially high prices due to price-fixing or market allocation agreements) and losses due to other anticompetitive behavior such as refusal to supply or exclusionary conduct by a dominant rival (e.g., lost profits from missed sales opportunities or lost enterprise value due to market share erosion).

152. For example, in a prominent recent case the First Instance Court of Madrid awarded legal interest to the claimant from the date legal proceedings were initiated. Antena 3 v. LFP, 2006 A.C. 172 (SJPI Madrid, Jun. 7, 2005) [hereinafter *Antena*].

153. White Paper, *supra* note 5, at § 2.5.

Assessing either type of damage will generally require detailed economic evidence to estimate what market conditions (price, market share) would have been but for the infringement. Several methods may be employed. Most straightforwardly, prices (or market shares, etc.) before and after the infringement can be compared, with the difference attributed to the anticompetitive conduct. Such a simple approach may, however, ignore other factors that can have important effects on the outcome such as macroeconomic trends or changes in the cost of inputs. In an effort to identify and eliminate such external variables, economists can compare the evolution of the relevant market to that in a "yardstick" or "benchmark" market that is (presumably) untainted by the effect of the infringement—the principal difficulty being to identify another market that can serve as a reliable benchmark. Alternatively, more sophisticated econometric techniques (regression analysis) can be employed to model the "but for" price based on multiple cost and demand variables in an effort to isolate the effect of the infringement. Such analyses are the most robust available method of calculating damages, but they are often limited by data availability, as well as being complex and subject to dispute by opposing economists.

Given the difficulty of constructing the "but-for world," both claimants and defendants face tough practical choices in relation to the presentation of evidence regarding quantum of damages. The most robust assessments of loss suffered based on complex econometric modeling can be difficult to present in straightforward terms that courts will find persuasive. On the other hand, simple, intuitive calculations may be more susceptible to rebuttal.

THE MOST ROBUST ASSESSMENTS OF LOSS SUFFERED BASED ON COMPLEX ECONOMETRIC MODELING CAN BE DIFFICULT TO PRESENT IN STRAIGHTFORWARD TERMS THAT COURTS WILL FIND PERSUASIVE. ON THE OTHER HAND, SIMPLE, INTUITIVE CALCULATIONS MAY BE MORE SUSCEPTIBLE TO REBUTTAL.

In view of such considerations it is not surprising that claimants often struggle precisely to quantify the amount of their loss. In most member states, however, this will not preclude the recovery of damages. For example, in England the court will simply do the best it can to quantify loss based on the evidence put before it. Similarly, in Italy, to the extent that the claimant cannot prove exactly how much damage was suffered, the judge is entitled to quantify the damage on an equitable basis. In Germany, the competition rules expressly permit the court to take into account cartel profits as a means of estimating the amount of overcharge suffered by the claimant. This applies where there is a *prima facie* case on the merits, which leads to a shift in the evidentiary burden of proving damage, making the issue easier for the claimant. By contrast, in Spain claimants normally face a higher burden of proving the exact amount of their loss.[154]

154. *See, e.g.*, *Antena*, *supra* note 152 (judicial standard to determine the loss of profits takes into account gains that were not merely a possibility, but were "very likely to have been obtained"

footnote 154 cont'd on next page

In recognition of these difficulties, the White Paper announces that, to facilitate the calculation of damages, the Commission intends to issue non-binding guidelines for quantification of damages in antitrust cases, including simplified rules on estimating the loss suffered as a result of the infringement.[155]

3. Joint and Several Liability

The tort law regimes of most member states provide that undertakings that are parties to anticompetitive agreements are liable for the entire damage caused by these agreements. That is, the co-infringers are generally jointly and severally liable for the damage caused by their actions. A victim that suffered harm caused by an unlawful agreement may claim his entire damage not only against his direct trading partner(s), but also against any of the other parties to the unlawful agreement. Between the infringers, liability is several (i.e., the infringer who compensated the entire harm of a victim has the right to seek contribution from co-infringers). The ultimate liability shares of each infringer are determined at this contribution stage.

The specter of joint and several liability threatens to undermine cartel leniency programs by creating a strong disincentive for potential leniency applicants to come forward. In recognition of this disincentive, the White Paper puts forward for further consideration a rule whereby the civil liability of successful immunity applicants under leniency programs would be limited to claims by their direct and indirect contractual partners.[156] Such a rule would not grant an additional financial reward to the immunity applicant, since it would still be liable for the damage it had caused. The rule would, however, offer the successful immunity applicant two benefits. First, there would be a procedural advantage in follow-on damages litigation, since the firm would not need to seek contribution from other cartel participants. Second, the rule would insure against the insolvency or unavailability of one or more cartel members since the immunity applicant would not be required to bear the financial burden of compensating any victims other than its own direct and indirect trading partners (as opposed to the remaining solvent cartel members, who would be jointly and

footnote 154 cont'd

should the illegal conduct not have occurred). Note that this judgment was appealed and reportedly reversed, but the Appellate Court's judgment is not yet available. In another case, *Conduit v. Telefónica*, a standalone action based on a violation of Article 82 EC, the claimant was awarded EUR 639,000 in damages but the court declined to award damages for lost profits, on the grounds that reduction of Conduit's market share (which had been presented as a proxy for its lost profits) was not related to the abuse (refusal to supply), but to the level of Conduit's investment in advertising. Conduit v. Telefónica, 2006 A.C. 1881 (SAP Madrid, May 25, 2006).

155. White Paper, *supra* note 5, at § 2.5. In July 2008, the Commission opened a tender procedure for a study on the quantification of damages caused by competition law infringements, the results of which will presumably feed into the upcoming guidelines.

156. *Id.* at § 2.9.

severally liable for the entire damage award).[157] Such a rule would help allay the legitimate fear of potential leniency applicants that, by coming forward with evidence of an infringement, they would be open to liability for the whole loss caused by the cartel—a fear that would potentially undermine the incentive to apply for leniency.

4. The "Passing-on Defense" and the Standing of Indirect Purchasers

a) A pair of difficult issues

The inter-related questions of whether to recognize the "passing-on defense" and whether indirect purchasers should be entitled to sue for antitrust damages raise difficult issues of substance, procedure, and policy and can have a determinative effect on the availability or quantum of damages in many cases. The passing-on defense arises out of the compensatory principle of damages. In jurisdictions that allow it to be raised, defendants can argue that their direct customers should not be entitled to claim the full damage amounts to which they would otherwise be entitled (usually measured as the amount of overcharge attributable to a cartel or abusive pricing scheme) if they passed the higher price through to their own customers downstream.[158] The standing of indirect purchasers (purchasers who had no direct dealings with the infringer, but who nonetheless may have suffered harm because an illegal overcharge was passed on to them along the distribution chain) is linked directly to this issue, since if the overcharge has been passed on by the direct purchaser, indirect purchasers become the primary injured parties.

There are sound policy arguments favoring different approaches to these issues. Regarding the passing-on defense, the compensatory principle of damages counsels in favor of allowing defendants to raise it, since a claimant that passed overcharges through to its customers would be unjustly enriched if its damage award was not reduced correspondingly. On the other hand, the passing-on defense inevitably increases the complexity of litigation because it creates the need to analyze the distribution of an overcharge along the entire relevant product supply chain in order to determine damages. As the preceding section illustrates, estimating overcharges even at one level of distribution is difficult enough; forcing courts to scrutinize price effects along an entire vertical distribution chain may be too much to ask. Moreover, allowing the passing-on defense makes it more difficult for direct purchasers—precisely those who are most likely to have the greatest incentive and ability to bring private actions—to obtain antitrust damages, likely decreasing the overall level of private enforcement.

157. *See* Working Paper, *supra* note 6, at paras. 280-84.

158. Note that even if a direct purchaser passed the full amount of overcharge through to its own customers, it may still have suffered harm due to lost sales if demand for the relevant product is elastic.

Regarding indirect purchasers, the compensatory principle supports allowing standing, since they are, by assumption, the injured parties if the overcharge was passed on by the direct purchaser. It would seem paradoxical if the real victims of anticompetitive conduct could not seek compensation. On the other hand, the same problems of increased litigation complexity, with particularly difficult questions regarding causation and quantum of damages attributable to the infringement, are inevitable in indirect purchaser actions. Indirect purchasers also lack privity with the defendant, which remains a central principle of tort law in several jurisdictions.

Finally, the interplay between the two issues also creates difficulty: allowing the passing-on defense while denying standing to indirect purchasers could mean that unambiguously guilty defendants face no liability for damages they have caused. On the other hand, disallowing the passing-on defense while allowing indirect purchasers to sue may result in unjust enrichment of some plaintiffs and force defendants to pay for the same damage more than once. There are no simple answers to these issues.

b) One Approach: The U.S. Experience

In the United States, the Supreme Court settled on an approach in two landmark judgments, *Hanover Shoe*[159] and *Illinois Brick*.[160] First, in *Hanover Shoe*, the Court rejected the defendant's passing-on defense as a matter of law, reasoning that passing-on arguments would (i) generate unduly long, complex litigation and (ii) discourage lawsuits by making it more difficult for those best placed to bring them, thereby reducing the effectiveness of private actions as an antitrust enforcement tool. Second, nine years later in *Illinois Brick* the Court held that only direct purchasers from cartel members had standing to bring federal antitrust lawsuits for damage recovery. This time, the Court again cited the problem of litigation complexity in indirect purchaser actions (which raise the same practical difficulties as if the passing-on defense were allowed), but was also concerned about the risk of double liability for defendants—which would become six-fold liability given the mandatory trebling of damages under Section 4 of the Clayton Act—if indirect purchasers were allowed standing. Thus, under U.S. federal law, the Supreme Court established a compromise position: streamlining litigation and avoiding double recoveries against defendants, but potentially allowing unjust enrichment of plaintiffs and, more significantly, undermining the compensatory principle by denying potential "real victims" of the illegal conduct the right to bring claims.

Illinois Brick was unpopular and perceived as unfair in many state capitals, which reacted by introducing so-called "*Illinois Brick* Repealer" statutes. These

159. Hanover Shoe v. United Shoe Machinery Corp., 392 U.S. 481 (1968).

160. Illinois Brick Co. v. Illinois, 431 U.S. 720 (1977).

laws, which have been enacted in almost half of the U.S. states, entitle indirect purchasers to sue for treble damages under state antitrust laws—effectively circumventing the Supreme Court's ruling by providing a remedy under state law that the Court denied under federal law.[161] The end result—ironic in light of the rationale of the *Illinois Brick* judgment—is that antitrust defendants in the United States face the prospect of not only multiple recoveries for the same harm, but also extraordinarily complex, duplicative, and even inconsistent litigation in both federal and state courts based on the same underlying facts.

c) The approach under EU law

The Commission's Green Paper[162] signaled that the Commission was prepared to consider all options, inviting comment on four different possible approaches: (1) allowing the passing-on defense, with both direct and indirect purchasers entitled to sue; (2) excluding the passing-on defense, with only direct purchasers able to sue (the approach in the U.S. federal courts); (3) excluding the passing-on defense, with both direct and indirect purchasers able to sue (the de facto approach in the U.S. federal plus state court system); and (4) the introduction of a two-step procedure under which the passing-on defense is excluded in an initial procedure in which the defendant is sued for the total overcharge, then in later proceedings the damages are allocated among all parties (including direct and indirect purchasers) that suffered a loss. The Commission recognized at the time that given the complexity of these issues, a trade-off between justice (in the sense of full recovery for all those who have suffered a loss from an illegal practice) and efficiency is inevitable.

Consistent with its other positions, the White Paper proposes a course grounded in the compensatory principle of damages, based on option (1) above. First, the White Paper advocates granting standing to indirect purchasers. Taking a different position would have been difficult in view of the Court's holdings in *Crehan* (in which the Court states that EU law leaves it "open to any individual to claim damages for loss caused to him") and *Manfredi*, and would also have conflicted with the Commission's consistently stated aim of using competition law to defend consumer interests. Second, the Commission proposes to allow defendants to raise the passing-on defense. The White Paper cites the potential adverse results of unjust enrichment and multiple compensation in support of this position.

At the same time, however, the Commission proposes to ease the burden of proof for indirect purchasers by granting them a rebuttable presumption that the illegal overcharge was passed on in its entirety. The burden would then shift to

161. The Supreme Court later rejected arguments to overturn the *Illinois Brick* repealer statutes as unconstitutional on federalism grounds. *See* California v. ARC America Corp., 490 U.S. 93 (1989).

162. Green Paper, *supra* note 4, at 8.

the defendant to show that the overcharge was not, or was only partially, passed on to the claimant. This approach reduces the possibilities that direct purchasers who have passed on overcharges may be unjustly enriched and that defendants may be required to pay twice for the same harm, while also recognizing and seeking to address in part the difficulties of proof that indirect purchasers commonly face. Defendants in actions brought by direct purchasers will undoubtedly seek to have this pass-on presumption applied in those cases as well, and it remains to be seen how receptive courts will be to such arguments.

d) Member state approaches

At the member state level, approaches to the passing-on and indirect purchaser issues are based largely on the compensatory principle of damages, which counsels generally in favor of allowing the passing-on defense (since passing on of an overcharge would reduce a direct purchaser's actual loss) and allowing indirect purchasers to sue (since indirect purchasers suffer loss if they absorb passed-on overcharges).

For example, in France, courts have allowed defendants to raise passing-on defenses in a series of cases. The *Arkopharma*[163] and *Doux*[164] judgments discussed in Section II.D.1 above are notable examples. In each case, based on a passing-on defense, the court dismissed the claim in a follow-on action where the infringement had already been established by the European Commission. On similar facts, the courts each found that even if the claimants (direct purchasers of products from a cartel member) had not, in fact, passed the full amount of overcharge on to their own customers, they *could* have done so without suffering loss. Because all purchasers had been subject to the cartelized price and demand for the underlying products was relatively inelastic, the direct purchasers could have increased their own prices by the amount of the overcharge without losing sales.[165] The claimants had thus failed to establish that the infringement had caused them any loss.

In Germany, the passing-on defense is disfavored by statute. Damages arising from cartelized prices cannot be excluded simply on grounds that the plaintiff/direct purchaser passed on the overcharge to its customers.[166] When calculating damages, however, the court may take into account the mitigating effects of passing on higher prices by applying the "adjustment-of-benefits principle" (Vorteilsausgleichung) to prevent unjust enrichment. This principle shifts the

163. Arkopharma v. Roche (Nanterre Commercial Court, May 11, 2006).

164. Doux Aliments v. Ajinomoto (Paris Commercial Court, May 29, 2007).

165. *See also* Juva v. Roche (Paris Commercial Court, Jan. 26, 2007).

166. GWB, at § 33(3), sentence 2.

burden of proof onto the defendant to show, against a statutory presumption that the loss suffered coincides with the margin by which prices have been raised artificially, that damages would unreasonably enrich the claimant.[167] The Federal Cartel Office has sought to provide guidance, opining that the passing-on defense should be allowed only in exceptional cases where: (i) the damage has been passed on; (ii) the passing-on did not involve any economic risk for the damaged party; (iii) the passing-on required only minimal effort; and (iv) the passing-on did not result in a decline in sales.[168] It is not entirely clear whether indirect purchasers have standing to sue, but if the passing-on defense is allowed in some cases, the logical consequence is that the indirect purchaser at the next market level should (exceptionally) be entitled to claim damages.[169]

In Italy, the passing-on defense has not been recognized expressly. However, under general civil liability principles, a claimant may only seek compensation for harm that it has actually suffered, and provided that it did not knowingly contribute to the harm. In the only antitrust precedent on this point, the Turin Court of Appeal found that a travel agency could not be awarded damages because it had willfully participated in an anticompetitive agreement with the intent to pass on the overcharge to final customers.[170]

In England and Wales, any party suffering damage—including direct purchasers, indirect purchasers, and competitors—may commence an action. There have been no cases considering the possible admission of a "passing-on defense," but the general principles on mitigation of loss suggest that, if a claimant has suffered no loss (e.g., because it has passed on an overcharge), it will not be entitled to recover any damages. In its 2007 recommendations, the OFT declined to take

167. In a judgment awarding EUR 1.6 million in damages to a direct purchaser (Storck) in a follow-on action against Hoffmann-La Roche AG, a member of the vitamins cartel, the Regional Court of Dortmund ruled that a defendant seeking to reduce the amount of damages to be awarded on the basis of the passing-on defense must prove that the cartelized prices were, in fact, passed on to the plaintiff's customers. Case 13 O 55/02 Kart - Vitaminkartell, EWS 2004 (Regional Court of Dortmund, Apr. 1, 2004), at 434-36. Two prior judgments by other courts in similar cases had taken a broader view of the passing-on defense, rejecting follow-on claims by vitamin purchasers and imposing the burden of proof on the plaintiff to show that higher prices had not been passed on. Case 6 U 183/03 - Vitaminpreise (Higher Regional Court of Karlsruhe, Jan. 28, 2004); and Case 7 O 326/02 - Vitaminpreise (Regional Court of Mannheim, Jul. 11, 2003).

168. These conditions might apply, for example, in connection with "cost-plus" contracts, in which the cost of inputs is passed directly through to the downstream purchaser and the intermediate producer's margins are fixed in advance.

169. *See* U. Böge & K. Ost, *Up and Running, or is it? Private enforcement - the Situation in Germany and Policy Perspectives*, 27 E.C.L.R. 197, 200 (2006); Bundeskartellamt, Private Kartellrechtsdurchsetzung: Stand, Probleme, Perspektiven (discussion paper) (Sep. 26, 2005), at 12, *available at* http://www.bundeskartellamt.de/wDeutsch/download/pdf/Diskussionsbeitraege/05_Proftag.pdf.

170. Indaba v. Juventus (Turin Court of Appeals, Jul. 6, 2000).

a position on the issue of indirect purchaser standing and the passing-on defense, opting to wait for the White Paper.[171]

5. Availability of Interim Remedies

An interim remedy such as an interim injunction allows a claimant to force the defendant to amend or terminate allegedly anticompetitive behavior pending final resolution of the case at trial. Claimants will often seek interim injunctions in cases where the alleged harm is ongoing, such as those involving allegations of infringements such as abusive exclusionary conduct, predatory or below-cost pricing, or refusals to deal.

An application for interim relief may be brought prior to proceeding on the merits, although it is usually necessary to commence substantive proceedings within a specified period of time thereafter (e.g., within 60 days in Italy and Germany). As noted in Section II.B.1 above, subject to national jurisdictional rules, it may be possible to obtain an interim injunction even in a member state that does not have substantive jurisdiction on the merits of the case.

The test for granting interim relief appears at first blush to be very similar across most member states, but as explained below, this obscures important practical differences that may be determinative of injunctive relief availability. For example, national laws differ on issues such as the required strength of the claimant's case, the standard according to which the claimant must show that the alleged harm cannot be compensated by damages, and issues of timing (urgency, delay, maintenance of the status quo). Interim relief availability is therefore an important consideration in the jurisdictional strategy of both claimants and defendants. The approaches in several member states are summarized below.

a) England

In England and Wales, injunctive relief is an equitable remedy and is discretionary. It is necessary for the claimant to establish: (i) a "serious issue to be tried," such that the applicant has a real prospect of success (arguably a slightly lower standard than the more common prima facie case requirement); (ii) risk of irreparable harm, such that damages would not be an adequate remedy; and (iii) that the balance of convenience lies in favor of granting the injunction. This third factor requires the court to consider the harm that would accrue to each party from granting or not granting the injunction, having regard to factors such as the extent to which each party might suffer loss that is not compensable in damages, preservation of the status quo, delay in seeking an injunction, and whether the effect of the injunction would be oppressive or disproportionate.

171. OFT Discussion Paper, *supra* note 44, at paras. 12.8-12.15.

The 'balance of convenience' test provides considerable scope for discretion in determining whether to grant an injunction. For example, in the *Adidas* case,[172] the Adidas sportswear company brought an action against the organizers of the four major international "Grand Slam" tennis championships, alleging that the organizers had applied their dress code rule in a discriminatory manner, in breach of Articles 81 and 82 EC. Adidas sought an injunction to prevent the application of the dress code to its three-stripes design, which would have prevented Adidas-sponsored players from wearing new Adidas clothing during the tournaments. Notwithstanding that the Grand Slams had given Adidas a year's notice that its design would need to comply with the dress rule, the court held that the balance of convenience favored granting the injunction. The court noted in particular that damages would not be an adequate remedy (as it would be difficult to quantify Adidas's lost sales if the dress rule had been enforced), that granting the injunction would not harm the tennis federations, and that maintenance of the status quo favored the injunction.

b) France

Interim or "preventive" relief may be obtained upon a showing that there is risk of imminent harm from obviously unlawful conduct. No separate proof of urgency is normally required. Alternatively, interim relief may be granted if the plaintiff shows urgency and the defendant does not seriously contest the underlying facts. In this case, it is not necessary for the plaintiff to establish an obviously unlawful act on the part of the defendant. In either case, injunctive relief may be granted even if damages would be an adequate remedy at trial, although the adequacy of damages will be taken into account in determining whether to grant the relief.

c) Germany

To obtain interim relief in Germany, the claimant must establish: (i) a prima facie case; (ii) urgency; and (iii) risk of harm. Further, if a preliminary ruling on the merits is the only means of protecting the applicant's interests, the applicant must also show one of the following: substantial economic detriment for which any subsequent award of damages would be an inadequate remedy; a threat to the claimant's economic existence; substantial competitive disadvantage; or that the balance of interest lies in granting the relief. Where the claimant is a small or medium-sized enterprise, there are rebuttable presumptions that reduce the evidentiary burden in actions based on an alleged abuse of dominance.[173] Accordingly, from a potential claimant's point of view, Germany may be an attractive forum in which to seek interim relief.

172. Adidas-Salomon AG v. Draper & Ors, 2006 E.W.H.C. 1318 (Ch).

173. GWB, at § 20(5).

d) Italy

While the process of obtaining interim relief may be somewhat slower on average than in other member states, applications for interim injunctions are common in Italian antitrust cases. The applicant must demonstrate: (i) a prima facie case, (ii) urgency, and (iii) a risk of imminent serious and irreparable harm that is not readily compensable in damages.

e) Spain

Interim measures are available in Spain. For an injunction to be granted, the applicant must show that: (i) there is "appearance of good right" (fumus boni iuris, i.e., that the application is based on solid arguments); and (ii) there is a risk that the final decision will not, without an interim remedy, be enforceable (periculum in mora).

Spain was recently the subject of a major interim measures proceeding originating in the merger context. In September 2005, the Spanish gas company Gas Natural launched a hostile bid for the leading Spanish electricity company Endesa and, to eliminate anticipated competition concerns, agreed in advance to sell some assets to the second Spanish electricity company, Iberdrola. Amongst other defensive responses, Endesa applied to the commercial court in Madrid for an interim injunction blocking the takeover bid on grounds that it was the instrument for the execution of an unlawful agreement between Gas Natural and Iberdrola, in violation of Article 81 EC. Judge Miriam Iglesias granted the order and suspended the takeover bid on the basis that Endesa posted a €1 billion guaranty against damages to Gas Natural. Gas Natural and Iberdrola appealed the interim order, and the Audiencia Provincial of Madrid upheld the appeals, lifting the suspension on January 16, 2007—by which time the Supreme Court had also made an interim order suspending the bid following Endesa's appeal of the merger clearance decision. In the end, Endesa and Gas Natural agreed to a settlement under which they mutually decided to withdraw proceedings.

III. Conclusion: What Is to Come?

Private competition law litigation in Europe is evolving from the state of "total underdevelopment" described in the 2004 Comparative Report. In the last few years there have been several highly publicized cases in multiple jurisdictions, and the White Paper has focused the European legal and business communities' attention on the issue. Facilitating private enforcement as a complement to public enforcement has become a central plank of European Commission competition policy. Yet most of the essential issues still impeding actions for damages are under the responsibility of the member states. The Commission seems determined to exercise its legal and persuasive authority to promote development of national rules to facilitate private claims, but the extent to which these efforts will bring about meaningful change remains unclear.

The White Paper's recommendations will likely set the tone for the development of the legal systems governing private damages litigation across Europe over the coming years. It had been speculated that, following the 2005 Green Paper, which set forth a wide range of options for discussion, the White Paper might propose ambitious measures that would foster "U.S.-style" antitrust litigation, such as multiple damages, opt-out class actions, or extensive discovery rules. By and large, however, the Commission has not proposed measures contemplated in the Green Paper that would have been viewed as dramatic or controversial: the White Paper's recommendations fall largely within the scope of existing European civil law practice and principle. This appears implicitly to reflect the Commission's appreciation of the great inertia in the member states' legal systems of, particularly, the civil law member states, where resistance to changing legal traditions to promote competition law litigation will be stiff. The White Paper can be viewed as an effort by the Commission to set achievable goals based on relatively conservative measures that have a realistic chance of being widely adopted.

The White Paper repeatedly highlights the Commission's intention to preserve a "genuinely European approach" to the issue of damages actions that is "rooted in European legal culture and traditions."[174] The Commission does not indicate what the next step in its effort to promote private actions will be, but one possibility would be seeking to pass a Regulation, which would require support of the European Parliament and Council. The Parliament has already indicated its support in principle, having issued a Resolution in 2007 calling for the adoption of common measures at the EU level "to facilitate the bringing of 'stand alone' and 'follow on' private actions claiming damages for behaviour in breach of the Community competition rules."[175] In the meantime, the White Paper may already provide guidance to national judges who are asked to decide on an action for damages under Article 81 or 82 EC.

Perhaps more important, the EU jurisdictions in which antitrust damages litigation is most developed—particularly England—are already considering and implementing measures to promote private actions that in many respects go beyond the White Paper's recommendations. Other member states are also moving in this direction, as witnessed by the recent enactments of new laws expressly intended to facilitate private actions in countries such as Germany, Sweden, and France.

Market developments also point toward more damage actions in the near future. As described above, the Belgium-based CDC firm (as well as other similar enterprises such as the Germany-based Talionis) has had some initial success

174. White Paper, *supra* note 5, at § 1.2.

175. European Parliament resolution of 25 April 2007 on the Green Paper on Damages actions for breach of the EC antitrust rules - 2006/2207(INI).

using a model to bring collective damages actions based on claims acquired from the injured parties. In addition, the major U.S. plaintiff's firm Cohen, Milstein, Hausfeld & Toll set up its first European office in London in 2007 and is reportedly pursuing follow-on actions for damages relating to several Commission cartel decisions. Such firms will have been disappointed by the White Paper's conservative recommendations. The establishment in Europe of an active plaintiff's bar, however, under these or other models will no doubt result in more damages actions being brought.

The form that such actions will take is not clear. The White Paper strongly favors a model for collective actions brought by representative organizations under an opt-in structure, but recent comments by the U.K. consumer organization Which?, following its negotiated settlement in the *JJB Sports* case, call into question the financial viability of this model. The emerging plaintiff's bar will likely focus on follow-on litigation on behalf of direct purchasers from cartel participants, but these actions will also face hurdles such as widespread recognition of the passing-on defense (as recommended by the White Paper), general unavailability of punitive damages, and continued resistance to contingency fees for lawyers (which is at odds with the "acquisition of claims" model being pursued by CDC). Significant obstacles for plaintiffs remain, and no predominant model for bringing damage actions has yet emerged.

Despite such obstacles, the high level of cartel enforcement activity by the European Commission and national authorities—by some reports over 150 immunity applications, each of which potentially indicates a separate upcoming cartel decision, are pending at DG COMP alone—will ensure a rich supply of new potential targets for damage claimants over the coming years. Follow-on actions of some form, brought in countries with the most plaintiff-friendly rules of civil procedure, thus seem likely to comprise the bulk of private damage claims for the foreseeable future. However, while such actions may be appropriate in gaining compensation for victims, it seems unlikely that they contribute significantly to overall competition law enforcement, since by definition the infringement has already been discovered and the offenders punished. As the Commission has recognized, private enforcement will only be a true complement to public enforcement if it "extend[s] the scope of enforcement beyond the cases already dealt with by public authorities,"[176] which means that focusing on measures to facilitate standalone damages claims, rather than follow-on actions, is the appropriate policy priority.

176. Kroes (2007), *supra* note 30.

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Will Europe Provide Effective Redress for Cartel Victims?

Vincent Smith

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853),
Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Will Europe Provide Effective Redress for Cartel Victims?

*Vincent Smith**

This article gives an overview of the history of the development of private redress for competition law breaches in Europe. The article begins by reviewing the current proposals to improve private actions, examines the areas where further development is still required, and makes some suggestions as to how to tackle the most important of these. The issues discussed include how to determine which court should hear competition claims, how to institute a process that does not result in a multiplicity of actions across the European Union, and what system would ensure that claimants achieve effective redress while also being fair to defendants.

*The author is partner with Cohen Milstein Hausfeld & Toll LLP in London and was Senior Director for Competition at the UK Office of Fair Trading from 2003-2007.

I. Historical Perspective

Recent attention to the lack of redress available to the victims of cartel and other competition law breaches in Europe has highlighted the main difference between the competition regimes in Europe and that of most other countries in the Organization for Economic Cooperation and Development ("OECD"): the relative lack of redress for victims (especially the smaller ones). Although there may be more hidden resolutions and settlements of competition cases than are readily apparent publicly,[1] the overall recovery for competition law victims has been much lower than in, for example, the United States, where civil damages actions in this area have been encouraged for nearly 40 years. This article seeks to assess the extent to which the current initiatives (particularly those undertaken by the European Commission) will help tip the scales towards a better equilibrium, particularly for consumers and small businesses who have been the victims of cartels. This article then suggests some issues which would benefit from further consideration.

At the outset, it is worth recalling that competition law (and competition enforcement in particular) is a relatively new phenomenon in Europe. In contrast to the United States, where the need for private parties to have a means of tackling abuses of market power (whether cartel conspiracies or misuse of monopoly power) was recognized early,[2] European countries preferred an economic policy based on substantial state intervention. This normally took the form, especially in the decades following the Second World War, of either central state planning of the economy with all enterprises under direct state control (as in central and eastern Europe), or a mixture of regulation and state ownership (in western Europe). This pattern only began to change in the 1980s with the beginning of a withdrawal from state ownership of industrial firms in Western Europe, in some cases accompanied by a new set of ex ante regulatory requirements to ensure full service to all customers. These latter were prominent in "network" industries[3] and became increasingly important throughout Western Europe as the decade progressed.[4]

AT THE OUTSET, IT IS WORTH RECALLING THAT COMPETITION LAW (AND COMPETITION ENFORCEMENT IN PARTICULAR) IS A RELATIVELY NEW PHENOMENON IN EUROPE.

With the change in regime in Central and Eastern Europe in the late 1980s and early 1990s, countries in those regions opened their markets to competition

1. Barry Rodger, *Private Enforcement of Competition Law: The Hidden Story*, 29(2) E.C.L.R. 96 (2008).

2. U.S. Sherman Act, 15 U.S.C. § 1-7 (1890).

3. United Kingdom, The Telecommunications Act of 1984 (Apr. 12, 1984).

4. Germany, Telecommunications Act (Telekommunikationsgesetz, TKG), Federal Law Gazette I (Jul. 25, 1996), at 1120, as amended by subsequent Acts.

relatively quickly and introduced competition laws based on the EC model; indeed, it was a condition of accession to the European Union that they did so.[5] Thus, the Member States which joined the European Union in 2004 all had functioning competition regimes from the late 1990s onwards.

Before the 1980s, there were two exceptions to the then prevalent European picture of state ownership and direct intervention in preference to the use of competition laws. Germany introduced its first law against restrictions of competition in 1949, but substantially updated it in 1958. At more or less the same time, the Treaty of Rome and its fundamental principles established both a prohibition on anticompetitive agreements (cartels and others) and a prohibition on abuse of market dominance.

The German emphasis on competition was a reaction against the corporatism, and encouraged by the state, under the National Socialist regime before the Second World War. It was part of a wider group of laws and constitutional provisions designed to address the excesses of the capitalist system which were widely believed to have contributed to the events leading to the war. The most prominent example of these German laws is, of course, the law on worker codetermination (Mitbestimmungsgesetz). This emphasis carried over into the negotiations leading to the Treaty of Rome and gave a "constitutional" importance to the EC competition rules. Although the emphasis was on protecting "free competition", the reasons for doing so were more a result of political reaction to historical events than the overtly ideological, freedom-based, U.S. and other Anglo-Saxon systems of competition law.

These developments, of course, had a significant impact on the enforcement regime envisaged in each legal system. Whereas the U.S. Sherman Act proceeded on the basis that everyone (including the public authorities) could bring a claim, the Treaty of Rome proceeded on the basis that the relevant provisions of the Treaty would be enforced by the European Commission, a public enforcer with a wide-ranging remit to ensure that the law contained in the Treaty is observed.[6]

In one of the first cases (1966) brought by private parties under the European competition rules, the European Court of Justice ("ECJ") confirmed that an anticompetitive distribution agreement (a "vertical" agreement between a supplier and its distributor) fell within the scope of the prohibition on restrictive agreements.[7]

5. *See, e.g.*, Hungary, The Competition Act on the Prohibition of Unfair Market Practices (Act LXXXVI of 1990); Poland, The Act on counteracting monopolistic practices and protection of consumers interests of 24th February 1990; and Czech Republic, The Act on the Protection of Economic Competition, No. 63/199, as amended Act No. 495/1992 Coll.

6. European Community, Treaty Establishing the European Community, as amended by subsequent Treaties, Rome (Mar. 25, 1957), at art. 89(1).

7. Joined Cases 56 & 58/64, Consten and Grundig v. European Commission, 1966 E.C.R. 299.

A German manufacturer, Grundig, and its French distributor, Consten, had appealed against a Commission decision finding that their distribution agreement, which prevented all competition to Consten in France for the sale of Grundig branded products, had breached Article 85 (now Article 81) EC Treaty.

This is not the most obvious case of anticompetitive behaviour. It does, however, demonstrate a further strong theme which runs through competition law enforcement at the European level—the emphasis on creating a single economic market across the European Union. Free trade is a main vehicle for creating the single market, unrestricted not only by state barriers (import duties and the like), but by also private barriers through restrictive agreements and abuses of market power to keep prices higher in some Member States than in others.

IT DOES, HOWEVER, DEMONSTRATE A FURTHER STRONG THEME WHICH RUNS THROUGH COMPETITION LAW ENFORCEMENT AT THE EUROPEAN LEVEL—THE EMPHASIS ON CREATING A SINGLE ECONOMIC MARKET ACROSS THE EUROPEAN UNION.

Despite their ubiquity and obvious harm to economic efficiency, it was not until the early 1970s that the European Commission levied its first significant fines against a cartel.[8] Cartel enforcement increased significantly in the following decade and, beginning in the mid 1990s, became the centerpiece of the Commission's competition law enforcement programme.[9] One of the reasons for the change of emphasis may well have been the completion of the "single market" programme and the Commission's realization that it needed to concentrate its resources on the larger and more serious competition breaches, in particular EU-wide cartels.

EC (and German) competition law nevertheless remained the exception in Europe until the 1990s. Even the other European countries that had competition law regimes did not (generally) use them effectively against anticompetitive behaviour. Far less was there any acceptance that private parties should have a right to redress for loss caused by such behaviour.

For example, although the United Kingdom introduced its first competition statute in 1949,[10] the law did not provide for a cartel enforcement regime. Instead, it gave a new state body (then called the Monopolies Commission) the power to look into sectors of the economy, at the request of the government, and make recommendations for changes to the industry structure (usually through

8. Joined Cases 40-48, 50, 54-56, 111, & 113-114/73, Suiker Unie v. European Commission (Sugar Cartel), 1975 E.C.R. 1663.

9. EUROPEAN COMMISSION, DG COMPETITION REPORT ON THE AMOUNT OF FINES IMPOSED BY THE COMMISSION, 1990-2008 (2008), *available at* http://ec.europa.eu/comm/competition/cartels/statistics/statistics.pdf.

10. United Kingdom, Monopolies and Restrictive Practices (Inquiry and Control) Act (1948).

government secondary legislation). This mechanism typified the approach to economic regulation during the immediate post-war period and was used relatively little, not least because the state had more direct means of control over industry through the government ownership of key industrial and commercial concerns. (The coal, steel, and railway industries were all nationalized at the beginning of the post-war period, as was the Bank of England.)

Direct enforcement of competition law only became a reality with the establishment of the 1976 Restrictive Trade Practices Act. Of course, by that time, European competition law also applied in the United Kingdom following its accession to the EEC in 1973. However, the Act was formalistic—and, therefore, difficult to apply—and the sanctions provided in the law (such as imprisonment for contempt of court if restraining orders were breached) were felt to be inappropriate in many cases. Public enforcement through formal sanctions was therefore extremely rare, and private litigation was almost unheard of. The 1976 Act, despite its shortcomings, was not replaced until the 1998 Competition Act introduced a European-style prohibition system to the United Kingdom statute book, with effect from March 2000.

Another important element of an effective private redress regime is providing an efficient set of civil law procedures to enable private actions to be brought. Until very recently (and as is still the case in the large majority of EU Member States), the general civil procedure rules were the only vehicles for bringing a claim for redress in the competition field and they have a number of drawbacks when applied to competition law claims.

First, most civil procedure rules have difficulty in dealing with expert evidence—although the nature of these difficulties varies from system to system. This is particularly important for competition claims which are essentially economics-based: legal systems traditionally have difficulties dealing with areas in which the essence of a decision is non-legal. The problem differs depending on whether the litigation system is adversarial (as in common law countries) or inquisitorial (as in civil law countries). In an adversarial system, in which the experts appear as witnesses for each of the parties, the (non-expert) judge is left to adopt one or other of two conflicting sets of testimony. In the inquisitorial system, in which an expert is appointed by and gives evidence on behalf of the court, the judge essentially delegates an important part of (and in some cases most of) the decision-making function to the "witness", with relatively little opportunity for the parties to challenge (or, in many cases, even to address) the content of the evidence being given.

Second, all European civil procedure systems start from the basic premise that the dispute which is to be adjudicated is between one (or a very small number) of claimants against one (or an even smaller number of) defendants. For competition litigation, in particular against cartels, this paradigm does not hold true. The typical cartel case (if it is to be dealt with at all efficiently) will have a large number of claimants and almost certainly more than one party being pursued. The typical reaction of civil courts is to treat such cases as a bundle of bilateral disputes and to try them as such with varying, but until now relatively modest, success. This "bundling" phenomenon is seen in particular in the need for each claimant to issue proceedings separately in his or her name (and to pay a separate court fee), and to run the risks of litigation (e.g., the adverse costs risk) as if he or she were acting alone, rather than as part of a much wider group. This has had a chilling effect on competition claims, in particular for consumers and small businesses.

Third, the embryonic state of recognition and enforcement of civil judgments (especially for cartel and other competition claims) makes bringing an effective action for redress across Europe even more problematic. The structure of the current European rules (contained in Regulation 44/2001) has not been significantly modified since 1968 when the Brussels Convention (on which the Regulation is more or less wholly based) was made. The central paradigm of the Convention, unsurprisingly, is the bilateral dispute and generally its rules work well in such cases. But the system breaks down when multiple claimants face multiple defendants in what is essentially the same dispute. Although the Regulation permits courts (other than the court first seized of the dispute) to stay proceedings in subsequent claims brought relating to that dispute (but between different parties),[11] there is no requirement to do so. Furthermore, the coordination mechanisms between the national courts provided in the Regulation are weak (almost nonexistent). The prospect, as cartel enforcement increases, of a large amount of relatively uncoordinated cartel litigation across Europe in relation to essentially a single infringement is a very real one.[12]

II. Recent Policy Developments

The lack of competition redress in the European Union, in particular for smaller claimants, was highlighted in a report written for and published by the European Commission in mid 2004 ("Ashurst report").[13] The Ashurst report's sweeping but portentous conclusion that private enforcement of competition law across the European Union showed "astonishing diversity and a state of total under development" generated headlines. Although this conclusion was contested by those who pointed out that there was much more activity than the Ashurst report suggested, especially in the form of injunction proceedings and private set-

11. Council Regulation (EC) No. 44/2001 of 22 December 2000 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters, 2001 O.J. (L 12) 1, at art. 6(2).

12. Devenish Nutrition Ltd v. Sanofi-Aventis SA ("Vitamins litigation"), 2007 E.W.H.C. 2394 (Ch).

13. D. Waelbroeck et al., Study on the Conditions of Claims for Damages in Case of Infringement of EC Competition Rules: Comparative Report (Ashurst report) (2004), *available at* http://ec.europa.eu/comm/competition/antitrust/others/actions_for_damages/study.html.

tlements, the European Commission was nevertheless encouraged to carry forward its work in this area.

As a result, in December 2005, the Commission published a Green Paper which made a number of recommendations as to how to improve access for justice for cartel victims.[14] The proposals sent out for consultation included a recommendation that collective actions should be made easier, especially for consumers, and that incentives for claimants to come forward and litigate (if necessary) for them to obtain redress should be improved. Most controversially, the Commission suggested that the damages awarded could be doubled to compensate claimants for both their loss and the risk of bringing the claim in the first place—an idea modeled on (but not replicating) the U.S. treble damages system for cartel claims under the Sherman Act.

The responses to the Green Paper were considerable, varied, and vociferous (notably with respect to multiple damages). The Commission took some time to consider them and reflect on its policy aims and it was not until April 2008 that the Commission finally published its firm proposals. The incentives to claimants were remodeled and the controversial "copy" of the U.S. damages multiplier was dropped, but the proposals on collective actions were retained, refined, and made a little more detailed.

THE RESPONSES TO THE GREEN PAPER WERE CONSIDERABLE, VARIED, AND VOCIFEROUS (NOTABLY WITH RESPECT TO MULTIPLE DAMAGES).

The Commission now proposes two types of collective action, both of which it expects Member States to introduce. First, the Commission recommends a group action where claimants opt-in to a claim against a cartel and which Commission officials have suggested might be particularly useful for claims by small- to medium-sized enterprises. Second, the Commission proposes a "Community-wide" representative action requirement, where consumer and other representative bodies, designated by their Member State government or the court in which the action is brought, can bring claims on behalf of groups of affected consumers or small businesses. The Commission believes that this type of collective action would be particularly well-suited to consumer claims or claims in which there is a defined group of consumers (rather than a listed group of individuals) that are represented unless they opt-out. Variations of both of these collective actions exist in a number of Member States (as the Commission has emphasized) and Commission officials have made an effort to point out that these actions are also very different from the U.S. "class action" system.

The Commission's changed view on the desirability of multiple damages may have been conditioned not only by the opposition voiced to it, but also by devel-

14. Commission Green Paper on Damages Actions for Breach of the EC Antitrust Rules [hereinafter "Green Paper"], COM(2005) 672 final (Dec. 19, 2005).

opments in the case law of the European Courts. In particular, the ECJ decided in the *Manfredi* case[15] that both pre-judgment interest (running from the time the competition law infringement took place until judgment) and compensation for loss of future profit (lucrum cessans) were recoverable as a matter of European law, which takes precedence over contrary national laws across the European Union. Where cartels have been ongoing for a number of years, the interest on the overcharge (as well as any loss of future profit) can sometimes double the original overcharge paid to the cartel.

In parallel with the initiatives at a European level (and encouraged by them), a number of the EU's Member States have also begun considering the issue of collective redress for cartel victims—either as a standalone policy issue or (more usually) as part of an overall consideration of collective redress in their civil procedure systems.

In the United Kingdom, the Office of Fair Trading ("OFT") published a discussion paper in April 2007 recommending that the existing, and limited, form of representative action[16] for "follow-on" cartel action in the Competition Appeals Tribunal ("CAT") should be extended. At present, only pre-designated consumer bodies can bring such claims on behalf of named consumers. Since only one consumer body has been designated since the law came into force in 2003,[17] and only one consumer claim has been brought against a cartel (in the sale of soccer shirts to football fans), there is concern that the existing system is not working well.

The OFT therefore recommended that representative actions in the competition area should be extended in a number of ways:

- the court, as well as the government, could decide if a body is an appropriate one to bring a representative action;
- representative actions should be available not only to consumers, but also to small businesses who have suffered harm from an anticompetitive practice;
- the representative mechanism should be available not only in claims which "follow-on" from a competition authority's decision, but also in claims which "stand alone"; and
- the bodies should be able to represent not only named claimants, but also claimants who fall within a defined group.

15. Joined Cases C 295-298/04, Vincenzo Manfredi v. Lloyd Adriatico Assicurazioni [hereinafter *Manfredi*], 2006 E.C.R. I-6619, at para. 5(2).

16. U.K. Office of Fair Trading, Private actions in competition law: effective redress for consumers and business (OFT Discussion paper, No. 916, Apr. 2007).

17. The U.K. Consumers' Association "Which?", *at* http://www.which.co.uk/.

Following a period of consultation, the OFT slightly refined and confirmed its view in recommendations to the government in November 2007. As of October 2008, the publication date of this article, the government's response was still pending.

Other Member States have also been considering collective actions for compensation, with competition law usually included in a wider initiative to promote collective redress. For example, in France, a number of legislative texts (some sponsored by the government) have been put forward in the last two years or so, although none of them have been adopted by the National Assembly. Italy has also recently passed a law on representative actions, although its commencement has been suspended by the newly elected government pending further consultation.[18] More Member States have laws which promote collective redress in particular areas such as Germany's capital markets law (Kap MuGe),[19] Portugal's "popular action" for consumer, public health, and environmental actions,[20] and the Netherlands "mass tort settlements" law,[21] which is general in scope, but only applies where the parties have agreed to settle a case (its fairness is then endorsed by the court in Amsterdam and applied to all those in the represented group).

III. Towards a European Cartel Damages System?

Will the policy impact, in particular at the European level, reverse the situation of "total under development" found in the Ashurst report? The "astonishing diversity" also found there applies very much to the civil procedure rules of the (now 27) EU Member States.

Clearly the proposals in the Commission's White Paper will improve the situation if they are carried through, but there are limits on the EU's competence in this area. Civil procedure coordination across Europe is done through cooperation, notably through Regulation 44/2001 and the Regulations on the law applicable to contractual and non-contractual obligations.[22] There are very few cases

18. Law 244/2007, L'azione collettiva risarcitoria ("la ACR"), Il Codice del Consumo, sub-Art. 140-bis (2007).

19. The Bundestag, Germany, Act on the Initiation of Model Case Proceedings in respect of Investors in the Capital Markets (Nov. 1, 2005).

20. Lei No. 83, Lei de Acção Popular ("LAP") (Aug. 31, 1995).

21. The Netherlands, Act on the Collective Settlement of Mass Damage Claims (Aug. 1, 2005).

22. Regulation (EC) No. 593/2008 of the European Parliament and of the Council of 17 June 2008 on the law applicable to contractual obligations (Rome I), 2008 O.J. (L 177) 6; and Regulation (EC) No. 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations (Rome II), 2007 O.J. (L 199) 40.

in which EU law instruments have interfered with Member States' choice of how they ensure the enforcement of European law in their internal legal order[23] and, for the most part, these have related as much to criminal law as civil enforcement. This lack of interaction sits uncomfortably with public competition law enforcement which is increasingly being carried out (in fact, if not in law) on a centralized European basis.

The European public enforcement regime was dramatically overhauled beginning on May 1, 2004 in accordance with Regulation 1/2003. Although on its face it devolves more competition enforcement to the Member States' national competition authorities by giving them the power to fully apply EC competition law (and a number of them have taken advantage of this), it has also had two, possibly rather more surprising consequences. First, it encourages the "soft" harmonization of public enforcement procedures across the European Union as Member States increasingly adopt an EU-style enforcement process for their national competition authorities. Second, despite the new "devolution", the European Commission is enforcing EU competition law, particularly against cartels, as energetically as ever, at least as measured by the amount of fines levied.[24]

THIS SHARP CONTRAST BETWEEN AN INCREASINGLY CENTRALIZED PUBLIC ENFORCEMENT PROCESS AND A SHARPLY "DECENTRALIZED" CIVIL REDRESS MECHANISM ACROSS EUROPE IS HEIGHTENED BY THE ALMOST TOTAL ABSENCE OF ANY TOOLS FOR COORDINATING THE TWO SYSTEMS.

This sharp contrast between an increasingly centralized public enforcement process and a sharply "decentralized" civil redress mechanism across Europe is heightened by the almost total absence of any tools for coordinating the two systems. Regulation 1/2003 makes brief reference to the issue by giving national courts the power to apply EU competition law in full in cases before them,[25] on the condition that they do not take decisions which may conflict with any decisions made or expected from the European Commission.[26] The latter provision reflects the case law of the ECJ,[27] and it is the case law of the Court which otherwise provides the (relatively thin) glue which holds the two systems in some relationship to each other.

23. Directive 98/27/EC of the European Parliament and of the Council of 19 May 1998 on injunctions for the protection of consumers' interests, 1998 O.J. (L 166) 1; and Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights, 2004 O.J. (L 157) 1.

24. *See supra* note 9.

25. Council Regulation (EC) No. 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. (L 1) 1, at art. 6.

26. *Op. cit.* at art. 16.

27. Case C-344/98, Masterfoods, 2000 E.C.R. I-11369.

In addition to the "no conflict" rule (or, perhaps more accurately, the "no contradictory action" rule) set out in Regulation 1/2003, the European Court has relied on two guiding principles in assessing whether national civil procedure rules properly allow the enforcement of EU law: the principle of equivalence and the principle of effectiveness.[28] The principle of equivalence means that, for the enforcement of directly effective EU law obligations in the national courts, all remedies available for similar breaches of domestic law must also be available. The principle of effectiveness means that, even where those remedies are applied, they must be capable of producing a result. National law "must not render impossible, or excessively difficult" the exercise of EU law rights.[29]

In principle, the European institutions leave the choice and form of civil procedure to each Member State subject to some very high-level benchmarking against relatively few general principles. Recently, the European Courts have shown signs of using the principle of effectiveness in a more interventionist manner, but the fundamental principle remains untouched.

Effectively, the EU acts as a confederation rather than a union in civil procedure matters. Given the huge diversity both in current civil procedure laws and in legal traditions, this is probably inevitable and to be welcomed. But there are significant issues over the coordination mechanism (discussed earlier in this article) which, despite the non-interventionist stance the European Union currently takes, nevertheless needs to be addressed if the Commission's policy objective of improving access to justice is to succeed.

IV. Possible Solutions to These Issues

Before turning to propose some possible solutions to the issues raised above, it is worth considering briefly what the aim of any changes should be. As with the overriding objective in the English civil procedure rules,[30] the aim of any EU-approved change should be to deal with claims justly. This means not only having regard to the eventual outcome, but also to the speed and economy of the process and, therefore, to its efficiency on a European level.

The ideal should, therefore, be one civil claim per cartel (or other infringement) in a single EU Member State court. The outcome of that claim should be binding and easily enforceable across the European Union. This of course implies less choice of court for claimants and therefore needs to be accompanied by some guarantees that effective redress will be available in the competent court. The

28. *Manfredi, supra* note 15, at para. 2(2).

29. *Id.* at para. 3.

30. U.K. Ministry of Justice, The Civil Procedure Rules (2008), at §§ 1.1(1) & 1.1(2).

reminder of this article suggests what might be appropriate after considering how best to determine which court should hear the case.

Although the question of which court should hear the case is the first issue for claimants, the choice of which law the court should apply is also highly relevant. The recently adopted Regulation 864/2007 provides that, for competition claims, the applicable law shall be the law of the Member State where competition is or is likely to be affected by the infringement.[31] If that leads to the choice of more than one law, then the claimant may instead choose the law of the court seized (provided the law chosen has a real connection with the infringement alleged). The choice of court therefore has a substantial effect not only on the civil procedure rules to be followed, but also on the substantive law to be applied.

How best then should the court competent to hear the case be chosen? The basic premise of the current Regulation 44/2001 is that the place of the defendants' domicile is the appropriate forum for the case to be heard. It is only where there is more than one defendant domiciled in different Member States that a choice of court comes into play.

Most claims for redress for competition law breaches will be brought as a breach of a non-contractual obligation (tort, delict, or quasi-delict). Regulation 44/2001 (as interpreted by the European Court of Justice) currently provides that the claim may be brought in a court where the harmful event occurred or may occur. This has been held to be either the court of the place where the tortious act took place or the courts of the place where the effect of the act occurred. The application of those rules can lead to a wide range of available jurisdictions in cartel cases—a way needs to be found to narrow the choice.

It is suggested that there are three possible options for criteria to choose the competent court: First, and most mechanistically, a turnover-based test could be used. Where an infringer (or its EU subsidiaries) earned turnover in the products affected by the infringement through activities in a number of EU Member States, the country in which the highest such turnover was earned would have jurisdiction. Where there are a number of infringers (e.g., in a cartel), there appear to be two choices: either the courts of the country with the highest total turnover of the affected products or services, or the courts of the domicile of the defendant cartelist with the highest such turnover.

The difficulty with such a test is obtaining a complete set of the relevant turnover figures before proceedings have begun. For follow-on cartel actions, it is possible that the competition authorities' decision will contain the relevant information if it is needed to calculate fines (many competition authorities use this as a metric when setting fining levels). For standalone actions and those fol-

31. Commission Regulation (EC) No. 864/2007 of 11 July 2007 on the law applicable to non-contractual obligations (rome II) L199/40 at art. 6(3).

low-on actions where turnover information is not available, pre-action disclosure in a likely forum may be the only way forward.

As with the choice of jurisdiction in the merger control context, where turnover thresholds are used to determine whether a merger has a "community dimension" giving the European Commission exclusive oversight of it, turnover-based thresholds may give early certainty as to which body is entitled to deal with the case. However, given that there is no EU-wide authority capable of handling borderline cases, this certainty may come at the price of some (and possibly a significant number of) inappropriate choices of jurisdiction, in particular in cases where many of the infringers have their main turnover outside the European Union.

Second, a refined version of the current "effects-based" test could be applied. At present, the Member State court first seized could in principle hear the entire compensation claim as it applies to that infringement. Rather than rely on claimants having the best or most comprehensive claims to get to the door of their chosen court first, it may be more realistic for the court where the greatest effect was felt from the infringement first to hear the claim. This could be the court of the place where the known claimants suffered the most loss—although this may be extremely difficult to quantify—or, more simply, but possibly more arbitrarily, the place where (after a suitable period of, say, 6 to 12 months from the first claim being brought) the claimant(s) with the largest total claimed loss bring their claims, provided that that forum was in some way affected by the infringement. Again, however, arbitrary results cannot be ruled out.

None of the possible bases for attributing jurisdiction is, therefore, without serious drawbacks, and the trade-off between certainty and the effectiveness of redress is a difficult one.

The third option would be to use the place of the act which caused the infringement (e.g., the first cartel meeting), which could be the best forum on the basis that the evidence relating to the infringement may best be found there. However, the problem of determining where the infringing act occurred in competition cases should not be underestimated. This is particularly true for long-running cartels, where the cartel meetings may have taken place in various international locations (some of them not in the European Union), or for abuses of a dominant position, where it may not be easy to find where the decision or the action constituting the abuse was initiated.

None of the possible bases for attributing jurisdiction is, therefore, without serious drawbacks, and the trade-off between certainty and the effectiveness of redress is a difficult one. Without some reduction in the availability of "all courts" where a defendant or its subsidiaries has a domicile to hear competition claims, the possibility of ongoing claims being made across the European Union for many years is too real to be sensible. A less radical change would be to supplement the current basic and existing rule in Regulation 44/2001. The "domicile" rule could be amended so that, where more than one country is indicated by the defendants' domicile, only the court where either the largest amount of infringing turnover was earned (either by the cartel as a whole or a reasonably representative proportion of it if the turnover information is not fully available) or, where different, the court of the place in the European Union where the claimant can prima facie show that the cartel or abuse had its most significant effect. For this solution to work effectively, courts across the European Union will need to be willing to grant pre-action disclosure of the relevant turnover information to prospective claimants to establish the turnover test. Clearly some form of improved mechanism for publicizing which competition claims are also pending (and where in the European Union they are pending) is needed for this system to work optimally.

It is unlikely, however, that anything as developed as the U.S. Multi-District Litigation Panel, where cases are allocated at a federal level to the various U.S. Federal District Courts around the United States, would be sensible in an EU context. Other jurisdictions which have developed good collective redress mechanisms have not found it necessary to introduce such a system.

There are certain minimum conditions which will need to be met before the European Commission can sensibly reconcile its policy of encouraging private redress actions with the need to streamline and simplify the civil enforcement mechanism across Europe. First, all of the courts which are eligible to take a competition claim in each Member State must have civil procedure rules which meet the minimum standards as set out in the Commission's White Paper (or the legislation which will flow from it). Courts in countries which do not meet these criteria should not be within the system of handling jurisdiction and enforcement of competition claims discussed earlier in this article. If claimants are to lose the ability to choose from a potentially significant number of courts in which to bring their claims, and thus themselves to make a choice between the effectiveness of competing civil procedure rules, then the court to which their claim is directed *must* be capable of providing them with a proper remedy within a reasonable time.

Second, where the court chosen is one where the language of the proceedings is likely to be different from the language of the majority of the evidence, the court should be prepared to accept evidence in the original language, rather than put to the claimants or defendants the cost of experts to translate it. Of course, some expenditure may be required for language training for judges (or for court interpreters and translators), and the claims may need to be heard before a panel of specialists within the general civil courts for this reason. Some countries (e.g., the United Kingdom and Italy) have already taken steps to direct competition cases falling within their country's jurisdiction to particular courts or tribunals, a trend which should be encouraged.

Third, given that many claimants in competition cases will be smaller enterprises and consumers, who will likely not have the funds to launch an action without assistance from third parties, providing a level playing field for the parties will require third-party litigation funding in all courts for which competition claims can be allocated. Claims are not normally brought (at least on a collective basis) unless a funding mechanism has been put in place. It is probably sensible to provide that, as long as a particular funding mechanism permitted under the (EU Member State) law properly applies to it, it should be recognized and given effect in any court in the European Union. As with the "single passport" home state regulation system for financial services in the European Union,[32] this would, in addition to allowing claimants to support their claims financially, expand the available sources of funds to claimants which, in turn, can be expected to improve their chances of gaining compensation.

Finally, all Member States will need to allow the publicity necessary to make sure that all claimants are aware of the claim being brought in the allocated court. This will mean that those Member States (e.g., France) which currently have strict bans on soliciting for litigation will need to relax them—at least so far as directed by the court properly seized—to permit residents in their countries to be aware of and participate in the claim if they are affected by it.

V. Conclusion

The diversity of civil procedures and therefore of effective outcomes for cartel victims noted in the Ashurst report four years ago,[33] is the direct result of two or more centuries of different legal history across Europe. Clearly, even moving a little towards a more coordinated approach to achieving redress for victims of European competition law breaches will take time. But such moves need to happen if redress for the victims of Europe-wide cartels and other anticompetitive practices is to be made available. The operation of Regulation 44/2001 is due to be reviewed in the near future and may give a good opportunity to revisit these issues.

Clearly, even moving a little towards a more coordinated approach to achieving redress for victims of European competition law breaches will take time.

If no progress can be made on a coordinated approach, then it may be worth considering direct access to the European court system for those seeking redress—at least in those cases where a breach of EC competition law has been found. However, this

32. Directive 2004/39/EC of the European Parliament and of the Council of 21 April 2004 on markets in financial instruments amending Council Directives 85/611/EEC and 93/6/EEC and Directive 2000/12/EC of the European Parliament and of the Council and repealing Council Directive 93/22/EEC, 2004 O.J. (L 145) 1, as subsequently amended.

33. *Supra* note 15.

is likely to require an amendment to the Treaty of Rome, and as the recent experience of treaty amendments has shown, this route is also far from straightforward.

Some incremental change is very likely to happen soon, through a combination of legislation (or recommendations or similar) at both the EC and national level and through European Court jurisprudence. Whether these developments will lead to a coherent system for claiming redress across Europe which effectively meets the needs of both claimants and defendants remains to be seen.

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

The Road to the Commission's White Paper for Damages Actions: Where We Came From

Assimakis P. Komninos

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

The Road to the Commission's White Paper for Damages Actions: Where We Came From

*Assimakis P. Komninos**

The European Commission's April 2008 White Paper on Damages Actions for Antitrust Violations is a groundbreaking development. It marks the establishment of a system of private antitrust enforcement system in Europe, which, however, does not imitate the U.S. example but is rather "European" in its conception, origins, and main parameters. To help understand the White Paper proposals, it is imperative to review its origins (i.e., where we came from). This article aims at presenting the jurisprudential developments in Europe that created the right atmosphere for the White Paper to come in existence. The review of these developments explains the main qualities and basic premises of the White Paper. In particular, it explains the fundamental choice to depart from the U.S. solution and instead opt for allowing both offensive and defensive passing-on.

*The author practices law in Brussels with White & Case LLP and is a visiting lecturer at IREA - Université Paul Cézanne Aix - Marseille III and at University College London (UCL). He is the author of *EC Private Antitrust Enforcement, Decentralised Application of EC Competition Law by National Courts* (Hart Publishing, Oxford/Portland, 2008). The present views are strictly personal to the author.

I. Introduction

The publication in April 2008 by the European Commission of the long-awaited White Paper on Damages actions for breach of the EC antitrust rules ("WP")[1] has again brought EC private antitrust enforcement to the forefront. This article attempts to explain how we got to the WP in the first place and how European Community law, in particular the European Court of Justice's ("ECJ's") rulings, have set the ground for the latest developments. It also proceeds to an appraisal of the WP and of some specific issues on the basis of existing Community law, while referring to some interesting developments in the EU member states.

II. The Road to the White Paper

A. THE "EUROPEAN" CONTEXT OF ANTITRUST RIGHTS AND REMEDIES

The application of EC competition law by civil courts, though not particularly developed in Europe, has not been a recent phenomenon. Indeed, the very first preliminary reference made by a national court to Luxembourg under the old Article 177 EEC, was a competition case where EC competition law arose in the context of private litigation.[2] Of course, the mere application of the competition rules by national courts cannot be said to amount to a system of private antitrust enforcement. The term "enforcement" signifies an instrumental role of private actions in the sense of the private litigants not just seeking redress, but also in effect becoming themselves actors in enhancing the overall efficiency and effectiveness of the competition enforcement system. It is only very recently that private antitrust enforcement appears for the first time as a meaningful complement—though certainly not an alternative—to public enforcement.

1. Commission White Paper on Damages Actions for Breach of the EC Antitrust Rules, COM(2008) 165 final (Apr. 2, 2008). This comes as a follow up after the publication in December 2005 of a Green Paper, but is also a prelude for Community legislation. In EU jargon, a "White Paper" is a document containing proposals for Community action in a specific area. It sometimes follows a "Green Paper" published to launch a consultation process at European level. While Green Papers set out a range of ideas presented for public discussion and debate, White Papers contain an official set of proposals in specific policy areas and are a prelude to Community legislation. The WP itself is a rather short document that in reality summarizes the far more developed Staff Working Paper ("SWP") (Commission Staff Working Paper Accompanying the White Paper on Damages Actions for Breach of the EC Antitrust Rules, SEC(2008) 404 (Apr. 2, 2008)). An impressive 600-page Impact Assessment Study (CEPS, EUR & LUISS, Making Antitrust Damages Actions More Effective in the EU: Welfare Impact and Potential Scenarios, Final Report (Dec. 21, 2007)) is itself usefully summarized in a Commission Impact Assessment Report (Commission Staff Working Document Accompanying the White Paper on Damages Actions for Breach of the EC Antitrust Rules, Impact Assessment, SEC(2008) 405). In most cases in the present article, references made to the "WP" cover the whole of the recent Commission policy initiative and not just the short document which bears that title.

2. Case 13/61, Kledingverkoopbedrijf de Geus en Uitdenbogerd v. Robert Bosch GmbH and Maatschappij tot voortzetting van de zaken der Firma Willem van Rijn, 1962 E.C.R. 45.

The road was opened by the modernization and decentralization European reforms between the years 1999 and 2004. But it has also come as a consequence of groundbreaking rulings by the ECJ, which has extended the scope of remedies available to individuals by Community law to cover also individual civil liability and has always imposed stringent conditions on national substantive and procedural law, in order to ensure the effectiveness of the EC competition rules. The ECJ has indeed been particularly bold in this field due to mainly historical reasons: the Community competition rules have long been recognized as having "horizontal direct effect" (i.e., they apply to legal relationships between individuals), and at the same time they have been treated with a high degree of deference as part of the Community's "economic constitution", thus enjoying an increased normative value.

> AT THE HEART OF PRIVATE ANTITRUST ENFORCEMENT IN EUROPE LIES THE QUESTION OF THE RELATIONSHIP BETWEEN COMMUNITY AND NATIONAL LAWS.

At the heart of private antitrust enforcement in Europe lies the question of the relationship between Community and national (i.e., EU member states') laws. At the current stage of European integration, rights and obligations emanating from Community law are in principle enforced under national law and before national courts. The Community legal order is not a federal one and the Community acts only within the limits of the powers conferred upon it by the EC Treaty. The Community standard is that Community law is enforced primarily by having recourse to national administrative, civil, and criminal law before national administrative authorities and national courts.

Thus, speaking about private law disputes, on the side of substance, there is no Community law of contract, tort, or unjustified enrichment, or a European Civil Code. Indeed, even if the Community had the power or intention to legislate in such a vast cross-sector area, it would be almost impossible to arrive at a common denominator applicable throughout the EU member states, taking into account the century-long divisions in the European legal systems and families. Equally, on the side of procedure, there are no Community courts of full jurisdiction that could apply Community law and deal with Community law-based claims. Thus, national courts act also as "Community courts" of full jurisdiction (juges communautaires de droit commun).

It is true that in the last twenty years much has changed, and one can now speak of a positive integration drive to unify or harmonize rules on remedies and procedures. However, with very few exceptions, these are sectoral rules applying to some very specific Community objectives and the reality remains that there are no cross-sector Community rules of administrative or civil law dealing with the enforcement of Community law-based rights.[3]

3. *See* A.P. Komninos, EC Private Antitrust Enforcement, Decentralised Application of EC Competition Law by National Courts 142-44 (2008).

Consequently, natural and legal persons relying on Articles 81 and 82 EC would have no other means to pursue their civil claims but through access to national courts and laws.[4] However, the substantive and procedural conditions of civil antitrust enforcement can be very different in Europe depending on which national law applies and which national court adjudicates. Inconsistencies and inadequacies in national laws on remedies and procedures are certainly a source of serious concern, not just for EC competition law but for Community law in general. In this context, the problem can be identified in three different, albeit interconnected levels.

First, there is a problem of effective or adequate judicial protection (i.e., the effective protection of Community rights). This is a principle not only of Community law but also of human rights. Indeed, effective judicial protection in the form of access to the courts configuration derives from Articles 6 and 13 of the European Convention of Human Rights. As far as Community law is concerned, and as the ECJ has recognized, Articles 81 and 82 EC "tend by their very nature to produce direct effects in relations between individuals [and] create direct rights in respect of the individuals concerned which the national courts must safeguard." Failure to afford this safeguard "would mean depriving individuals of rights which they hold under the Treaty itself."[5]

Second, there is a problem for the effectiveness of the whole system of Community law as such and, more particularly, for the efficiency of the Community (competition) rules. There are two facets to this. One is Community law-specific and the other is competition law-specific. The first facet of the problem is that when citizens pursue their Community rights before the "juges communautaires de droit commun," in addition to serving their private interests, they are also instrumental for and indirectly act in the Community interest, becoming "the principal 'guardians' of the legal integrity of Community law in Europe."[6] The "direct effect" doctrine was developed partly with this consideration in mind. The second competition law-specific facet refers to the "private attorney general" role of individuals in antitrust cases. In a mature antitrust system, private enforcement is a necessary complement of public enforcement and by no means inferior or weaker. In such a system, private actions are crucial for the efficiency of the system as a whole.[7]

4. However, *see* the new line of case law, in particular, the *Courage* and *Manfredi* rulings (*infra*).

5. Case 127/73, Belgische Radio en Televisie and Societé Belge des Auteurs, Compositeurs et Editeurs de Musique v. SV SABAM and NV Fonior (I) [hereinafter *BRT v. SABAM I*], 1974 E.C.R. 51, at paras. 16 & 17.

6. *See* J.H.H. Weiler, The Constitution of Europe, 'Do the New Clothes have an Emperor?' and Other Essays on European Integration 20 (1999).

7. On this particular point, *see* the analysis of the ECJ's *Courage* ruling (*infra*).

Third, the disparities and inadequacies of national legal systems offend against the principle of consistent and uniform application of Community law. Such discrepancies are particularly regrettable from an EC competition law point of view, because they tend to create variations in the costs of enforcing the EC antitrust rules, and thus lead to unequal conditions of competition among the member states.

In the decentralized system of EC antitrust enforcement, the problem is exacerbated. Competitors and economic actors in general take the likelihood of public or private antitrust action seriously into account in defining their market strategies. In this context, damages have an especially powerful impact on business behavior. An economic operator's exploitation of its "immunity" from civil actions in damages and failure to compensate victims adequately in one jurisdiction, as opposed to other jurisdictions where companies are constantly successfully or unsuccessfully defending civil antitrust actions, is hardly compatible with the creation of "a level playing field for agreements, decisions by associations of undertakings and concerted practices within the internal market," as Regulation 1/2003 propagates.[8]

The ECJ has consistently recognized the "procedural and institutional autonomy" of the member states to identify the remedies, courts, and procedures that are necessary for the exercise of Community law rights at the national level. The term "procedural autonomy" creates the incorrect impression that this principle refers only to national rules of civil, administrative, or criminal procedure. In fact, its scope is much larger and covers all substantive or procedural mechanisms at the national level that can be used for the enforcement of Community law. That is why the term "remedial/procedural autonomy" is preferable. More importantly, however, the Court has also imposed demanding Community limits and safeguards on that autonomy. These are the principles of equality and effectiveness.[9] The first principle means that the enforcement of Community law at the national level should not be subject to more onerous procedures than the enforcement

8. 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. (L 1) 1, at Recital 8. *See also* Recital 1, which speaks of the necessity for Arts. 81 and 82 EC to "be applied effectively and uniformly in the Community"; and the Impact Assessment Form of the September 2000 Regulation proposal (Commission Proposal for a Council Regulation on the Implementation of the Rules on Competition Laid Down in Articles 81 and 82 of the Treaty and Amending Regulations (EEC) No. 1017/68, (EEC) No. 2988/74, (EEC) No. 4056/86, & (EEC) No. 3975/87 ("Regulation Implementing Articles 81 and 82 of the Treaty"), COM(2000) 582 final, 2000 O.J. (C 365) 56), where reference is made to a "level playing field for companies in the internal market by ensuring more widespread application of the Community competition rules."

9. *See*, *e.g.*, Case 33/76, Rewe-Zentralfinanz eG and Rewe-Zentral AG v. Landwirtschaftskammer für das Saarland, 1976 E.C.R. 1989, at para. 5; Case 45/76, Comet BV v. Produkschap voor Siergewassen, 1976 E.C.R. 2043, at paras. 12-13; Case 130/79, Express Dairy Foods Ltd. v. Intervention Board for Agricultural Produce, 1980 E.C.R. 1887, at para. 12; Case 199/82, Amministrazione delle Finanze dello Stato v. San Giorgio, 1983 E.C.R. 3595, at para. 12; and Case C-261/95, Rosalba Palmisani v. Istituto Nazionale della Previdenza Sociale (INPS), 1997 E.C.R. I-4025, at para. 27.

of comparable national law. The second principle is a much more demanding test. It means that although Community law-derived rights will have to count on national substantive and procedural remedies for their enforcement, such remedies still have to be effective and must not render the exercise and enforcement of those rights impossible or unjustifiably onerous. It reflects a more general guiding principle of Community law: full and useful effectiveness (effet utile).

Undoubtedly those two requirements make national divergences less burdensome. The ECJ has, nevertheless, proceeded further than that. Starting with such cases as *Francovich*, *Factortame I*, and *Zuckerfabrik Süderdithmarschen*,[10] it has also recognized the existence of certain autonomous Community law remedies for Community law-based rights, and has delegated to national law only the very specific conditions for their exercise, as well as the procedural framework rules, always within the limitations of equality and effectiveness. In doing so, it has been guided by the principle ubi ius, ibi remedium, under which a Community law right must be protected through an appropriate corresponding remedy. It has relied on "the full effectiveness of Community rules and the effective protection of the rights which they confer" and on the duties that Article 10 EC imposes on member states and their judicial organs.[11]

UNTIL 2001, THE ECJ NEVER HAD THE OPPORTUNITY TO RULE ON THE ISSUE OF CIVIL LIABILITIES ARISING FROM THE VIOLATION OF EC COMPETITION RULES

Professor Van Gerven, a former Advocate General ("AG") of the ECJ and eminent scholar of Community law, has proposed a more global approach to the issue of remedies in Community law, thus stressing the requirement of effective judicial protection which better describes the Court's case law on remedies. Van Gerven speaks of four already existing Community substantive remedies: a general one, to have national measures that conflict with EC law set aside; and three specific ones, compensation, interim relief, and restitution.[12] Individual civil liability is integrated in the first limb of these three specific remedies, beside its admittedly much more developed sibling, state liability. Van Gerven further makes a distinction between the "constitutive" and "executive" elements of remedies. The first pertain to the principle of the remedy as such; the second to its "content and extent". The first type of elements must be uniform, since they are entirely con-

10. Joined Cases C-6/90 & C-9/90, Andrea Francovich et al. v. Italy, 1991 E.C.R. I-5357; Case C-213/89, Regina v. Secretary of State for Transport, ex parte Factortame Ltd. et al. (I), 1990 E.C.R. I-2433; Joined Cases 143/88 & C-92/89, Zuckerfabrik Süderdithmarschen AG v. Hauptzollamt Itzehoe and Zuckerfabrik Soest GmbH v. Hauptzollamt Paderborn, 1991 E.C.R. I-415, respectively.

11. Joined Cases C-46/93 & C-48/93, Brasserie du Pêcheur SA v. Germany and Regina v. Secretary of State for Transport, ex parte Factortame Ltd. et al. (III), 1996 E.C.R. I-1029, at para. 39.

12. *See* W. Van Gerven, *Of Rights, Remedies and Procedures*, 37 C.M.L.R. 501, 503 (2000).

nected with the Community "right" of which individuals avail themselves. The executive elements, on the other hand, may to a certain extent be governed by national law, but only under more substantial Community requirements. For these elements, Community law should require an "adequacy test", rather than a mere "minimum effectiveness" or "non-impossibility" test which may continue to apply for simple procedural rules.[13]

B. WHEN THE ECJ SPOKE

1. *Courage v. Crehan*

Until 2001, the ECJ never had the opportunity to rule on the issue of civil liabilities arising from the violation of EC competition rules, although in some instances it referred to possible damages and other civil claims that private parties could pursue before national courts,[14] but without addressing the question of the Community law or national legal basis.

Earlier, then-AG Van Gerven, in his opinion in *Banks*,[15] had argued extensively in favor of recognizing a Community right to obtain reparation in cases where loss and damage are sustained as a result of an undertaking's infringement of the directly effective Community competition rules.[16] In his carefully structured opinion, the AG had considered that the general basis established by the Court in *Francovich* also applied to the case of "breach of a right which an individual derives from an obligation imposed by Community law on another individual." In competition law, in particular, the AG observed that such a Community right to damages would make the Treaty antitrust rules "more operational", adducing an argument from the U.S. system of antitrust enforcement, where civil suits for damages have played a dominant role.[17] In *Banks*, however, the Court declined to address all these fundamental issues, because it reached the conclusion that the only set of rules applicable to the facts, Articles 65 and 66 ECSC, did not have a direct effect.

The fundamental issue of the Community or national law basis of the right to damages in EC competition law violations was finally addressed by the ECJ in its

13. *See* Van Gerven (2000), *id.*, at 502-04 & 524-26.

14. This was already implicit in *BRT v. SABAM I*, *supra* note 5, at paras. 16 & 22. Reference should also be made to Case C-242/95, GT-Link A/S v. De Danske Statsbaner, 1997 E.C.R. I-4349, at para. 57; and Case C-282/95 P, Guérin Automobiles v. Commission, 1996 E.C.R. I-1503, at para. 39.

15. Case C-128/92, H.J. Banks & Co. Ltd. v. British Coal Corporation [hereinafter *Banks*], 1994 E.C.R. I-1209.

16. AG Opinion, *Banks*, *id.*, at paras. 37 et seq.

17. *Op. cit.* at para. 44.

September 20, 2001 *Courage* ruling.[18] In *Courage*, the Court recognized a right to damages as a matter of Community rather than national law, and stressed the fundamental character of the EC competition rules in the overall system of the Treaty.

The facts of *Courage* were rather undistinguished. Breweries in Britain usually own pubs which they lease to tenants, while the latter are under contractual obligations to buy almost all the beer they serve from their landlords. In 1991, Mr. Bernard Crehan signed a 20-year lease with Courage Ltd. whereby he agreed to buy a fixed minimum quantity of beer exclusively from Courage, while the brewery undertook to supply the specified quantities at prices shown in the tenant's price list. The rent was initially lower than the market rate and it was subject to a regular upward review, but it never rose above the best open market rate. In 1993, Mr. Crehan and other tenants fell into financial arrears, blaming Courage's supply of beer at lower prices to other non-tied pubs ("free houses") for their situation. In the same year, Courage brought an action for the recovery from Mr. Crehan of sums for unpaid deliveries of beer. Mr. Crehan, alleging the incompatibility with Article 81(1) EC of the clause requiring him to purchase a fixed minimum quantity of beer from Courage, counterclaimed for damages.

There were two specific obstacles to Mr. Crehan's success. The first one was that according to earlier case law, Article 81 EC had been interpreted as protecting only third parties, (i.e., competitors or consumers), but not co-contractors (i.e., parties to the illegal and void agreement). The second issue was that under English law a party to an illegal agreement, as this was considered to be by the Court of Appeal, could not claim damages from the other party. This was as a result of the strict construction English courts were giving to the nemo auditur turpitudinem propriam (suam) allegans or in pari delicto potior est conditio defendentis or ex dolo malo non oritur causa rule, which in essence meant that Mr. Crehan's claim in damages would fail because he was co-contractor in an illegal agreement.

The ECJ, following the ruling in *Francovich* which had recognized the principle of state liability as a principle of Community law, and also relying on its *Eco Swiss* ruling,[19] stressed the primacy of Article 81 EC in the system of the Treaty, since it "constitutes a fundamental provision which is essential for the accomplishment of the tasks entrusted to the Community and, in particular, for the functioning of the internal market."[20] It also stressed, with particular reference to "the possibility of seeking compensation for loss caused by a contract or by conduct liable to restrict or distort competition," the task of national courts to

18. Case C-453/99, Courage Ltd. v. Bernard Crehan [hereinafter *Courage*], 2001 E.C.R. I-6297.

19. Case C-126/97, Eco Swiss China Time Ltd. v. Benetton International NV, 1999 E.C.R. I-3055.

20. *Courage*, *supra* note 18, at para. 20.

ensure the full effect (plein effet) of Community rules and the protection of individuals' rights conferred by those rules. The full effectiveness (pleine efficacité) of the Treaty competition rules and, in particular, "the practical effect [effet utile] of the prohibition laid down in Article [81(1)]," would be put at risk if individuals could not claim damages for losses caused by the infringement of those rules. The instrumental character of such liability for the effectiveness of the law as such is more than evident in this passage, exactly as was the case with state liability in *Francovich*.[21] And finally, the Court dispelled any doubt as to its pronouncement:

> "Indeed, *the existence* of such a right strengthens the working of the Community competition rules and discourages agreements or practices, which are frequently covert, which are liable to restrict or distort competition. From that point of view, actions for damages before the national courts *can make a significant contribution to the maintenance of effective competition in the Community*."[22] (emphasis added)

This last quote makes it clear that the meaning of effectiveness in *Courage* has a double facet. It refers not only to Community law in general, but also to the specific field of antitrust. This is clear from the Court's use of the term "significant contribution" to refer to the role of damages claims for the efficiency of antitrust enforcement in Europe, with a view to maintaining effective competition. More authoritative words in favor of private enforcement and the "private attorney general" role[23] of the civil litigant could hardly be pronounced.

2. *Manfredi v. Lloyd Adriatico*

More recently, in *Manfredi*,[24] the ECJ proceeded to deal further with the "constitutive" and "executive" conditions of the Community right to damages. This was

21. *Op. cit*. at para. 26.

22. *Op. cit.* at para. 27.

23. Private antitrust actions, apart from their compensatory function, further the overall deterrent effect of the law. Thus, economic agents themselves become instrumental in implementing the regulatory policy on competition and the general level of compliance with the law is raised. It is for that reason that the private litigant in U.S. antitrust has been called a "private attorney-general" (per J. Jerome Franck *in* Associated Industries of New York State, Inc. v. Ickes, 134 F.2d 694, 704 (2d Cir. 1943)).

24. Joined Cases C-295/04 to C-298/04, Vincenzo Manfredi et al. v. Lloyd Adriatico Assicurazioni SpA et al., 2006 E.C.R. I-6619 [hereinafter *Manfredi*].

a preliminary reference case from Italy, where insurance companies had been sued for damages by Italian consumers for prohibited cartel behavior previously condemned by the Italian competition authority. The ECJ was basically called to decide:

- whether consumers enjoy a right to sue cartel members and claim damages for the harm suffered when there is a causal relationship between the agreement or concerted practice and the harm;
- whether the starting time of the limitation period for bringing an action for damages is the day on which the agreement or concerted practice was put in effect or the day when it came to an end; and
- whether a national court should also, of its own motion, award punitive damages to the injured third party, in order to make the compensable amount higher than the advantage gained by the infringing party and discourage the adoption of agreements or concerted practices prohibited under Art. 81 EC.

The Court, building on *Courage*, and after making it clear that the basis for individual civil liabilities deriving from a violation of Article 81 EC indeed lies in Community law, seems to have followed former AG Van Gerven's scheme of "constitutive", "executive", and simple "procedural" conditions of the Community right to damages. Thus, the Court makes a fundamental distinction between the "existence" and "exercise" of the right to damages. That the "existence" of the right is a matter of Community law is obvious from the fact that the Court reiterated the most important pronouncements of *Courage*.[25] In this context, it is also clear that the Court proceeded to define, as a matter of Community law, what former AG Van Gerven calls "constitutive" conditions of the right to damages: "It follows that any individual can claim compensation for the harm suffered where there is a causal relationship between that harm and an agreement or practice prohibited under Article 81 EC."[26]

In other words, the right to damages is open to (a) "any individual" as long as there is (b) "harm", (c) a competition law violation, and (d) a "causal relation-

25. *Op. cit.* at paras. 60, 61, 63 & 89-91 (citing *Courage*, *supra* note 18, at paras. 25-27). In particular, para. 91 of *Manfredi* (quoting para. 27 of *Courage*), stresses that:

> the *existence* of such a right strengthens the working of the Community competition rules and discourages agreements or practices, frequently covert, which are liable to restrict or distort competition. From that point of view, actions for damages before the national courts can make a significant contribution to the maintenance of effective competition in the Community. (emphasis added)

26. *Op. cit.* at para. 61.

ship" between that harm and that violation.[27] In thus defining the Community law constitutive conditions of the right to damages, the Court has produced a broad rule of standing, which includes consumers and indirect purchasers, while at the same time omitting the requirement of fault, which may mean that national rules following more restrictive rules on standing or requiring intention or negligence for an action for damages to be successful are contrary to the constitutive conditions in Community law of the *Courage/Manfredi* right to damages.

To mark the distinction between the existence of the right and its constitutive conditions, governed by Community law, and its exercise and executive conditions, governed by national law, the Court stresses again that "any individual . . . can claim compensation for [harm causally related with an Article 81 EC violation]," but:

> "[I]n the absence of Community rules governing the matter, it is for the domestic legal system of each Member State to prescribe the detailed rules governing the *exercise* of that right, including those on the application of the concept of "causal relationship", provided that the principles of equivalence and effectiveness are observed."[28] (emphasis added)

We submit that the Court refers here to the "executive" rules of the Community right to damages. In Van Gerven's scheme, these are separate from purely procedural rules, which are again a matter for national law. They are also subject to a higher standard of control under an "adequacy test", rather than a mere "minimum effectiveness" or "non-impossibility" test, which may continue to apply for simple procedural rules.

Indeed, the Court in *Manfredi* makes a clear distinction in its analysis between specific questions pertaining to the causal relationship between harm and

27. *Compare also* the recent ruling in *City Motors* which again refers to the constitutive conditions of the right to damages in the motor vehicle distribution context. Case C-421/05, City Motors Groep NV v. Citroën Belux NV, 2007 E.C.R. I-653, at para. 33:

> In the event of a breach by a supplier of the condition for application of the block exemption set out in Article 3(4) of Regulation No 1400/2002, the national court must be in a position to draw all the necessary inferences, in accordance with national law, concerning both the validity of the agreement at issue with regard to Article 81 EC and *compensation for any harm suffered by the distributor where there is a causal relationship between that harm and an agreement or practice prohibited under Article 81 EC*. (emphasis added)

28. *Manfredi*, *supra* note 24, at paras. 63-64.

antitrust violation and the availability of punitive damages, both seen as "executive" conditions,[29] and questions on limitation of actions and competent national tribunals, both seen as "detailed procedural rules". In addition, the Court seems to share the former AG's conviction that executive conditions affect the very core of the exercise of Community-based rights and should therefore be subject to a more stringent test concerning the Community principle of effectiveness, while detailed procedural rules can be subject to a more relaxed "non-impossibility" test. It is thus no surprise that in *Manfredi* the Court uses the "non-impossibility" language only in the context of mere procedural rules and not in the context of the executive conditions.[30] This means that questions such as causality, nature of harm and damages, and defenses, which can be characterized as "executive" conditions, will be subject to a more demanding test of effectiveness or adequacy, while questions such as competence of courts, limitation periods, and rules on proof, which are more "procedural" in nature, will be subject to a minimum effectiveness/non-practical impossibility test.

C. THE COMMISSION'S 2005 GREEN PAPER

The ECJ's *Courage* ruling provided the impetus for the Commission to adopt a more pro-active stance on private enforcement. Modernization was now a reality and there were, maybe for the first time, serious debates in Europe as to the desirability of introducing measures to enhance private antitrust enforcement. Soon after the ECJ delivered its *Courage* ruling, the Commission commissioned a study ("Ashurst Study") on the conditions for claims for damages in the member states in the case of infringement of EC competition rules.[31] Predictably, the study showed an "astonishing diversity and total underdevelopment" of civil antitrust actions in the member states. Until mid 2004, there were approximately 50 judgments that were the result of damages actions. Of these judgments, only 28 had resulted in a damages award.[32]

After digesting the results of the Ashurst Study and reflecting further on the appropriate way to move forward, on December 19, 2005, the Commission pub-

29. *Op. cit.* at paras. 64 & 92 *et seq.* as to causal relationship and punitive damages, respectively.

30. *Compare* paras. 64 & 92, which refer merely to effectiveness, with paras. 71 & 78, which refer to effectiveness seen through the prism of "rendering practically impossible or excessively difficult the exercise of rights conferred by Community law."

31. The study is made up of a comparative report on economic models for the calculation of damages and 25 national reports. *See* D. Waelbroeck et al., Study on the Conditions of Claims for Damages in Case of Infringement of EC Competition Rules: Comparative Report (Ashurst report) (2004), *available at* http://ec.europa.eu/comm/competition/antitrust/others/actions_for_damages/study.html.

32. These statistics are only indicative, since in some member states not all judgments are published and the comparative report necessarily relies on the national reports, for which quality varies. One must also bear in mind that these statistics do not include cases that were settled with significant damages awarded to the plaintiffs.

lished a Green Paper on damages actions for breach of the EC antitrust rules for public consultation.[33] The purpose of the Green Paper, which set out a number of possible options to facilitate private damages actions, was to stimulate debate and facilitate feedback from stakeholders. The Commission was in favor of increased private enforcement, as it believed that this will have a number of advantages for private parties, in particular:

(a) victims of illegal anticompetitive behavior will be compensated for loss suffered;

(b) deterrence of antitrust infringements and compliance with the law will be increased;

(c) a competition culture among market participants, including consumers, will develop further, and awareness of the competition rules will be raised; and

(d) as the Commission and the national competition authorities do not have sufficient resources to deal with all cases of anticompetitive behavior, administrative authorities will have discretion to pursue other priorities.[34]

D. NATIONAL DEVELOPMENTS

It would be inappropriate to end this review of the way to the WP without referring to developments in the member states. Modernization and decentralization of Community competition law enforcement and the related European debate on private enforcement, as well as the 2001 *Courage* ruling by the ECJ, led to important initiatives at the national level. The United Kingdom and Germany completely overhauled their legislation and, among other reforms, introduced provisions aimed at enhancing private antitrust enforcement of national and Community competition law. At the same time, there has been a surge of damages actions and awards in the national courts, most of them being cases of follow-on claims (i.e., actions relying usually on prior decisions by competition authorities). Whether this last development indicates increased awareness by plaintiffs or changing judicial attitudes is still unclear, but it certainly confirms that the European "wake-up calls" are reaching the member states.

To start with the United Kingdom, the Competition Act 1998 did not contain any direct reference to civil actions or actions for damages, though the availabil-

33. Commission Green Paper on Damages Actions for Breach of the EC Antitrust Rules [hereinafter "Green Paper"], COM(2005) 672 final (Dec. 19, 2005). The Green Paper was accompanied by a Staff Working Paper which set out the various options more discursively. *See* Commission Staff Working Paper, Annex to the Green Paper on Damages Actions for Breach of the EC Antitrust Rules, SEC(2005) 1732 (Dec. 19, 2005).

34. *See* Press Release, European Commission, MEMO/05/489, European Commission Green Paper on Damages for Breach of EC Treaty Antitrust Rules - Frequently Asked Questions (Dec. 20, 2005).

ity of such actions was implicit in some other provisions of the Act.[35] The situation was about to change. In 2001, a consultation paper by the U.K. Department of Trade and Industry powerfully advocated the desirability of private damages actions. Such actions were seen as serving two basic aims: (i) compensating victims of anticompetitive practices; and (ii) drawing private resources into the enforcement process, thus allowing public authorities to pursue the most important cases.

These ideas were set in motion with the Enterprise Act of 2002. Of particular interest for private enforcement is the conferring on the U.K. Competition Appeal Tribunal ("CAT") of jurisdiction to hear claims for damages in competition cases.[36] Damages claims before the CAT presuppose the establishing by either the U.K. Office of Fair Trading ("OFT") or the European Commission that an infringement of competition law has occurred. Such a finding of infringement is binding and cannot be re-litigated. These actions must be filed with the CAT within a period of two years, beginning at the time of the public enforcer's final infringement decision or on the date when the cause of action accrued. In addition, U.K. law provides for the possibility for ordinary civil courts to transfer competition issues arising in private civil actions to the CAT.

Section 58A of the U.K. Competition Act also aims at facilitating follow-on civil actions for damages brought before the ordinary civil courts. It provides that findings of infringement of U.K. or Community competition law by the OFT (or by the CAT on appeal) bind the courts deciding on follow-on civil claims for damages.[37] Apart from section 47A on follow-on civil claims for damages, the U.K. system provides for another novelty: section 47B provides for claims for damages brought on behalf of consumers by representative "specified" bodies. These are not meant to be U.S.-style class actions, and the claim must specify the consumers on whose behalf the claim is being brought.

The latest amendment of the German Competition Act offers another paradigm. German law has long-provided for antitrust damages actions, but the new section 33 of the GWB marked important progress, in that it provided a legal basis for damages claims for violation not only of German, but also of Community, competition law. The new provision also abandoned the previous rather restrictive condition for standing, which was conferred only on persons

35. *See*, *e.g.*, U.K. Competition Act 1998, at §§ 55(3)(b) & 58(2).

36. *Id.* at § 47A.

37. It is noteworthy that this provision is different from section 47A of the Act. The former refers only to the U.K. competition authorities' decisions, while the latter extends the binding effect of infringement findings to decisions of the European Commission. In addition, the provision of section 58A refers to follow-on civil proceedings for damages before the ordinary civil courts (the Chancery Division of the High Court), while section 47A refers to follow-on claims brought before or transferred to the CAT.

within the "protective scope" of the statute, and stressed that any "person affected", including competitors and "other market participants" can sue for damages. The law also gave standing to associations for the promotion of commercial or independent professional interests, including consumer associations. One novel feature was that it is now possible for the courts to calculate the damages taking into account the proportion of the profit which the defendant derived from the infringement. In addition, the passing-on defense is restrained, though not completely banned. Finally, German law goes even further than U.K. law by conferring a binding effect, not only on European Commission and Bundeskartellamt, but also on all other EU member states' competition authorities' infringement decisions. This binding effect is confined to follow-on civil litigation, basically aiming at offering incentives to claim damages from convicted cartelists.

As far as national case law is concerned, there has also been a boom in recent years. It is interesting to note that many of the recent cases, some of which are still pending, are follow-on cases. The *Vitamins* case is the most prominent source of such actions and there are already damages awards and settlements in Germany, England, Sweden, and other jurisdictions. A famous example is the *Provimi* judgment decided at the admissibility stage by the English High Court, where, in addition to English parties, a German party was also claiming antitrust damages.[38] Judgment was given only as to the jurisdictional issues and subsequently the parties settled. There, it was established that where there is an English connecting factor in the private international law sense (i.e., an English element to a cartel), other non-English claimants may also bring claims in London for their non-English based losses, instead of having to pursue separate claims in other jurisdictions.

The *Vitamins* litigation provided also for the first claims that were brought before the CAT as follow-on civil claims for damages under the special procedure of section 47A of the U.K. Competition Act. These cases did not lead to final judgments as they were settled, although there has recently been a resurgence.[39] However, another recent follow-on action brought before the CAT, and which was also settled, gave rise to an interim award of damages, which was the first ever award of damages in the United Kingdom for a competition law infringement. This was based on previous infringement decisions by the OFT and CAT in an abuse of dominance case concerning margin squeeze and rebates in the pharmaceutical sector. On November 16, 2006, while the case was still pending,

38. Provimi Ltd. v. Aventis Animal Nutrition SA et al. QB, 2003 E.C.C. 353 (Com. Ct.).

39. *See* Case No. 1029/5/7/04, Deans Foods Limited v. Roche Products Limited, F. Hoffman-La Roche AG and Aventis SA (2004) (case settled); and Case No. 1028/5/7/04, BCL Old Co. Ltd. DFL Old Co. Ltd. and PFF Old Co. Ltd. v. Aventis SA, Rhodia Ltd., F. Hoffman-La Roche AG and Roche Products Ltd. (2004) (case settled). But *see also*, very recently, Case No. 1098/5/7/08, BCL Old Co Ltd. et al. v. BASF AG, BASF plc and Frank Wright Ltd. (2008) (case pending); and Case No. 1101/5/7/08, Grampian Country Food Group Ltd et al. v. Sanofi-Aventis SA et al. (2008) (case pending).

the CAT awarded "interim damages" for an amount of GBP 2 million.[40] That represented, in the court's view, roughly 70 percent of the likely final damages award. Meanwhile, more actions for damages have been brought before the CAT, but again for the most part these were settled.[41]

In Germany, the courts' initial rejectionist approach has now changed and the first successful follow-on damages claims in the *Vitamins* litigation have become a reality. It is noteworthy that certain German courts adjudicating claims for damages in the post-*Courage* era refused to grant damages to direct purchasers of vitamins on passing-on grounds, because the cartel was not specifically directed at them but at all market participants.[42] This built on a very restrictive reading of standing under German law that was certainly incompatible with Community law, in particular the *Courage* ruling, which accepted no such limitations but granted a right to damages to all individuals harmed by the anticompetitive conduct. Recent German judgments, however, have reversed this restrictive approach and have rendered the first damages awards.[43]

Important successful damages claims have also been reported in Austria, France, Denmark,[44] Spain,[45] Sweden, and Italy, where the Corte di Cassazione after lengthy tribulations established that consumers could claim damages from

---

40. Case No. 1060/5/7/06, Healthcare at Home Ltd. v. Genzyme Ltd., 2006 C.A.T. 29 (case settled). Under English law, interim damages can be an adequate provisional measure, if there is a very good prima facie case and if damages appear to be an appropriate final remedy.

41. *See*, *e.g.*, Case No. 1078/7/9/07, The Consumers Association v. JJB Sports plc (2007) (case settled); Case No. 1088/5/7/07, ME Burgess et al. v. W. Austin & Sons (Stevenage) Ltd. and Harwood Park Crematorium Ltd. (2007) (case pending); and Case No. 1077/5/7/07, Emerson Electric Co et al. v. Morgan Crucible Company plc et al. (2007). In the latter case, the CAT decided, on a preliminary point, that the time for making a claim for damages pursuant to section 47A of the Competition Act 1998 had not yet begun to run, since appeals were pending before the Community Courts against the Commission's decision which the plaintiffs sought to rely on in their follow-on claims (*see* Judgment of 17 October 2007, Emerson Electric Co et al. v. Morgan Crucible Company plc et al., 2007 C.A.T. 28). Eventually, the CAT granted permission to the plaintiffs to proceed with their damages claims against the immunity recipient, which had naturally not challenged to the CFI the European Commission's—in that case—cartel infringement decision (Judgment of 16 November 2007, Emerson Electric Co et al. v. Morgan Crucible Company plc et al., 2007 C.A.T. 30), but not against the other addressees of the infringement decisions, whose appeals to the CFI were still pending (Judgment of 28 April 2008, Emerson Electric Co et al. v. Morgan Crucible Company plc et al., 2008 C.A.T. 8).

42. Judgment of 11 July 2003, 7 O 326/02 – Vitaminkartell (LG Mannheim, 2004); and Judgment of 15 January 2004, 12 HK O 56/02 Kart - Vitaminpreise, 54 WuW 1179 (LG Mainz, 2004).

43. Judgment of 4 January 2004, 13 O 55/02 Kart – Vitaminpreise, 54 WuW 1182 (LG Dortmund, 2004). The damages awarded in this case amounted to the difference between the price paid as a result of the cartel and a hypothetical market price. In addition, the court ruled that the defendant had failed to prove that the plaintiff had passed on his damage to his customers.

44. Judgment of 20 April 2005, GT Linien A/S v. DSB and Scandlines A/S, UfR 2005.2171 H (2005). The award of damages amounted to DKK 10 million plus interest and was mostly upheld by the Danish Supreme Court. *See also* Judgment of 3 October 2002, EKKO v. Brandt Group Norden et al. (EKKO I),

a cartel of insurance companies previously convicted by the Italian competition authority.[46]

III. The White Paper

A. INTELLECTUAL DEBT ACKNOWLEDGED

It is only in the aftermath of the ECJ's important pronouncements that the European Commission decided to go forward with the publication of its Green and White Papers. In so doing, the Commission availed itself of an increased degree of legitimacy in an area, which is always sensitive due to the inevitable intrusion on what is perceived as the member states' institutional and procedural autonomy. It would have been very difficult for the Commission to go ahead and propose Community legislative action, if it had not been for the ECJ's seminal rulings. This intellectual debt is fully acknowledged in the WP, which gives much space to the acquis communautaire, as established by the Court.

Indeed, the WP starts from the premises that the right to be compensated for harm caused by an antitrust violation is a right guaranteed by the Treaty itself, as the ECJ has stressed in *Courage* and *Manfredi*. This statement is an important reminder because the idea that the right to damages is based in Community law is still resisted by some commentators, particularly in the German-speaking theory, which sees this purely as a matter of national law, subject only to the Community principles of equality and effectiveness.[47] The Commission is now

---

footnote 44 cont'd
UfR 2004.2600 S (Copenhagen Maritime and Commercial Court, 2004); and Judgment of 15 October 2004, EKKO v. Electrolux Home Products Denmark et al. (EKKO II), UfR 2005.388 S (2005).

45. Case No. 125/2005, Antena 3 TV v. LNFP (Juzgado de Primera Instancia No. 4 de Madrid, Jun. 2005). In this case, in a "follow-on" civil action, the court awarded a record EUR 25 million in damages because LNFP, the Spanish football league, had abused its dominant position by selling broadcasting rights on an exclusive basis to regional public broadcasters, thus foreclosing certain new entrants. The damages claim was based on an earlier national infringement decision, confirmed by the Spanish competent courts. *See also* Case No. 36/2005, Conduit Europe SA v. Telefónica de España SAU (Juzgado de lo Mercantil No. 5 de Madrid, Nov. 11, 2005), confirmed in Audiencia Provincial de Madrid (Secc. No. 28) (May 25, 2006); and Case No. 73/2006, Conduit Europe SA v. Telefónica de España SAU (May 25, 2006). In that case, the court awarded the plaintiff, an Irish communication services provider, EUR 639,000 for losses incurred as a result of the defendant's abuse of a dominant position, consisting of giving defective and incomplete information in order to block the plaintiff's entry into the market for subscriber directory inquiries.

46. Case No. 2207, Compagnia Assicuratrice Unipol SpA v. Ricciarelli, 11 DANNO E RESPONSABILITÀ 495 (Corte di Cassazione, Feb. 4, 2005).

47. In favor of the Community law basis of the right to damages, *see also* K. Lenaerts & K. Gutman, *'Federal Common Law' in the European Union: A Comparative Perspective from the United States*, 54 AM. J. COMP. L. 1, 91-94 (2006); and Opinion of AG Poiares Maduro, Case C-438/05, The International Transport Workers' Federation and the Finnish Seamen's Union v. Viking Line ABP and OU Viking Line Eesti, 2007 E.C.R. I-10779, at para. 53 (which, referring to *Courage*, clearly speaks of a claim "based directly" on Community law).

unequivocal: there are many references to "the establishment under Community law of a right to compensation", derived "directly from Community law", and to the fact that "this European law remedy can as such not be refuted or conditioned by national legislation of any kind."[48] There is also a clear distinction between the existence of the right, which is a matter of primary Community law, and its exercise, which is determined by national legislation and which the WP intends on harmonizing to a certain extent through secondary Community law.[49]

A fundamental quality of the WP is that it codifies and restates the existing acquis communautaire involving most aspects of the right to damages for EC competition law violations. Naturally, references to the ECJ *Courage* and *Manfredi* rulings are prominent, but there are also references to other case law that deals with many other questions of remedies and procedures available at the national level for the enforcement of Community law. The Commission's choice to dedicate whole sections of the WP to the presentation of the impressive acquis communautaire is a wise one. First, it shows that even if the whole initiative to introduce Community measures for private actions were abandoned, the existing acquis itself is a Community minimum from which there can be no departure. Second, it acts as a powerful support and starting point for Community legislation.[50]

Notwithstanding this acquis communautaire and the Community law basis of the right to damages, the WP recognizes that there are various national legal and procedural hurdles and that therefore there is a need for a strong set of legislative measures to enhance private actions for damages. Community measures (e.g., a regulation or directive), and most likely a Commission Notice on the quantification of damages, are seen as desirable in order to achieve (i) an effective minimum protection of victims; (ii) a level-playing field; and (iii) greater legal certainty.

B. THE MAIN POLICY OPTIONS OF THE WHITE PAPER

The Commission speaks in the WP of a "combination of measures at Community and national level." It is fair to say that the Commission had never pretended to have an exclusive role in this area, but its more deferential attitude to national competencies may be the result of some resistance at the national level with respect to Community unification and harmonization initiatives especially those touching upon matters of national procedural laws. Indeed, the WP is now proposing to leave to the national level rules on costs, court fees, and funding.[51] In

48. SWP, *supra* note 1, at paras. 308-09 (emphasis added). *See also* WP, at § 1.1.

49. SWP, *supra* note 1, at para. 309.

50. Indeed, an argument that is often heard in favor of Community legislation in this area is that if Community legislation does not step in to deal with the conditions for the exercise of the right to damages (positive integration), then the ECJ would have to do this through the preliminary reference procedure (negative integration).

51. *See* SWP, *supra* note 1, at ch. 9.

addition, and to the extent Community legislation is necessary, "procedural" matters, such as collective relief and access to evidence, are seen as candidates for harmonization (through a directive), rather than for unification (through a Regulation). It is the view of this author, however, that a Regulation would be more appropriate for core conditions of the exercise of the right to damages, like standing, passing-on, and questions pertaining to fault.

The two basic objectives of damages actions, as perceived by the WP, are (a) full compensation for victims, which is presented as "primary objective", and (b) effectiveness of competition enforcement in Europe through increased deterrence, which presumably must be the secondary objective.[52] The Commission also mentions as a third objective the development of a competition culture among market participants and the increased awareness of the competition rules.[53]

The main measures and policy choices that the Commission intends to pursue can be summarized as follows:

- standing to sue for damages should be recognized for all persons harmed by an EC competition law violation, including competitors, direct and indirect purchasers, and of course consumers;
- direct purchasers in particular should be able to rely on the rebuttable presumption that the illegal overcharge was passed on to them in its entirety ("offensive passing-on");
- at the same time, it should be open to defendants to prove that the plaintiff (e.g., a direct purchaser) has passed the illegal overcharge on to his customers; in other words, defensive passing-on should be permitted;
- collective redress should be possible through (i) representative actions by consumer associations, state bodies or trade associations, that are officially certified in the member states, and (ii) opt-in collective claims for consumers and businesses;
- plaintiffs' access to evidence held by defendants should be made easier; thus, the WP proposes in effect a certain relaxation of the "fact-pleading" system and the introduction of some elements of "notice-pleading" under the control of the judge whereby national courts should have the power to order the litigants or third parties to disclose specific categories of relevant evidence;
- final infringement decisions issued by the Commission and by national competition authorities ("NCAs") or final judgments on judicial

52. This hierarchy may also explain the absence of proposals for more "offensive" or "aggressive" mechanisms, such as punitive damages. On the two objectives, *see recently* P. Nebbia, *Damages Actions for the Infringement of EC Competition Law: Compensation or Deterrence?*, 33 Eur. L. Rev. 23 (2008).

53. SWP, *supra* note 1, at paras. 14-15.

review should be binding on national courts throughout the European Union in follow-on civil actions;

- objective (strict) liability should be the rule (no fault requirement) for damages, once the infringement has been established, unless the infringer demonstrates that there is a genuinely excusable error (i.e., he bears the burden of such proof);
- full compensation should be available, covering not just actual losses, but also lost profits and interest;
- there should be no Community measure on punitive damages;
- the limitation period should not start to run before the day a continuous or repeated infringement ceases, or before the victim can reasonably be expected to have knowledge of the infringement and of the resulting harm;
- for follow-on claims, there should be a new limitation period running for at least two years after an infringement decision has become final;
- corporate statements by leniency applicants (including unsuccessful ones) should not be discoverable, even after the adoption of a final decision; and
- the immunity recipient's civil liability should be limited to claims by his direct and indirect contractual partners.

C. THE IMPACT OF THE PRE-EXISTING ACQUIS COMMUNAUTAIRE ON THE WHITE PAPER'S TREATMENT OF THE CONDITION OF STANDING

The fact that Community law in the post-*Courage/Manfredi* era itself defines the constitutive conditions of the right to damages, has profound consequences for very important questions such as the rules on standing, in particular for indirect purchasers and consumers.

Perhaps the most important feature of the WP is the broad rule of standing it advocates, notably for indirect purchasers and of course consumers. At the same time, the WP proposes the retention of the "passing-on defense". The question of the standing of indirect purchasers is closely connected with the prohibition or permission of the passing-on defense. Indeed, standing of indirect purchasers is referred to at times as "offensive passing-on".

Under U.S. antitrust law, indirect purchasers (e.g., traders that have purchased from retailers rather than from the manufacturer), cannot recover damages,[54] notwithstanding the fact that the harm may have been passed on to them. U.S. law clearly favors compensation only of direct purchasers, and indeed, it disal-

54. State antitrust laws in the United States, however, may allow for indirect purchaser suits.

lows the passing-on defense in this case.[55] There is a powerful policy rationale behind this rule, that the direct purchaser is a more "efficient" plaintiff, ultimately preferable to indirect purchasers as a "private attorney general". In general terms, the U.S. rule also seems to take into account certain prudential considerations, such as the burden on the judicial system that would result if the private right of action were available in an unlimited way to remotely injured plaintiffs. Then, denying indirect purchasers standing is a direct consequence of the exclusion of the passing-on defense, since the defendant and perpetrator of the antitrust violation should not be vulnerable to multiple actions referring to the same acts, while at the same time it is not open to him to rely on the fact that the damage may have been passed-on. In other words, the U.S. system bans the defensive use of the passing-on principle by defendants, while at the same time banning its offensive use by indirect purchasers who base their claims precisely on the fact that the overcharge was passed on to them.

THE FACT THAT COMMUNITY LAW IN THE POST-*COURAGE*/*MANFREDI* ERA ITSELF DEFINES THE CONSTITUTIVE CONDITIONS OF THE RIGHT TO DAMAGES, HAS PROFOUND CONSEQUENCES FOR VERY IMPORTANT QUESTIONS SUCH AS THE RULES ON STANDING, IN PARTICULAR FOR INDIRECT PURCHASERS AND CONSUMERS.

Irrespective of the critique that can be made against this rather inflexible U.S. judge-made rule,[56] in the European context of damages claims, the constitutional status of the Treaty competition provisions and the fact that they form the basis of rights for individuals, mean that the U.S. theories could not have been adopted uncritically. Thus, the a priori exclusion of indirect purchasers from the ambit of the persons who can claim damages would not have been compatible with Community law.

In *Courage*, the Court had no difficulty in finding that Article 81 EC not only protected third-party competitors (in that case third-party beer suppliers foreclosed by a specific network of exclusive beer supply agreements), but also pro-

55. *See* Hanover Shoe v. United Shoe Machines Corp., 392 U.S. 481 (1968) and Illinois Brick Co. v. Illinois, 431 U.S. 720 (1977).

56. The problem with the total exclusion of indirect purchasers is that there may be times when the direct purchasers benefited from the infringement and are not at all inclined to sue. An upstream cartel may also shield itself from private damages claims by forwarding a share of cartel profits to its direct purchasers. These benefits dissuade the direct purchasers from exercising their exclusive right to sue for private damages. *See further* M.P. SCHINKEL, J. TUINSTRA & J. RÜGGEBERG, ILLINOIS WALLS: HOW BARRING INDIRECT PURCHASER SUITS FACILITATES COLLUSION (Amsterdam Center for Law & Economics, Working Paper No. 2005-02, April 2008), *in* RAND J. ECON. (forthcoming). It is noteworthy that the recent report of the U.S. Antitrust Modernization Commission proposes overruling *Illinois Brick* and *Hanover Shoe* to the extent necessary to allow both direct and indirect purchasers to sue to recover for actual damages from violations of federal antitrust law (*see* Antitrust Modernization Commission, Report and Recommendations (April 2007)).

tected "any individual",[57] including co-contracting parties, or in that case, tenants. *Manfredi* built on *Courage* and defined in detail the Community law constitutive condition of standing, explicitly recognizing that consumers enjoy standing to sue for harm caused to them by anticompetitive conduct.[58] Indeed, the WP follows that approach and adopts a broad rule of standing, covering also indirect purchasers. It is actually interesting that the WP refers to indirect purchasers' standing, not as a proposal, but rather as part of the already-existing acquis communautaire.[59]

Thus, in Europe, the solution will be the opposite from the U.S. solution: both direct and indirect purchasers will have standing to sue, but at the same time the passing-on defense will be available. Allowing the passing-on defense is a logical consequence of the broad rule of standing, otherwise, as the WP accepts, there would be a risk of unjust enrichment of those purchasers that passed on the illegal overcharge to their customers and of multiple compensation of the overcharge.[60]

Finally, since difficulties also arise when the indirect purchaser invokes the passing-on of the illegal overcharge as a basis of his claim ("offensive passing-on"), the WP proposes the introduction of a rebuttable presumption that the overcharge has indeed been fully passed on to the plaintiff-indirect purchaser. This is intended as an alleviation of the victim's burden of proof, without, however, affecting the main conditions of civil liability: in other words, the plaintiff would still have to prove the infringement, the existence of the initial overcharge, and the extent the overcharge caused him harm (including causation).[61]

Where the WP may give rise to a discussion as to its compatibility with the *Courage/Manfredi* case law is in its proposal to limit the civil liability of successful immunity recipients[62] to claims by their "direct and indirect contractual partners". The aim is basically to safeguard the effectiveness of the Leniency Program, which

57. *Courage*, *supra* note 18, at para. 26.

58. *Manfredi*, *supra* note 24, at paras. 60, 61 & 63. Compare AG Mischo's Opinion in *Courage*, *supra* note 18, at para. 38, stressing that "the individuals who can benefit from such protection are, of course, primarily third parties, that is to say *consumers* and competitors who are adversely affected by a prohibited agreement." (emphasis added)

59. SWP, *supra* note 1, at paras. 33-37. Of course, the broad rule of standing does not affect the necessity of a causal link between the harm and the infringement of Articles 81 and 82 EC (*op. cit.* at paras. 37 & 205).

60. *Op. cit.* at note 3, para. 210. At the same time, the WP stresses that the standard of proof for the passing-on defense should not be lower than the claimant's standard to prove the damage. Under this model, the plaintiff must prove that he has suffered loss, but it is open to the defendant to prove that the plaintiff mitigated the loss by passing on the whole or part of the overcharge to downstream purchasers.

61. *Op. cit.* at para. 220.

62. This proposal does not cover the other leniency applicants that did not receive full immunity.

might have been put at risk as a result of the Commission's drive for an enhanced system of private actions in Europe. According to that proposal, the immunity recipient would be liable only to persons that bought the products or services in question directly from the recipient (direct contractual partners) or those down the supply chain who bought these products or services from the direct contractual partners themselves.[63] Thus, a victim that did not buy cartelized products or services directly or indirectly from him,[64] a harmed competitor, a shareholder,[65] or a victim of so-called "umbrella pricing"[66] would not be able to claim damages. At the same time, this rule would in effect remove the immunity recipient's joint liability,[67] since, as the Commission explains in an example, "where 30% of a victim's total purchases of cartelized products originate from the immunity recipient, the latter would only be liable for 30% of the total harm suffered by this victim due to the overcharge of the cartelized products."[68]

The question here is whether the limitation of the right of competitors and others not falling under the Commission's definition of "direct and indirect contractual partners" is at odds with primary Community law.

The question here is whether the limitation of the right of competitors and others not falling under the Commission's definition of "direct and indirect contractual partners" is at odds with primary Community law (i.e., with the Treaty itself and the ECJ rulings in *Courage* and *Manfredi*), which stress that the right to damages should be open to "any individual". However, the fact that primary Community law itself provides for a broad rule of standing does not mean that the Community legislator cannot make a policy decision and restrict—though not eliminate—the right of some

63. *Op. cit.* at para. 305.

64. An issue is what happens with cartels that do not involve the sale of goods or services to contractual partners (e.g., a cartel not to sell in a particular market or to a particular client).

65. Whether shareholders or other persons related to a company that has breached the antitrust rules, such as employees, can sue for damages, is debated. In the author's view, the broad language in *Courage* should cover these persons too, assuming they can identify and prove harm and, more importantly, causation.

66. These are customers who purchased not from cartel members but from fringe firms outside the cartel, but within the same relevant market and that charge a higher price as a non-cooperative response to the cartel price. *See further* Impact Assessment Study, *supra* note 1, at 413.

67. The WP considers that removal of joint liability by itself is not sufficient to effectively limit the immunity recipient's liability (SWP, *supra* note 1, at para. 304). *Compare*, however, SWP, at para. 322, where removal of joint liability is surprisingly mentioned as a separate proposed measure. Perhaps the reference in para. 322 was left in from a previous draft by mistake.

68. SWP, *supra* note 1, at note 160.

plaintiffs, if that would be beneficial to the effectiveness of the whole system of enforcement.

A closer look at the proposed solution reveals that in reality the WP does not propose to affect the exercise of those persons' right to damages against the other cartel members that did not receive full immunity from fines. Indeed, joint and several liability of these cartel members continues to be the rule, so they would still be jointly and severally liable to pay damages to a potential harmed competitor for the whole of his harm. Thus, in reality, what the WP proposes is not to totally bar some persons from suing for damages, but rather to make those persons only slightly worse off by slightly increasing their risk in case of the insolvency of all or some of the other cartel members with the exception of the immunity recipient. This is a rather low risk.[69] In fact, irrespective of this WP proposal and of what primary Community law dictates, all plaintiffs always bear the risk of all the cartel members' insolvency. So, it seems that the proposed solution would most likely not seriously affect the existence of the Community right to damages, while at the same time it would undoubtedly strengthen the effectiveness of one aspect of the Leniency Program, the race to the authority to be the first undertaking that self-reports, thus ensuring full immunity status.[70] Being second or third would not only mean the loss of full immunity, but also exposure to damages liability for the whole of the harm.[71]

An even better solution would be to completely exclude the immunity recipient's liability also for claims by his direct and indirect contractual partners. Again, this would not dramatically affect the exercise of the right to damages by these persons, since they could still claim compensation for the whole of their damage against the other cartel members, who would remain jointly and severally liable. As a safety valve, the law could provide that this total exclusion of liability would not apply to the exceptional case of insolvency by one or more of the jointly and severally liable (other) cartel members.[72] While not affecting the right of compensation, such a solution would enhance the effectiveness of the Leniency Program

69. The cartels that are prosecuted by the Commission under Article 81 EC are likely to concern activity and companies of a certain size and therefore the risk of insolvency of any of these companies is extremely low.

70. *See also* Impact Assessment Study, *supra* note 1, at 521.

71. Note, however, that the Commission does not propose to disallow contribution among the (non-immunity recipient) cartel members.

72. In such a case, the plaintiffs would have to sue first the other cartel members and, in case of insolvency of the latter, they could then bring a new action against the immunity recipient for the part of harm that is attributable to him (in other words, removal of joint liability for him should here be the rule).

even more.[73] Indeed, this has now been adopted in the amended text of the Hungarian Competition Act, which, with the intention to increase the attractiveness of the leniency policy, provides that a leniency applicant receiving full immunity from fines would not be liable to pay damages to third parties, until and if such damages can be collected from other cartel members (i.e., from those which did not receive full immunity under the leniency policy).[74]

Besides, ensuring that the Leniency Program remains attractive and thus effective is very beneficial for private enforcement and potential plaintiffs. First, the plaintiffs become aware of the cartel infringement, which is more effectively exposed to the public authority by the leniency applicants. Second, the facts are established during the administrative proceedings. Third, courts or plaintiffs could under certain circumstances ask for documentary evidence in the hands of the public enforcer, in order to establish the liability or damage. Fourth, a final public decision, depending on the applicable rules, may have a binding effect on the follow-on civil proceeding or may constitute prima facie evidence of the cartel violation.[75]

73. Of course, a debate is still possible, if one views *Courage* and *Manfredi* not only as authority for a Community right to damages available to victims, but also as authority for a Community law obligation imposed on infringers (to compensate the victims). In that case, indeed, any exclusion of an infringer's liability to certain classes of victims would be contrary to primary Community law. In the author's view, however, the language in *Courage* and *Manfredi* (*supra* notes 18 and 24, at paras. 26-27 and paras. 89-91, respectively), which is rights- and not obligations-centered, and the underlined powerful rationale of effectiveness would allow for a compromise in order to safeguard the effectiveness of public enforcement and thus by implication the effet utile of Article 81 EC.

74. The amended text has not yet come into force due to a pending review by the Constitutional Court, because of certain concern on new rules imposing liability on management. Although the amendments were due to take effect in September, the entire amending act is now suspended until the Constitutional Court concludes its review. In any case, the new provisions on damages actions are not subject to the ongoing constitutional review, so they are likely to enter into force as they stand now.

75. *See also* A.P. Komninos, *The EU White Paper for Damages Actions: A First Appraisal*, 84 CONCURRENCES 2, 89-90 (2008).

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Private Actions in EC Competition Law

Renato Nazzini & Ali Nikpay

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853),
Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Private Actions in EC Competition Law

*Renato Nazzini and Ali Nikpay**

The paper considers the case for reform of the system of private actions in the European Union. In doing so, it seeks to identify the central changes which would need to be made if private actions are to play a more significant role in the competition regime. Contrary to recent statements made by the European Commission, the paper argues that any changes made must recognize that private actions perform a dual function in EC competition law: they not only compensate those who have been harmed by anticompetitive behavior but also contribute to the overall level of deterrence generated by the competition regime. Going further, it argues that whilst increased deterrence and compensation almost always go hand in hand, the primary objective of private actions is to support effective competition enforcement.

Building on this, the paper identifies and examines the main pillars of any effective reform program in Europe: enhancing the role of collective actions, clarifying the issues surrounding indirect purchasers' standing and passing-on, and ensuring, as far as possible, that public and private enforcement operate in harmony—where they clash, the paper argues that the former must take precedence over the latter.

In light of this discussion, the paper goes on to assess the proposals made by the European Commission ("Commission") and the U.K. Office of Fair Trading ("OFT") for reform of the system. It concludes that the proposals made, if implemented, would appreciably increase the incentives of businesses to comply with

*Renato Nazzini is a Solicitor, Member of the Chartered Institute of Arbitrators, and Reader in law at the University of Southampton. Ali Nikpay is the Senior Director for Policy at the Office of Fair Trading in the United Kingdom and Visiting Fellow at the London School of Economics. The views expressed in this article are personal to the authors and do not necessarily reflect those of the Office of Fair Trading. The authors would like to thank Philip Collins and Steven Altham for their comments on an earlier draft.

the EC competition rules while at the same time achieving higher levels of compensation. In addition the reformed system would retain significant safeguards to guard against the risk of unmeritorious or speculative claims. However both sets of proposals are cautious in particular in relation to the availability of opt-out collective actions. This is an issue which policymakers in Europe may need to return to in the future.

I. Introduction

Over the last 10 years competition law enforcement by the Commission has been transformed from both a substantive and procedural perspective. Though not perfect, public enforcement is now far more effective than it was just a decade ago. The same cannot be said for private actions brought in the courts: a study prepared for the Commission in 2004 described the system as being in a state of "total underdevelopment".[1] A later study[2] for the Commission found that between 2004 and 2007 there were less than 100 antitrust damages actions across the EU; significantly, almost all were concentrated in a few sectors in 17 of the 27 Member states. The authors of the report estimated that "at most" 10 percent of antitrust cases in Europe are initiated by a private claim before a national court; this compares with 95 percent in the United States.[3] National competition authorities ("NCAs") have made similar findings. For example, research carried out by the OFT shows that companies and their advisers view private actions as the least effective aspect of the competition regime in the United Kingdom;[4] indeed the OFT reported in 2007[5] that consumers in the United Kingdom had never recovered damages for breach of the competition rules.

In response, the Commission and a number of European NCAs (in particular the OFT) have, over the last 3 years, sought to identify the main obstacles to a more effective system of damages claims and set out different options to improve the regime. These initiatives culminated in the publication by the OFT in

1. Ashurst, Study on the conditions of claims for damages in case of infringement of EC competition rules, Brussels (2004).

2. Making antitrust damages actions more effective in the EU, joint report submitted to the European Commission by the Centre for European Policy Studies, Erasmus University Rotterdam (EUR) and Luiss Guido Carli (December 2007) [hereinafter Joint Report].

3. H. J. Hovenkamp, The Antitrust Enterprise: Principle and Execution 38 (2006) [hereinafter Hovenkamp].

4. The deterrent effect of competition enforcement by the OFT, OFT962 and OFT963 (November 2007), *available at* http://www.oft.gov.uk/advice_and_resources/resource_base/consultations/deterrent.

5. Private actions in competition law: effective redress for consumers and business - Recommendations from the Office of Fair Trading, OFT916resp (November 2007) [hereinafter Recommendations] *available at* http://www.oft.gov.uk/shared_oft/reports/comp_policy/oft916resp.pdf.

November 2007 of Recommendations to Her Majesty's Government for reform of the U.K. system[6] and by the Commission in April 2008 of a White Paper.[7]

However, despite the "total underdevelopment" of the private actions system and the strenuous efforts made by many competition authorities, reform of the system remains a controversial subject. Challenges have been made both to the principle of a greater role for private actions in the European regime[8] and to many of the ideas for change put forward by the OFT and the Commission.

The aim of this paper is to consider the case for reforming the European system of private actions. In doing so, we will identify the central changes which would need to be made if private actions are to play a more significant role in the competition regime. The paper will also assess the proposals made by the Commission and the OFT.

II. The Case for Private Enforcement

In essence, competition *enforcement*[9] has two main functions:[10] first to ensure that the prohibitions in the law are not violated (the "deterrent" effect); and second to provide corrective justice through compensation to victims (the "compensation" effect). In the United States, private actions clearly perform both functions.[11]

However, according to the White Paper,[12] the Commission's primary objective in reforming the private actions regime in the EU is "to improve the legal conditions for victims to exercise their right under the Treaty to reparation of all

6. *Id*.

7. Commission White Paper on Damages actions for breach of the EC antitrust rules, Brussels, 2.4.2008 COM(2008) 165 final [hereinafter White Paper on Damages Actions], Commission Staff Working Paper accompanying the White Paper on Damages actions for breach of the EC antitrust rules, Brussels, 2.4.2008 SEC(2008) 404 [hereinafter Staff Working Paper on Damages Actions], and Commission Staff Working Document—Accompanying document to the White Paper on Damages actions for breach of the EC antitrust rules: Impact Assessment, Brussels, 2.4.2008 SEC(2008) 405 [hereinafter "Impact Assessment accompanying the White Paper on Damages actions"].

8. *See*, for example, W. P. J Wils, *Should private enforcement be encouraged in Europe?,* World Competition 478 (2003).

9. For the purposes of this paper we draw a distinction between the *objectives* of competition law (for example, consumer welfare, SME protection, etc.) and the function of enforcement (which is essentially about *how* the objectives are achieved).

10. *Se,* generally, W. P. J Wils, The Principles of European Antitrust Enforcement (2005) [hereinafter Wils].

11. *See,* generally, C. A. Jones, Private Enforcement of Antitrust Law (1999); see also Hovenkamp, *supra* note 3 at chapter 3.

12. White Paper on Damages Actions *supra* note 7 *at* 3.

damage suffered as a result of a breach of the EC antitrust rules." Increased deterrence is mentioned almost in passing and appears to be viewed as no more than a useful by-product: "[i]mproving compensatory justice would therefore inherently also produce beneficial effects in terms of deterrence."

The position taken in the White Paper on Damages actions runs contrary to the views set out by the Commission in its 2005 *Green Paper on Damages Actions for Breach of the EC Antitrust Rules*[13] where it expressly stated that public enforcement and private actions "are part of a common enforcement system and serve the same aims: to deter anticompetitive practices forbidden by antitrust law and to protect firms and consumers from these practices and any damages caused by them."[14] More importantly, the weight placed on compensation appears to run counter to two recent judgments from the European Court of Justice ("ECJ"). In these cases the ECJ explicitly underlined the dual function of private actions, emphasizing in particular their deterrent effect.[15] In its 2001 judgment *Courage and Crehan*[16] the Court held that the "full effectiveness" of Article 81 and, in particular, the practical effect of the prohibition laid down in Article 81(1) would be put at risk if it were not open to any individual to claim damages for losses caused to him by a contract or by conduct liable to restrict or distort competition. The Court went on to explain that the existence of such a right strengthens the working of the Community competition rules and "*discourages agreements or practices, frequently covert, which are liable to restrict or distort competition*. From that point of view, actions for damages before the national courts can make a significant contribution to the maintenance of effective competition in the Community" (emphasis added). Five years later in 2006, the ECJ repeated, almost verbatim, paragraphs 26 and 27 of *Courage and Crehan* in its judgment *Vincenzo Manfredi v Lloyd Adriatico Assicurazioni*.[17]

13. Commission Green Paper on Damages actions for breach of the EC antitrust rules, Brussels, 19.12.2005 COM(2005) 672 final [hereinafter Green Paper on Damages actions] and the Commission Staff Working Paper Annex to the Green Paper, Brussels, 19.12.2005, SEC(2005) 1732 [hereinafter Staff Working Paper Annex to the Green Paper].

14. *Id*.

15. Commentators are divided as to which of the deterrence and compensation functions would, as a matter of law, take precedence should there be a conflict between the two. In the authors' view the judgments in Courage and Crehan, Case C-453/99 [2001] ECR-I-6314 and Joined Cases Vincenzo Manfredi v Lloyd Adriatico Assicurazioni, C-295/04 to C-298/04 [2006] ECR I-6619 clearly indicate that the former would take precedence in case of a conflict. For a detailed exposition of this view see R Nazzini, *in Potency and Act of the Principle of Effectiveness: The Development of Competition Law Remedies in EC Law*, The Outer Limits of EU Law (C. Barnard & O. Odudu eds., forthcoming 2008). For an alternative view *see* P. Nebbia*, Damages Actions for the Infringement of EC Competition Law: Compensation or deterrence?,* 33(1) E.L. Rev. 23-43 (2008).

16. *See supra* note 15 *at* para. 26, 27.

17. *Id*.

In our view the case for reforming the private actions system rests not only on the need to ensure the victims are properly compensated but also on the increased deterrent effects created by payments of compensation. Compensation and deterrence are distinct but interrelated. Further the payment of compensation adds a third potential benefit: enabling victims to recover losses more easily could help promote the benefits of competition law to the wider public, thereby increasing support for the regime (and the market economy) as a whole.

In our view the case for reforming the private actions system rests not only on the need to ensure the victims are properly compensated but also on the increased deterrent effects created by payments of compensation.

III. The "Deterrence" Case for Private Enforcement

Private actions can increase the deterrent effect of antitrust rules in at least three ways: first by increasing the resources available for prosecution of cases; second by improving the detection and conviction rate of the regime; and third by increasing the financial consequences of detection/conviction.

A. ENHANCING DETERRENCE BY INCREASING RESOURCES AVAILABLE FOR PROSECUTION OF CASES

The first clear benefit of private actions in terms of deterrence is that they increase the resources available for the prosecution of competition law infringements. As Philip Collins, Chairman of the OFT, has explained, "competition authorities cannot, and should not, take on every case. Our work has to be prioritized, limited taxpayers' resources allocated accordingly, and the progress of cases speeded up."[18] The authors' own anecdotal observations suggest that the OFT fully investigates less than 20 percent of all cases in which it has a reasonable suspicion that the competition rules have been breached.

The Commission has expressed a similar view on several occasions. For example, in the Staff Working Paper[19] annexed to the Green Paper,[20] the Commission noted that "[p]rivate litigation can in particular deal with cases which the public authorities will not deal with, in particular due to resource constraints and other prioritization needs." It is worth noting in this regard that the Commission typically takes five to ten infringement decisions a year. For an economy the size

18. Philip Collins, *Public and private enforcement challenges and opportunities,* 15, (Jun 6, 2006), available at http://www.oft.gov.uk/shared_oft/speeches/0306.pdf.

19. Staff Working Paper Annex to the Green Paper *supra* note 13.

20. Green Paper on Damages actions *supra* note 13.

of the EU, this seems unlikely to be the optimal level of enforcement even taking into account the cartels prosecuted by NCAs.

In the United States, by contrast, for every case brought by agencies, a further nine are brought in the courts. While it is unlikely that the same ratio between private and public enforcement could be achieved in the EU in the foreseeable future, reform of the private actions system should result in an increase in the number of well-founded private actions being brought. Research suggests that current detection rates in Europe are likely to be between 10 percent and 30 percent.[21] As such, reform of the private actions system could bring significant additional resources into the competition enforcement regime in Europe as cases which are "prioritized out" or not detected by agencies are litigated in the courts.

Resource constraints, and the prioritization process it necessarily implies, not only mean that few infringement decisions are taken by competition authorities but also that they may be less likely to deal with certain types of cases. Competition agencies tend to put most of their enforcement resources into prosecuting a relatively small range of violations; what Hovenkamp[22] has called the "antitrust core." Today, this means the detection and prosecution of cartels and other "hardcore" restrictions. Even in the EU, which is widely regarded as significantly more interventionist than the United States in vertical and unilateral effect cases, infringement decisions in non-hardcore areas are rare.

This "enforcement gap" can be filled by private actions; indeed to a limited extent it is already happening. A recent review of private actions cases in Europe shows that between 2004 and 2008 over 60 percent related to vertical restraints while just under 25 percent involved abuses of dominance; less than 15 percent of cases concerned "hardcore" horizontal agreements.[23] Given the focus of public enforcement on the "antitrust core", this data suggests that private actions can not only increase the resources available for the prosecution of infringements in general but also address competition concerns in areas not prioritized by public enforcers.

B. ENHANCING DETERRENCE BY INCREASING THE LIKELIHOOD OF DETECTION

As noted below, competition authorities often cannot impose optimal fines from a deterrence perspective. However, an increase in the detection rate can compensate for the inability of agencies to levy sufficiently high fines. Private actions

21. Joint Report *supra* note 2. *See* also the discussion in section VIII i*nfra*.

22. HOVENKAMP *supra* note 3 *at* 60.

23. *Supra* note 2.

can help because in many instances private parties are better placed than agencies to detect anticompetitive conduct and bring successful prosecutions. For example, in a study analyzing a group of 29 recent successful large-scale private antitrust cases in the United States, Lande and Davies found that more than 70 percent of the total damages recovered came from cases, 12 in total, that did not follow federal, state, or EU government enforcement actions. Of the 17 cases involving the government, the scope of the courts' findings was broader than the agencies' enforcement actions in nine cases. For example, in the vitamins price fixing cartel, the private plaintiffs were able to establish both that the conspiracy had lasted considerably longer than the U.S. Department of Justice ("DOJ") pleas had indicated and that it covered a far larger range of products. Similarly in the *Automotive Refinishing Paint* case, the government's investigation yielded no indictments, whereas private cases led to a recovery of $67 million. In *Linerboard* the action by the Federal Trade Commission ("FTC") was against one firm for unilateral conduct while the private case involved a conspiracy. In *Polypropylene Carpet*, private plaintiffs obtained greater monetary recovery and prosecuted larger numbers of defendants than did the government. In *Relafen*, there was no federal case; the state governments intervened only after settlement of the private case. In *Specialty Steel*, private action led to a finding of an infringement of longer duration than did the public action.[24]

The Lande and Davies data suggests that private claimants may be able successfully to prosecute cases that the public enforcers do not, or cannot, pursue. This view is supported by a joint report, drawing on the latest research and an almost-exhaustive survey of the literature, by the Centre for European Policy Studies, the Erasmus University Rotterdam, and Luiss Guido Carli.[25] This result, which some may find almost surprising, is easily explained if one considers that the right to damages gives rise to a private incentive to prosecute competition law violations. This, combined with the information about market behavior that potential claimants are likely to possess, can substantially increase the likelihood of detection and successful prosecution of infringements provided that the right to damages can be effectively enforced.

C. CONSEQUENCES OF DETECTION/CONVICTION

Another way in which private enforcement can increase deterrence is by increasing the cost of non-compliance to infringing undertakings.

In the United States and Canada, damages—as opposed to fines—represent the lion's share of the financial implications for undertakings that breach the

24. Lande & Davis, AN EVALUATION OF PRIVATE ANTITRUST ENFORCEMENT: 29 CASE STUDIES, interim report, Aug. 11, 2006, available at http://www.antitrustinstitute.org/archives/files/550b.pdf.

25. *Supra* note 2.

competition rules.[26] Given the relatively low numbers of cases in Europe, this suggests that private actions can make a significant contribution to increasing deterrence.

GIVEN THE RELATIVELY LOW NUMBERS OF CASES IN EUROPE, THIS SUGGESTS THAT PRIVATE ACTIONS CAN MAKE A SIGNIFICANT CONTRIBUTION TO INCREASING DETERRENCE.

Competition authorities have acknowledged the impact of private action in this regard. For example, the OFT argued that "a more effective private actions system would increase the incentives of businesses to comply with competition law, since the potential incidence and magnitude of any financial liability to a competition authority and/or a claimant will increase. As these financial risks increase, so does (or should) the interest of those ultimately responsible for the governance of the business (especially supervisory boards and non-executive directors) or for supporting the business (including, for example, financiers and investor groups). In this way public enforcement and private actions are complementary."[27]

However, the fact that private damages can have this effect does not mean that it is the optimal way to achieve deterrence. In this respect, it is frequently suggested in Europe that it would be better to enhance deterrence by increasing the level of fines imposed by competition authorities than to promote private actions.[28]

The data we have reviewed suggests there may indeed be scope to increase fines in Europe. According to a recent study of cartel cases by Connor and Helmers (2006),[29] between 1990 and 2005 EU fines averaged less than 10 percent of the overcharges imposed by cartels. Other commentators have found that EU fines were in the 23 percent to 79 percent range.[30]

In our view, increasing fines by the multiples required to optimize deterrence is not a realistic option for Europe. While fines in many jurisdictions are too low (the Commission implicitly accepted this when it amended its policy in this area in 2006[31]), for policy reasons they are unlikely ever to reach the levels required for optimal deterrence. For example, as Her Majesty's Government indicated,

26. See Connor & Helmers, *Statistics on Modern Private International Cartels, 1990-2005*, (2007). Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1103610. [hereinafter Connor & Helmers].

27. PRIVATE ACTIONS IN COMPETITION LAW: EFFECTIVE REDRESS FOR CONSUMERS AND BUSINESS – DISCUSSION PAPER (OFT916, April 2007), available at http://www.oft.gov.uk/shared_oft/reports/comp_policy/oft916.pdf.

28. *See* WILS *supra* note 10.

29. *See* Connor & Helmers *supra* note 26.

30. *See* Joint Report *supra* note 2.

31. GUIDELINES ON THE METHOD OF SETTING FINES IMPOSED PURSUANT TO ARTICLE 23(2) (A) OF REGULATION NO 1/2003, OJ C 210/2 (2006).

"one option [to enhance deterrence] would be to increase the maximum level of fines significantly—perhaps six to ten times the existing maximum fines. The Government does not believe that fines at this level would be proportionate. A U.S. study indicates that more than half of firms convicted of price-fixing would go into liquidation if required to pay the optimal fine. This would not be fair. In many cases, the cartel will only have covered one aspect of the firm's business.... Very large fines would damage innocent employees, shareholders, and creditors who have done nothing to harm consumers or break the law."[32]

The (understandable) political reluctance to impose fines at optimal deterrence levels finds expression in the caps imposed on the amount authorities can levy: in both the EC and the U.K. fines cannot exceed 10 percent of the convicted firm's global turnover.[33] It is worth noting in this regard that the reluctance to impose high fines is not limited to Europe: the 1987 U.S. Sentencing Guidelines for criminal price fixing impose an upper limit of 80 percent of the guilty firm's U.S. affected sales. Similarly Connor and Helmers report that median penalties worldwide were less than 21 percent of actual overcharges; in the United States and Canada median average fine ratios were in the range of 15 percent to 18 percent.[34]

It is important in this regard, however, that any reforms to the private actions system are made in such a way as to ensure that unwarranted actions are minimized.

As stated in the joint report by the Centre for European Policy Studies, the Erasmus University Rotterdam, and Luiss Guido Carli, damages awards in private actions can act as a complement to public enforcement in a second-best context, when the optimal solution is impossible to achieve. This is because damages, if they are not multiplied, are no more than reparation of harm that is unlawfully caused. In addition, they are paid to those who have suffered harm rather than disappearing into State coffers. As such, they are likely to have greater legitimacy and political support. It is important in this regard, however, that any reforms to the private actions system are made in such a way as to ensure that unwarranted actions are minimized. How this is to be achieved is discussed below.

32. A World Class Competition Regime, White Paper Cm 5233, (July 2001).

33. Under section 36(8) of the Competition Act 1998, '[n]o penalty fixed by the [OFT] under this section may exceed 10 percent of the turnover of the undertaking (determined in accordance with such provisions as may be specified in an order made by the Secretary of State.' *See also* Competition Act 1998 (Determination of Turnover for Penalties) Order, 2000SI 2000/309. Under Article 23(2) of Regulation 1/2003, '[f]or each undertaking and association of undertakings participating in the infringement, the fine shall not exceed 10 percent of its total turnover in the preceding business year. Where the infringement of an association relates to the activities of its members, the fine shall not exceed 10 percent of the sum of the total turnover of each member active on the market affected by the infringement of the association.' See also Article 23(4).

34. *See* Connor & Helmers *supra* note 26.

IV. The "Compensation" Case for Private Enforcement

A. INFRINGEMENTS OF COMPETITION LAW CAUSE SIGNIFICANT HARM

The harm arising from infringements of competition law is significant.[35] It has been estimated recently that the total overcharge from EU-wide cartels could be between EUR 30 billion and EUR 138.7 billion.

Staggeringly, the above figures reflect only a portion of the cartels operating in Europe. According to a report by Connor (2005), penalties imposed by EU countries on 72 cartels in the period from 1990 to 2005 totaled $1.9 billion in real 2005 dollars. Of the 72 cases in question, 67 were brought in Western European countries (totaling $1.86 billion in real 2005 dollars) and five in Eastern EU Member States (totaling $43 million). Against this background, the penalties imposed by the Commission in the same period in respect of 86 cartels were $2.15 billion in real 2005 dollars ($961.2 million for EU-wide cartels, plus EUR 1.188 billion for global cartels also sanctioned by the Commission). This means that, for the period from 1990 to 2005, penalties imposed at the Member State level were 88.4 percent of penalties imposed at EU level. If the assumptions on the detection rate and the ratio between penalties and overcharges used to calculate the figures in the above paragraph are applied to national cartels, the annual impact of national cartels would range from EUR 7.88 billion to EUR 122.55 billion.

It is clear, therefore, that even under conservative assumptions cartels can result in massive unlawful transfers from buyers to sellers. This is unjust. In addition, as a matter of law[36] European businesses and consumers have the right to recover compensation for the harm caused to them. In our view, the regime in Europe must be reformed to ensure that those who have been harmed by antitrust infringements can effectively exercise their right to recover damages.

B. PRIVATE ACTIONS AS THE MEANS OF OBTAINING COMPENSATION

Some commentators have argued that competition authorities are best placed to obtain compensation for victims.[37] We do not share this view. As EC Competition Commissioner Neelie Kroes has explained

35. This section relies heavily on the data in Joint Report *supra* note 2. Uncited data in this section are drawn from this report.

36. *See* the discussion at section II above.

37. *See*, for example, Wils *supra* note 10.

> "... no matter how closely public intervention mirrors the concerns of consumers, no matter how effectively the fines that we impose punish and deter unlawful behaviour, the victims of illegal behaviour will still not be compensated for their losses. Public enforcement is simply not there to serve this goal. It is there to punish and deter illegal behaviour. It cannot make amends for the damage and suffering caused to consumers. Therefore, consumers should be empowered to enforce their rights themselves."[38]

This reflects the view that public enforcement, by its nature, is not designed to provide full compensatory redress to consumers either individually or collectively, whereas the civil justice system is designed for this purpose: "[u]nlike courts, which address and enforce the rights of individuals, the authorities act in the general interest."[39]

In 2006, Richard Macrory, a barrister and professor of economics, was asked by Her Majesty's Government to look at regulatory regimes in the United Kingdom. He concluded[40] that the primary function of these regimes is to ensure compliance with statutory and other regulatory norms through punitive sanctions and deterrence; the aim of these regimes is not to compensate victims.

There are a number of reasons why we share the views of Kroes and Macrory. First, the standard of proof imposed on competition authorities may be higher than that imposed on claimants in the civil courts. Indeed, in the Regulatory Enforcement and Sanctions Act 2008, the UK Parliament made it more difficult for regulators (which include competition authorities for these purposes) to obtain compensatory awards than in the past—the regulator must now be satisfied that the criminal standard of proof is met before taking such action.[41]

38. MAKING CONSUMERS' RIGHT TO DAMAGES A REALITY: THE CASE FOR COLLECTIVE REDRESS MECHANISMS IN ANTITRUST CLAIMS, at 3, (Sept. 11, 2007) *available at* http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/07/698&format=HTML&aged=0&language=EN&guiLanguage=en.

39. ENHANCING ACTIONS FOR DAMAGES FOR BREACH OF COMPETITION RULES IN EUROPE *at* 2 (Sept. 22, 2005) available at: http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/05/533&format=HTML&aged=0&language=EN&guiLanguage=en.

40. Macrory, REGULATORY JUSTICE: MAKING SANCTIONS EFFECTIVE, (Final Report) (November 2006). Macrory's six penalties principles are as follows: "A sanction should: 1. Aim to change the behavior of the offender; 2. Aim to eliminate any financial gain or benefit from noncompliance; 3. Be responsive and consider what is appropriate for the particular offender and regulatory issue, which can include punishment and the public stigma that should be associated with a criminal conviction; 4. Be proportionate to the nature of the offense and the harm caused; 5. Aim to restore the harm caused by regulatory non-compliance, where appropriate; and 6. Aim to deter future non-compliance."

41. REGULATORY ENFORCEMENT AND SANCTIONS ACT, Section 42(2) (2008).

Second, even if competition authorities were able to require compensation to be paid, obtaining such awards is likely to be a secondary consideration in the context of the regime. This is likely to be particularly acute in cartel cases where authorities in Europe are only required to prove that the "object" of the parties to the cartel was anticompetitive. To obtain redress, however, requires extensive analysis of the cartel's effect. Gathering and assessing the relevant data is highly resource intensive and would likely add to the complexity and duration of the investigations; it is also likely to increase the litigation burden on authorities since such findings are likely to be challenged. In this situation obtaining redress—in other words, acting to protect private interests—could easily conflict with the competition authority's main role of acting in the public interest. Unsurprisingly, the public interest is likely to be given priority.

Finally, and perhaps most importantly (as noted above), authorities simply do not have the resources to take on all cases which raise competition issues. As Commissioner Kroes has noted "[a]nyone harmed by unlawful action should not have to wait for a public body to intervene."[42]

This is not to suggest that authorities can never or should never seek redress for victims. It is conceivable, in certain cases, that compensation to those who have been harmed could be secured in the context of public enforcement—as part of a settlement, as an additional element of the leniency program, or as a spontaneous initiative by the perpetrator (who may then plead that the payment compensated for the harm, thereby mitigating any financial penalties).

However, public and private enforcement should be kept distinct. Public enforcement must focus its resources on the infringements which need the deterrent effect of public sanctions. This may be the case for secret cartels where the likelihood of detection would be lower in the absence of public enforcement or for cases establishing a new principle or modifying existing legal doctrine. Those who have been harmed by an alleged antitrust infringement must be free to pursue their claim in the courts regardless of whether a competition authority has taken action in the same matter. To the extent there are barriers to effective redress through the courts, these barriers must be removed or alleviated as far as possible. The answer is not to erect an additional barrier by limiting the right to damages for competition law infringements to redress that can be obtained in public enforcement proceedings.

C. BRINGING COMPETITION POLICY CLOSER TO CONSUMERS

Competition rules help maximize the welfare of society; benefits include lower prices, larger output, and increased productivity. However, these benefits are often

42. MORE PRIVATE ANTITRUST ENFORCEMENT THROUGH BETTER ACCESS TO DAMAGES: AN INVITATION FOR AN OPEN DEBATE, 3, (Mar. 9, 2006) available at http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/06/158&format=HTML&aged=0&language=EN&guiLanguage=en.

not ascribed by the general public to competition law and its enforcement. Too often, competition law and enforcement are seen as the preserve of a closed group of specialists operating in an esoteric area which is detached from everyday life. More effective redress for businesses and consumers is likely to bring home to the general public the purpose of competition law and the benefits that its effective enforcement produces. This awareness may provide stronger legitimacy to the competition regime, which could result in enhanced effectiveness if support for robust enforcement action increases on the part of those who ultimately benefit.

MORE EFFECTIVE REDRESS FOR BUSINESSES AND CONSUMERS IS LIKELY TO BRING HOME TO THE GENERAL PUBLIC THE PURPOSE OF COMPETITION LAW AND THE BENEFITS THAT ITS EFFECTIVE ENFORCEMENT PRODUCES.

The rest of this paper seeks to identify the central changes that are needed if private actions are to play a more significant role in the competition regime in the future and then assesses the proposals for reform put forward by the Commission and the OFT.

V. Basic Structure of Optimal Private Actions Regime

As discussed above, private actions are underdeveloped in Europe. However, both as a matter of law and as a matter of policy, they have an important dual role to play, first to increase the deterrence and effectiveness of the regime and second to secure compensation to those who have been harmed as a result of infringements. Increased deterrence and compensation almost always go hand in hand. However, when these two objectives conflict, both as a matter of law[43] and as a matter of policy, the objective of increasing deterrence and ensuring the effectiveness of the regime should prevail.

We believe that a well-functioning private actions regime should rest on the robust structuring and fine tuning of three main pillars: a clear and sound legal framework for collective actions, a European-wide solution to the problems of indirect purchasers' standing and passing-on, and coordination between public and private enforcement, ensuring the centrality of the former. This paper will focus on these three areas. The Commission White Paper and the OFT Recommendations put forward a number of other proposals regarding disclosure, the requirement to prove fault, and costs. However, a number of these proposals, while important, are less fundamental than our pillars. As regards costs, the European Union has a number of restrictions on funding legal services that constitute a major barrier to effective redress, particularly in those Member States where legal costs are very high.

43. *Supra* note 15.

However, it is likely that, if the three pillars of an effective private actions system mentioned above are in place, the market may be able to deliver adequate solutions to the current funding problems. A summary of the proposals put forward by the Commission and the OFT can be read in the Annex.

A. COLLECTIVE ACTIONS

1. Need for Opt-out Collective Action

The first pillar of an effective private actions regime is a robust legal framework for collective actions. Collective actions are procedural mechanisms bundling a number of individual claims in one set of proceedings.[44] They may be of two types. Opt-in collective actions ("opt-in actions") are based on the principle that an action may only be brought on behalf of persons who have expressly consented to be represented in the proceedings. Opt-out collective actions ("opt-out actions") are based on the principle that the action may be brought on behalf of an appropriately defined class of affected persons who will be bound by the outcome of the litigation unless, having being adequately informed of the proceedings, they state their intention not to be represented.

Generally, systems of civil procedure envisage mechanisms whereby two or more individual claims can be brought together so that common issues may be decided, once, in a way that binds all the claimants. Principles of judicial economy and avoidance of conflicting judgments make the availability of such a procedure highly desirable—if not necessary—in any legal system. More recently, however, collective actions have played an additional role in modern societies: ensuring access to justice and the effective enforcement of the law.[45]

Competition infringements may harm a significant number of persons. An individual loss may be relatively small but the aggregate loss to all potential claimants may be large. Both of the national cases that gave rise to the references to the ECJ on the right to damages for breach of Article 81(1) are instances in which an infringement of competition law affected a significant number of businesses or consumers in a similar way. In the *Courage and Crehan* case, the issues related to the anticompetitive effects of a "beer tie" agreement.[46] A significant number of publicans were in the same position as Mr. Crehan as they had been lessees of tied houses during the relevant period and claimed to have suffered loss as a result. In the *Manfredi* case, a policy holder claimed damages against the

44. On collective actions, see C. Hodges, Multi-Party Actions (2001) and R. Mulheron, The Class Action in Common Law Legal Systems: A Comparative Perspective (2004) [hereinafter Mulheron].

45. Mulheron *supra* note 44 *at* 63 – 66.

46. A beer tie agreement is a clause in a contract between the tenant of a public house and its landlord that obliges the former to purchase almost all of its beer supply from either the landlord or a company nominated by it at the list price in force.

insurers alleging that the insurance premiums of compulsory civil liability insurance relating to accidents caused by motor vehicles had been artificially increased as a result of a cartel among the insurers.[47] A large number of motor vehicle owners in Italy were similarly affected. In cases such as these, given the size of each individual claim relative to the costs of bringing the claim, individual claimants may be effectively deterred from bringing proceedings even if they have a well-founded case. The result may be that—in the absence of an effective collective redress mechanism—when the perpetrator of an infringement harms a great number of individuals but the individual loss is not sufficiently large to justify the costs and risks of bringing an individual claim, the perpetrator will escape liability for the loss it caused and those harmed will not be compensated.

A system of collective redress relying exclusively on opt-in actions is inherently ineffective.

In our view, the answer to this problem is the availability, within an appropriately designed legal framework, of an opt-out action. A system of collective redress relying exclusively on opt-in actions is inherently ineffective.[48] In the United Kingdom, where only opt-in representative actions are allowed, only one such action has been brought since the relevant provisions of the Competition Act 1998, as amended by the Enterprise Act 2002, entered into force on June 20, 2003. The level of take-up by consumers was low compared to the scale of the infringement. In the end, about 600 consumers joined the action, with aggregated damages only in the thousands of pounds. In that case, a follow-on action from the decision of the OFT in *Replica Football Kits*, the OFT estimated that the unlawful arrangements would have cost the consumer over 50 million pounds had the arrangements not been brought to an end.[49]

Evidence from other jurisdictions points in the same direction. In France, an action brought by UFC-Que Choisir? on behalf of mobile phone users allegedly harmed by a cartel among mobile phone operators had a take-up of around 12,000 consumers. It would appear, however, that around 20 million consumers had been affected by the infringement.[50]

---

47. The Autorità Garante della Concorrenza e del Mercato had established the infringement in its decision *No 8546 (I377) of 28 July 2000*, Bollettino 30/2000 (2000).

48. See, for instance, R. Mulheron, Reform of Collective Redress in England and Wales: A Perspective of Need, Civil Justice Council of England and Wales' research paper submitted 2008, available at http://www.civiljusticecouncil.gov.uk. at the website of the Civil Justice Council of England and Wales [hereinafter Mulheron Reform].

49. U.K. Office of Fair Trading, Private actions in competition law: effective redress for consumers and business (OFT Discussion Paper, No. 916resp, Nov. 2007).

50. Transcript of the OFT's public hearing on private actions (Sept. 24, 2007). For more information about the UFC-Que Choisir? action *see* www.cartelmobile.com.

One can generalize from these facts that, with opt-in collective action models, the perpetrator of a competition law infringement is not at risk of having to compensate for the full harm it caused. Minimizing this risk not only impairs the right to damages from acting as a deterrent against engaging in anticompetitive conduct, but also fails to deliver compensation to those who have been harmed by the conduct in question.

An opt-out action, on the other hand, has clear benefits both in terms of achieving deterrence and in terms of securing compensation. Because opt-out levels are relatively low, opt-out collective actions optimize litigation economies of scale, avoid (to a significant extent) duplicative litigation, and minimize the risk of inconsistent judgments. Therefore, if properly designed and managed, opt-out collective actions can deliver significant benefits to society in terms of deterring anticompetitive behavior, thus promoting consumer welfare and productivity at the lowest possible cost. The counterfactual to an opt-out collective action is that the perpetrators of the infringement are generally not at risk of the entire loss they caused. As a consequence, deterrence is low. The other possible, but unlikely, counterfactual is that all claimants sue and are compensated in individual or opt-in actions. In such a scenario, the deterrent effect is achieved but at a higher cost to society because the aggregate cost of individual actions or opt-in collective actions is likely to be higher than the cost of an opt-out action. In terms of compensation, it must be recognized that, if the individual loss is small but the total harm is large and the issues to be litigated are complex, the most likely counterfactual to an opt-out action may be that private actions (whether on an individual basis or as opt-in collective actions) are either not brought at all or are brought by or on behalf of a small minority of those who have been harmed. In these cases, opt-out actions are necessary to ensure that those who have been harmed as a result of competition law infringements obtain the compensation they are entitled to.

2. Possible Objections to Opt-out Collective Actions

Notwithstanding the clear benefits of opt-out collective actions, there have been a number of objections to this model. Broadly, they fall under the following categories: a) opt-out collective actions do not achieve compensation since compensation always presupposes that the claimant opts-in at some point; b) opt-out collective actions give rise to a disproportionate risk of abuse because, given the potentially very significant damages at stake, the defendants are under pressure to settle even unmeritorious cases; c) opt-out collective actions may raise problems under Article 6 of the European Convention of Human Rights and under the constitutional provisions of some Member States; and d) opt-out collective actions are not consistent with the legal traditions of the Member States.

We will briefly deal with these objections in turn.

Compensation. The argument that an opt-out collective action is not of a compensatory nature is largely fallacious. Often, the only viable vehicle to pursue a significant number of small claims raising complex issues of law and fact is to aggregate them so that a critical mass is achieved making the action worthwhile in terms of attracting the necessary funding. Because of the low take-up levels of opt-in collective actions, the only mechanism to achieve this objective in certain cases is an opt-out collective action. The counterfactual to the availability of an opt-out collective action is often no compensation at all. An opt-out collective action plays a fundamental role in ensuring access to justice in this category of cases. Furthermore, even if the class members who ultimately claim under a settlement or judgment are only a percentage of all the class members, compensation in an opt-out collective action would still be superior for those who do claim under the settlement or judgment if the claims in question would not have been viable as individual actions or opt-in collective actions. Compensation is also achieved when any unclaimed funds are applied to the benefit of the category of consumers or businesses harmed by the infringement in question under the so-called cy pres distribution. While, by their own choice, some members of the class do not recover damages, the society sector which had to bear the brunt of the infringement receives tangible benefits.

Opt-out actions promote both private action functions. They deliver compensation in cases when, in the absence of an opt-out action, no claim would be brought. At the same time, they increase deterrence by placing the perpetrators of competition law infringements at risk of having to compensate the full harm caused.

In any event, arguments about the allegedly non-compensatory nature of opt-out actions become otiose when one considers that private actions have a dual function: not only to compensate those who have been harmed but also to deter anti-competitive behavior, thus enhancing long-term social welfare and productivity for the benefit of the society as a whole. In terms of deterrence, if the choice is between making the perpetrators pay for the full harm caused or letting them benefit from the barriers faced by claimants in aggregating claims in an opt-in collective action or bringing them on an individual basis, in our view the former must be preferred. In this way, opt-out actions promote both private action functions. They deliver compensation in cases when, in the absence of an opt-out action, no claim would be brought. At the same time, they increase deterrence by placing the perpetrators of competition law infringements at risk of having to compensate the full harm caused.

Risk of Abuse. It is often claimed that an opt-out collective action is open to abuse. The argument is as follows: Because an opt-out collective action can potentially produce a substantial level of damages, defendants will often find it preferable to settle even unmeritorious cases rather than running the risk of going to trial. The argument might have some force in the United States, where the claimant not only has an automatic right to treble damages but also the constitutional right to a jury trial. In a trial by jury, the verdict might be less predictable and, in some cases, possibly, biased against the defendant. The argument, even if it were true in the United States, has much less force if the trial is by judge alone. If fact, in England and Wales the evidence would suggest that the chances of a claimant succeeding on the merits are not high. In the vast majority of English and Welsh competition cases tried on their merits, the claimants failed, including where cases were clearly not speculative or unmeritorious. Awards of damages have been very rare.

Furthermore, it is possible to design a system which has appropriate safeguards against any risk of abuse. For instance, a preliminary stage may be designed in which a number of threshold requirements must be fulfilled before an opt-out collective action is allowed to proceed.[51] It is important, however, that such safeguards do not unduly restrict the availability of opt-out collective actions or disproportionately raise their costs.

ECHR, Article 6 and National Constitutional Fair Trial Provisions. It is sometimes argued that an opt-out collective action may raise issues under Article 6 of the European Convention on Human Rights ("ECHR") and may be incompatible with national constitutional fair trial provisions. The analysis focuses on Article 6 of the ECHR but it is submitted that it should be possible to arrive at the same conclusion in respect of the relevant national constitutional fair trial provisions.

The problem appears to be that those who do not opt-out of the action in the prescribed way are bound by the outcome of the litigation. This—it is argued—may be in conflict with the right to access to a court (a right enshrined in Article 6 of the ECHR) because those who did not explicitly express their consent to participate in the action are nevertheless bound by any settlement or judgment, preventing them from bringing an individual claim if and when they wish. This argument is misconceived. Those who do not exercise their right to opt-out of the action have full access to a court. By not opting-out, they choose to participate in the action with all the resultant consequences in terms of both the binding effect of any judgment or settlement and the inability to bring further proceedings on the same or, in certain circumstances, related cause of action. To the extent that the need to opt-out of an action brought by another person may be seen as a lim-

51. These may include a certain minimum threshold relating to the allegations and evidence provided at pleading stage and need not be specific to out-out actions: see, in the U.S., Bell Atlantic v. Twombly, 127 S Ct 1955, 1964 (2007) and, in England and Wales, CPR, 3.4 (on the court's power to strike out a statement of case that discloses no reasonable grounds for bringing or defending the claim) and 24.2 (on the grounds for summary judgment if the claimant has no reasonable prospect of succeeding on the claim or issue). See also the White Paper on Damages actions *supra* note 4 *at* 5, which proposes that the conditions for a disclosure order should include that "the claimant has **presented all the facts** and **means of evidence** that are **reasonably available** to him, provided that these show PLAUSIBLE GROUNDS to suspect that he suffered harm as a result of an infringement of competition rules by the defendant." (emphasis in the original).

itation of an unfettered right to court access, it need only be stressed that such a right is never unqualified. Under Article 6 of the ECHR, not all access restrictions to a court are an infringement of the right to a fair trial.[52] The compatibility of opt-out collective actions with Article 6 of the ECHR depends on how the system is designed and the effectiveness of the publicity requirements supervised or mandated by the court. There is nothing in the basic features of an opt-out collective action which makes it incompatible with Article 6 of the ECHR.

THERE IS NOTHING IN THE BASIC FEATURES OF AN OPT-OUT COLLECTIVE ACTION WHICH MAKES IT INCOMPATIBLE WITH ARTICLE 6 OF THE ECHR.

Legal Traditions of the Member States. Some argue that opt-out collective actions are not embedded in the legal traditions of the Member States.[53] This argument rests on a strong path-dependence assumption, essentially denying the possibility of any legal reform which is not an incremental change to the existing legal framework. Even conceding that this is the only scope for legal reform in the European Union, it must be stressed that a number of Member States have adopted opt-out collective redress systems, including Denmark,[54] Portugal,[55] Spain,[56] the Netherlands,[57] and Norway.[58] In England and Wales, the representative party action has long been recognized[59] and the Civil

52. Golder v. United Kingdom EHRR 524 (1975) and Ashingdane v. United Kingdom 7 EHRR 528 (1985).

53. For a judicial *dictum* see Campos v. Kentucky & Indiana Terminal Railroad Co., 2 Lloyd's Rep 459 (1962).

54. Lov nr 181 28.2.2007 (Act no 181 of Feb. 28, 2007).

55. *See* Right of Proceeding, Participation and Popular Action, Law No 83/95 of 31st August, and Establishing the Legal System Applicable to Consumer Protection, Law No 24/96 of 31st July, discussed in Mulheron Reform *supra* note 48 at 97 – 101.

56. Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil (BOE núm 7, de 8 de enero del 2000, pp 575-728. Corrección de errores BOE núm 90, de 14-04-2000, p 15278 y BOE núm 180, de 28-07-2001, p 27746) (LEC). See, in particular, LEC, Libro I, Título I, Capítulo 1, Articulo 6 Capacidad para ser parte: Podrán ser parte en los procesos ante los tribunales civiles: . . . 7.º Los grupos de consumidores o usuarios afectados por un hecho dañoso cuando los individuos que lo compongan estén determinados o sean fácilmente determinables. Para demandar en juicio será necesario que el grupo se constituya con la mayoría de los afectados.

57. Wet van 23 juni 2005 tot wijziging van het Burgerlijk Wetboek en het Wetboek van Burgerlijke Rechtsvordering teneinde de collectieve afwikkeling van massaschades te vergemakkelijken (Wet collectieve afwikkeling massaschade), Staatsblad 2005, 340.

58. Act Relating to the Litigation Procedure in Civil Disputes (The Dispute Act), no 6, LOV-1915-08-13-6, (August 13, 1915). *See* C. Bernt-Hamre, Class Actions, Group Litigation and Other Forms of Collective Litigation in the Norwegian Courts (National Report prepared for the conference 'The Globalisation of Class Actions', Oxford, (Dec.12-14, 2007), available at http://www.law.stanford.edu/display/images/dynamic/events_media/Norway_National_Report.pdf).

59. CPR, r 19.6.

Procedure Rules ("CPR") provide for group litigation orders.[60] As these examples demonstrate, it is clearly possible to design an opt-out collective action model which is compatible with the legal systems of the Member States.

B. INDIRECT PURCHASERS' STANDING AND PASSING-ON DEFENSE

The extent to which there should be any limitation on indirect purchasers' standing and the availability of the passing-on defense is one of the most controversial issues in relation to private actions in competition law.

In the United States, the question is far from settled. There has been a considerable backlash against the ruling of a majority of the U.S. Supreme Court in *Illinois Brick Co v Illinois*,[61] pursuant to which claims by indirect purchasers were precluded under federal law. A majority of U.S. states have now enacted '*Illinois Brick* Repealer' statutes to preserve indirect purchasers' rights to sue in state courts. In many of those states, the Supreme Court's earlier majority ruling in *Hanover Shoe v United Shoe Mach*,[62] pursuant to which the passing-on defense was excluded, has also been overturned. The Antitrust Modernization Commission recommended that Congress overrule the Supreme Court's decisions to the extent necessary to allow both direct and indirect purchasers to recover.[63]

> WHILE THE DETERRENCE-BASED ARGUMENTS FOR DISALLOWING THE PASSING-ON DEFENSE AND EXCLUDING THE STANDING OF INDIRECT PURCHASERS ARE UNDOUBTEDLY POWERFUL, IT IS UNCLEAR WHETHER THESE MEASURES WOULD ACHIEVE ANY OF THE INTENDED BENEFITS.

The argument for disallowing the passing-on defense and excluding the indirect purchaser's standing rests entirely upon a deterrence rationale. This rationale argues that direct purchasers, as compared to indirect purchasers, are the best placed to sue because of their knowledge of the market, access to evidence, and relative ease of proving the overcharge. However, if direct purchasers have to litigate the issue of whether they passed on any overcharge, in full or in part, to purchasers further down the supply line, this would deter them from suing in the first place. Since they are the best placed to sue and, in most circumstances, indirect purchasers will not bring an action, allowing the passing-on defense undermines the effectiveness of the regime. If the passing-on defense is disallowed, a necessary corollary would appear to be the exclusion of the standing of indirect purchasers in order to avoid multiple recoveries in respect of the same harm.

60. CPR, rr 19.10 – 19.15. Rule 19.12 envisages circumstances in which a judgment or order may bind the parties to a claim which is entered on the group register after the order or judgment was made.

61. Illinois Brick Co. v. Illinois, 431 US 720 (1977).

62. Hanover Shoe v. United Shoe Mach., 392 US 481 (1968).

63. See ANTITRUST MODERNIZATION COMMISSION, REPORT AND RECOMMENDATIONS, April 2007, available on www.amc.gov/report_recommendation/toc.htm.

While the deterrence-based arguments for disallowing the passing-on defense and excluding the standing of indirect purchasers are undoubtedly powerful, it is unclear whether these measures would achieve any of the intended benefits. Direct purchasers may, for example, share with their suppliers the benefits of an overcharge or may attach a greater importance to maintaining good commercial relations with their suppliers. These considerations would be less likely to apply to indirect purchasers. The threat of action by indirect purchasers, therefore, may well be crucial in terms of achieving deterrence. Nor should one underestimate the deterrent effect of the threat of private actions by a wider group of claimants, including both direct and indirect purchasers. Finally, the *Courage and Crehan* and *Manfredi* cases suggest that EC law itself requires that, in order to ensure the effective enforcement of the EC competition rules, all persons harmed by an infringement of Articles 81 and 82 should be able to recover their loss provided that the other requirements to obtain compensation are met.[64] Last but not least, excluding indirect purchasers would run counter to the compensatory function of private actions. For all of these reasons, until further research is done, any limitation on the standing of consumers and other end users would not be appropriate at this stage.

It is, however, important that 'passing-on' does not become a powerful shield for defendants to escape liability and, as a result, a disincentive for direct or indirect purchasers to bring an action. To the extent this can be achieved by reforming the procedural and evidential rules while at the same time preserving both the standing of indirect purchasers and the possibility for defendants to prove that the claimant passed on the overcharge to its customers this would appear to be preferable to reforming the substantive rules on liability.

C. INTERACTION BETWEEN PUBLIC AND PRIVATE ENFORCEMENT: LENIENCY

As discussed above, private actions have a dual function: to increase deterrence and deliver compensation. While these two functions almost always go hand in hand, there may be cases in which they conflict. In such cases, our view is that the function of increasing deterrence should prevail both as a matter of law and as a matter of policy. This means that the right to damages may have to be limited whenever its compensatory function conflicts with its deterrent function. One such area is the interaction between public and private enforcement. Such an interaction may occur in different ways. In this paper, we focus on the interaction between leniency programs and private actions.

Leniency programs are designed to reward, with either immunity from fines or reduced fines, undertakings that reveal to the competition authorities the existence

64. Courage Ltd v. Bernard Crehan, Case C-453/99 ECR I-6297, paragraphs 26 - 28 (2001) and Vincenzo Manfredi v. Lloyd Adriatico Assicurazioni SpA, Case C-295/04 ECR I-6619, paragraphs 59 – 61 (2006). At the same time, national courts can take steps to ensure that claimants are not unjustly enriched: Vincenzo Manfredi v. Lloyd Adriatico Assicurazioni SpA, paragraph 94.

of a cartel or provide useful evidence in the course of an investigation.[65] Leniency programs are generally seen as an essential tool in the fight against cartels.[66]

The decision of an undertaking to apply for leniency is a complex one. The likelihood of detection and the likely amount of any adverse financial consequences are the main factors taken into account. As the incentives for individuals who may be personally liable are not aligned with the incentives of the undertaking, the threat of personal sanctions, either of a criminal[67] or civil law nature,[68] increases the undertaking's uncertainty as to whether the cartel will be uncovered. On the other hand, from the undertaking's point of view, the sanction is not only the fine that may be imposed by a competition authority but also any damages that may be recoverable by those who have been harmed by the infringement. As the latter increase, the relative benefits of any reduction or immunity from public law fines decrease. Furthermore, an undertaking in receipt of leniency is at risk of being the primary, and perhaps the sole, target of a private action. The reasons are largely practical. First, a claimant would generally assume that the leniency applicant is likely to have important or even crucial evidence in his possession that may be obtained through disclosure or, in civil law systems, through a court order relating to specific documents. Second, it is tactically very difficult for a leniency applicant to dispute its liability in court even if the relevant competition authority has not yet made an infringement decision or, technically, the decision of the competition authority would not bind the court. Finally, if the relevant competition authority has not yet made a decision, the claimant will assume that the leniency applicant is likely to be an addressee of any infringement decision while there may be some uncertainty as regards other parties being investigated. If the leniency applicant is jointly and severally liable with the other cartelists, there is a strong incentive for the claimant to sue the leniency applicant and, possibly, only the leniency applicant[69] for the entire loss.

65. For an overview of the leniency policy *see* FIGHTING HARDCORE CARTELS: HARM, EFFECTIVE SANCTIONS, AND LENIENCY PROGRAMMES OECD (2002).

66. *See* the COMMISSION NOTICE ON IMMUNITY FROM FINES AND REDUCTION OF FINES IN CARTEL CASES, OJ C 298/11, at paragraphs 1 – 5 (2006). As of July 1, 2007, competition authorities in 24 Member States operated a leniency program. The U.S. Department of Justice, Antitrust Division, operates a Corporate Leniency Policy and a Leniency Policy for Individuals. For an overview of the U.S. system see R. NAZZINI, CONCURRENT PROCEEDINGS IN COMPETITION LAW: PROCEDURE, EVIDENCE AND REMEDIES 415-417 (2004) [hereinafter NAZZINI].

67. *See*, in the U.K. Enterprise Act 2002, ss 188 and 189.

68. In the U.K., the disqualification of company's directors under the Company Directors Disqualification Act 1986.

69. If the leniency applicant is solvent and able to satisfy the entire claim, the claimant would not have an incentive to sue the other cartelists as he would be exposed to adverse costs orders in relation to more than one defendant. The costs of the litigation are also likely to be higher the more defendants are jointly sued in the same action.

While the increased likelihood of private litigation and the increased magnitude of damages at stake, especially in opt-out actions, may appear at first sight to lower the incentives to apply for leniency, a leniency application can remain very attractive even if the undertaking in question factors in potential damages. Private actions and the threat of personal sanctions increase the likelihood of detecting the cartel. If the cartel is uncovered and the undertaking has not made a timely application for leniency, the potential liability would include the entire amount of the fine plus any liability in damages.

In our view, the theoretical arguments and the anecdotal evidence suggest that an effective private actions regime per se is not likely to have a negative impact of the effectiveness of the leniency programs. However, there may be specific aspects of the civil litigation system that conflict with specific aspects of the leniency regime. Experience has shown that tension may arise when public enforcement proceedings are conducted in parallel with private actions. The claimant may be entitled to disclosure from the defendant. Such disclosure may extend to documents submitted by a leniency applicant to a competition authority. Another tension may arise if the leniency applicant, and, in particular, the first applicant that qualifies for full immunity ("immunity recipient"), is jointly and severally liable with the other cartelists and likely to be the primary or sole target of private actions. In light of these two areas of potential tension between the private actions regime and the leniency program, any reforms aimed at promoting private enforcement, especially if opt-out actions are introduced, should adequately address any negative impact on the effectiveness of the leniency regime. In our view, this is an area in which the deterrent function of private actions prevails over the compensatory functions, in that precedence should be given to the protection of the integrity of the public enforcement process.

Any reforms aimed at promoting private enforcement, especially if opt-out actions are introduced, should adequately address any negative impact on the effectiveness of the leniency regime. In our view, this is an area in which the deterrent function of private actions prevails over the compensatory functions.

The following sections analyze the OFT's and the Commission's proposals against the benchmark of the principles and models set out in this section.

VI. Proposals Relating to Representative Actions

This section examines the representative actions proposals put forward by the Commission and the OFT against the benchmark of the dual function of private actions, namely to increase deterrence and secure compensation.

The Commission and OFT proposals both recommend an opt-out representative action. There is an emphasis on appropriate safeguards, one of which is the adoption of an 'ideological claimant' model designed to act as a filter to avoid speculative litigation.

The Commission proposes to allow "representative actions, which are brought by qualified entities, such as consumer associations, state bodies, or trade associations, on behalf of identified or, in rather restricted cases, identifiable victims".[70] Under the Commission's "ideological claimant" model, only "qualified entities" (rather than a class member) have standing to bring an action on behalf of those who have been harmed. Such "qualified entities" could be designated, according to the Commission, either on a permanent basis or on an ad hoc basis. Entities designated on a permanent basis are those representing "legitimate and defined interests" which meet criteria to ensure that abusive litigation is avoided.[71] Such entities would be able to bring actions on behalf of identified or identifiable persons, even if not their members.[72]

Ad hoc designated entities are entities "whose primary task is to protect the defined interests of their members, other than by pursuing damages claims (e.g. a trade association in a given industry) and which give sufficient assurance that abusive litigation is avoided".[73] Under the Commission's proposals, ad hoc designated entities would be able to bring actions only on behalf of their members.

Actions on behalf of identified victims could be brought on an opt-out basis, i.e. the victim is represented in the action unless he states his intention not to be bound by the outcome of the litigation. Actions on behalf of identifiable victims can only be brought as opt-out collective actions. If the victims are not identified, by definition they cannot have given their express consent to be bound by the outcome of the litigation.

In its 2007 Recommendations to HM Government, the OFT recommended that representative bodies should be able to bring actions on behalf of either named consumers or businesses or consumers or businesses at large. The OFT does not define the 'representative body' but, like the Commission, proposes that 'representative bodies' could be either designated in advance on a permanent basis or by the court on an ad hoc basis.

Unlike the Commission, the OFT does not recommend that only representative bodies designated in advance on a permanent basis should be able to bring

70. White Paper on Damages actions *supra* note 7 at 2.1.

71. Staff Working Paper on Damages actions *supra* note 7 at para. 52.

72. White Paper on Damages actions *supra* note 7 at 4.

73. Staff Working Paper on Damages actions *supra* note 7 at para. 53.

opt-out actions on behalf of identifiable victims. Nor does the OFT limit ad hoc representative bodies to representing their members. Therefore, under the OFT's proposals, any representative body could bring an opt-out action if designated in advance or given permission to bring the action on an ad hoc basis.

In the OFT's Recommendations, the emphasis is on judicial discretion and case management. In particular, the OFT recommended that it should be open to the judge to decide, in the circumstances of each case but on the basis of appropriately defined criteria and filters, whether given claims should be brought as a representative action on behalf of consumers/businesses at large, as a representative action on behalf of named consumers/businesses, or as individual actions.[74]

Comparing the Commission's proposals with the OFT's recommendations, it appears that under the Commission's proposals, the availability of opt-out actions may be unduly restricted. This would be the case, for instance, if an action could only be brought on behalf of identifiable victims. If there is no representative body designated on a permanent basis willing to bring the action, the perpetrators of the infringement will not be at risk of having to compensate the full harm they caused. No other body would be able to bring an action. Furthermore, under the 'ideological claimant' model, no individual person has standing to bring an action on behalf of a class of similarly affected persons. The result may well be that no action is brought at all, which would impair both the deterrent and the compensatory function of private actions and, ultimately, the effectiveness of the EC competition rules.

Under the OFT's proposals, representative bodies which have not been designated on a permanent basis are not automatically prevented from bringing such an action. They would still be able to seek the court's permission to bring an opt-out action. However, while not defining the criteria for designation or permission, it is clear by the adoption of the "representative body" terminology that the OFT also recommends an "ideological claimant" model. Therefore, there is still a risk in individual cases that no representative body may be prepared to bring an action. Unlike under the class action model, an individual person would not have the standing to bring a class opt-out action.

In view of the potential limitations of the "ideological claimant" model, it can be argued that the class action model is better suited to achieving both objectives of private actions, namely compensation and deterrence. To the extent that the "ideological claimant" model is adopted as a safeguard against abusive litigation, it can be further argued that, if a permission stage

In view of the potential limitations of the "ideological claimant" model, it can be argued that the class action model is better suited to achieving both objectives of private actions, namely compensation and deterrence.

74. Recommendations *supra* note 5 at para. 7.33.

is to be part of any opt-out action, this should act as a sufficiently robust filter and render any further limitation on standing superfluous. However, it must be recognized that opt-out actions are still controversial in the EU Member States and there is still significant opposition to their introduction. The "ideological claimant" model would appear to be a reasonable compromise given the current political climate and the current level of experience in the EU with collective redress mechanisms. It is clear, however, that limiting standing to ideological claimants may restrict the availability of opt-out actions in some meritorious cases. There does not appear to be a need for an even more restrictive approach which would limit the ability to bring an action on behalf of identifiable victims or non-members only to representative bodies designated in advance and on a permanent basis. In this respect, the OFT's recommendations are more suited to furthering both the deterrent and the compensatory function of private actions.

It can also be noted that proposals on collective redress are now emanating from a number of sources.[75] In the long term, it may be that a consensus builds that the optimal model is an opt-out action with unrestricted standing of any member of a class of similarly affected persons. Provided that the sufficiency of robust judicial control to act as an effective safeguard against speculative claims at the permission stage is borne out by experience, this model appears to be the most suited to achieving the effective enforcement objective of private actions in EC competition law. The adoption of such a mechanism in the short to medium term is, however, unlikely.

VII. Proposals Relating to Indirect Purchasers' Standing and Passing-On Defense

In its Recommendations, the OFT stated that the issues of the passing-on defense and the standing of indirect purchasers would be best dealt with at the EC level. In particular, inconsistent treatment of the passing-on issue at the Member State level would undermine the effectiveness of damages actions regimes throughout the EU.

The White Paper makes two main proposals:

- The passing-on defense should be available to defendants to enable them to resist an overcharge compensation claim where the claimant passed on that overcharge to a subsequent purchaser. This would prevent the unjust enrichment of purchasers who passed on the overcharge and would avoid multiple compensations by the defendant. The burden of proof should be imposed on the defen-

75. In August 2008, the Civil Justice Council of England and Wales recommended the introduction of a general opt-out collective action: IMPROVING ACCESS TO JUSTICE THROUGH COLLECTIVE ACTIONS' : A SERIES OF RECOMMENDATIONS TO THE LORD CHANCELLOR, available on www.civiljusticecouncil.gov.uk/.

dant and the standard of proof should not be less than the standard imposed on the claimant to prove the loss.

- Indirect purchasers should be entitled to rely on a rebuttable presumption that the illegal overcharge was passed on to them in its entirety.

The Commission's proposals accord with the dual objective of increasing deterrence and ensuring compensation. Indirect purchasers would be entitled to rely on a presumption of passing-on that would both facilitate their claims and add to the deterrent effect of indirect purchasers' actions. The defendants' ability to rely on the passing-on defense is consistent with the principle that effective deterrence is achieved through full compensation and not multiple damages. However, giving the defendants the burden of proving passing-on should address, at least in part, concerns that allowing the passing-on defense may weaken the deterrent effect of private actions for breach of Articles 81 or 82.

Under the Commission's proposals, in the absence of a pan-European consolidation mechanism of private actions by direct and indirect purchasers, the risk of inconsistent judgments and multiple recovery of the same harm cannot be excluded. It is conceivable that, in an action brought by direct purchasers, the defendant is unable to prove that the overcharge has been passed-on to indirect purchasers. In a separate action, possibly in another Member State, indirect purchasers may rely on the presumption that the overcharge has been passed on to them. If the defendant is unable to rebut such a presumption, he may have to compensate the indirect purchasers for the same overcharge which had been already compensated in the former action. In terms of deterrence, this may be a risk worth taking if the counterfactual of placing on the direct purchasers the burden of proving a lack of passing-on or placing on indirect purchasers the burden of proving the actual passing-on of the overcharge from their sellers would undermine the deterrent effect of private actions in EC competition law.

The White Paper recognizes the limitations resulting from the lack of a European-wide consolidation mechanism, stating that

> "in the case of joint, parallel or consecutive actions brought by purchasers at different points in the distribution chain, national courts are encouraged to make full use of all mechanisms at their disposal under national, Community and international law in order to avoid under- and over-compensation of the harm caused by an infringement of competition law."[76]

76. White Paper on Damages Actions, *supra* note 7 at 2.6.

If there were an effective and widely-available method of consolidating cases, so that a defendant is likely to be facing only one action from both direct and indirect purchasers rather than multiple actions, there would be no need for the burden of proof and presumption of passing-on proposals. Once the overcharge had been proven, the defendant would be liable for damages arising out of that overcharge, but determining how much of the overcharge was passed on to various levels in the distribution chain would be for the various claimants to resolve in apportioning the damages. Concerns relating to multiple compensation claims would not arise. However, at this stage, the solutions put forward by the Commission in relation to indirect purchasers' standing and passing-on appear reasonable, given that:

- There is no evidence that allowing the passing-on defense and giving standing to indirect purchasers impair the deterrent effect of private actions. If there were such evidence, the issue should be reconsidered in light of the dual function of the right to damages under Article 81 or 82, which is not only to secure compensation but also to increase deterrence and ensure compliance, thus promoting social welfare and productivity in the long term;

- It seems unlikely that, in the short term, a pan-European mechanism for consolidation of cases brought by direct and indirect purchasers and apportionment of damages among direct and indirect purchasers can be introduced.

On both these issues, further research and work are needed before definitive answers to the complex questions of passing-on and indirect purchasers' standing can be given.

VIII. Proposals Relating to Leniency

Both the Commission[77] and the OFT[78] moved to safeguard the effectiveness of leniency programs from increased private litigation by proposing to exclude using leniency documents in civil litigation. This prevents the leniency applicant from being worse-off in civil litigation than a non-leniency applicant merely as a result of the leniency process[79] that requires the leniency applicant to produce written corporate statements and witness statements explaining in detail the functioning of the cartel and admitting its participation in the anti-competitive arrangements.

77. White Paper on Damages Actions, *supra* note 7 and Staff Working Paper on Damages *supra* note 7 at para. 287 – 302. This option was put forward in the Green Paper on Damages actions *supra* note 13, option 28.

78. Recommendations *supra* note 5 at para. 9.5.

79. Nazzini *supra* note 66 at Ch. 13.

Another set of proposals relates to the removal of joint and several liability for the leniency applicant (probably limited to the undertaking that receives full immunity) so that it is only liable for the harm caused to the direct and indirect purchasers of its own goods or services.[80] These proposals are intended to preserve the incentive to apply for leniency and to encourage a first leniency application by providing further benefits for the first applicant.

A. USE OF LENIENCY DOCUMENTS IN LITIGATION

The discoverability of leniency documents may increase the claimant's incentive to sue the leniency applicant as the primary or only target and place the leniency applicant at a disadvantage compared to the other cartelists. Such a disadvantage would not have occurred but for the leniency application and may leave a negative effect on the incentive to apply for leniency and the quality of the application. The obvious solution would be to exclude these documents from use in civil litigation without the consent of the leniency applicant.[81]

In order to assess this option in light of the dual function of actions for damages under EC law, the key question is whether this is a necessary and adequate measure to preserve the effectiveness of the leniency program and is in line with the principle of full effectiveness of Community law. This assessment requires a trade-off between the objective of increasing the deterrent effect of the EC competition rules and the compensatory dimension of the individual's right to damages. The trade-off is akin to a proportionality test. In our view, the objective of increasing deterrence should prevail but any interference with the protection of individual rights should be limited to what is necessary to achieve the prevailing objective. If leniency documents are defined as documents that would not have come into existence but for the leniency application, all other evidence and, in particular, any contemporaneous documentary evidence of the cartel remains available. Therefore, the unavailability of leniency documents for use in civil proceedings does not disproportionately restrict the individual right to damages and does not make its exercise impossible or excessively difficult.[82]

80. Green Paper on Damages actions *supra* note 13, options 29 – 30; Recommendations, *supra* note 5 at para. 9.9 – 9.10. In the United States, see the ANTITRUST CRIMINAL PENALTY ENHANCEMENT AND REFORM ACT OF 2004 15 USCA § 1 note.

81. White Paper on Damages actions *supra* note 7 at 10 and Staff Working Paper on Damages actions *supra* note 7 *at* para. 287 - 302; Private actions *supra* note 78 at para. 9.5.

82. Restrictions on the admissibility of evidence may raise a question of compatibility with Art 6 of the European Convention on Human Rights. As in relation to the right to access to court more generally, restrictions on disclosure or admissibility of evidence may be justified if necessary in the public interest provided that the party's right to a fair trial is not denied (see, for instance, Rowe and Davis v. United Kingdom, 30 EHRR 1 (2000)). In civil proceedings, it is unlikely that the inadmissibility of leniency documents, narrowly defined as those documents which would not exist but for the leniency application, might deny the claimant the right to a fair trial.

B. REMOVAL OF JOINT AND SEVERAL LIABILITY OF THE IMMUNITY RECIPIENT

As regards the removal of the immunity recipient's joint and several liability in damages, the Commission has been more cautious than the OFT. In the White Paper, the Commission does not propose this measure but simply puts it forward for further consideration. The OFT, on the other hand, recommended that the U.K. Government should consult on the option. The proposal aims at addressing the potential disincentive to a leniency application that the leniency applicant may be the only target of any damages action. If jointly and severally liable, the applicant would have to compensate the whole harm caused by the cartel. Depending on the applicable law, it may be able to recover from other cartelists their shares in contribution. By limiting the immunity recipient's liability only to the harm caused to those who directly or indirectly purchased goods or services from him, this disincentive would be removed. Furthermore, such a measure could further incentivize applications for leniency and, particularly if limited to the immunity recipient as the Commission and the OFT suggest, could incentivize the first application, thus increasing the destabilizing effect of leniency programs on the cartel.

The major objection to this proposal is that public enforcement objectives are limiting the rights of third parties. This is perceived as being 'unfair' or contrary to the compensatory function of private actions. However, this argument fades away if one recalls that, while private actions have a dual function under EC law, in the case of conflict the objective of increasing deterrence and ensuring the effective enforcement of the EC competition rules should prevail. Any limitations on the exercise of the right to damages that are necessary to achieve this objective are fully justified and consistent with the primary rationale for a right to damages: to increase deterrence and ensure compliance, thus increasing social welfare and productivity in the long term. It must be added, in line with the proportionality approach outlined above and consistent with the concurrent compensatory function of private actions, that third parties are not deprived of their private rights. Only joint and several liability is removed. Any party will be able to sue all cartelists jointly and severally except for the immunity recipient, who can only be sued by its direct and indirect purchasers. The measure in question only imposes limitations on the exercise of private rights to the extent that they are necessary to achieve the primary objective of increasing deterrence and enhancing the effectiveness of the regime as a whole.

IX. Conclusions

Private actions currently play a marginal role in competition enforcement in the European Union. This is particularly true of actions by consumers or small businesses. However, the ECJ has recognized that private actions can make a significant contribution to the effective enforcement of the EC competition rules. We have argued that they can do so in three ways. First, they can increase the resources available for the enforcement of competition law. Second, they can increase the detection rate of anti-competitive behavior. Finally, they can increase the magnitude of the financial consequences of an infringement. Private actions also have an important role to play in ensuring that those who have been harmed by competition law infringements are compensated. Increased deterrence and compensation almost always go hand in hand but the primary objective of private actions remains to contribute to effective competition enforcement, thus increasing social welfare and productivity in the long term.

Increased deterrence and compensation almost always go hand in hand but the primary objective of private actions remains to contribute to effective competition enforcement, thus increasing social welfare and productivity in the long term.

A well-functioning private actions regime should rest on the robust structuring and fine-tuning of three main pillars: collective actions, indirect purchasers' standing and passing-on, and the relationship between public and private enforcement. Both the Commission and the OFT have made proposals in these three areas.

In light of the dual role of private actions, which is to increase deterrence and ensure compensation, it may be argued that the collective actions proposals are unduly timid. The adoption of the ideological claimant model, in which only a "qualified entity" or "representative body" but not any member of an affected class, can bring an action on an opt-in or opt-out basis may lead to some meritorious cases not being brought. However, in the current political climate in Europe, still adverse to more effective collective redress, this model may be a realistic way forward. In our view, this model can work provided that two conditions are met. First, the criteria for the designation or authorization of the "qualified entity" or "representative body" should not be unduly restrictive. Second, any "qualified entity" or "representative body" should be given standing to bring a collective action on an opt-out basis, including by applying to the court for permission without any need for previous designation. Furthermore, it must be emphasized that the ideological claimant model is a significant safeguard against abusive litigation. This would justify a lighter-touch approach to any additional safeguards that the court may be required to apply or consider when permission to bring the action is sought.

The proposals relating to the standing of indirect purchasers and passing-on reflect the still incomplete understanding of this topic on both sides of the Atlantic. Given the lack of evidence that indirect purchasers' standing has a negative impact on the effective enforcement of the EC competition rules, it would be inappropriate at this stage to exclude or limit such standing. As a consequence, it also seems appropriate to allow the defendant to plead the 'passing-on' of overcharges as a defense (for which it carries the burden of proof). At the same time, it is appropriate to allow indirect purchasers to rely on a presumption of passing-on in order to facilitate their actions. Further thought, however, needs to be given to procedural mechanisms providing for the coordination or consolidation of direct and indirect purchasers' actions on an EU wide basis, although it must be recognized that such an EU wide procedural device may be very difficult to achieve in the short to medium term.

The proposals relating to leniency are fully consistent with the central role played by deterrence in the enforcement of EC competition law. The leniency program is of fundamental importance in the detection and prosecution of cartels. If the evidence shows that certain reforms of the private actions regime are likely to have a negative impact on the leniency program, the right to damages and its exercise may have to be limited to safeguard the effectiveness of the public enforcement process. In this regard, concerns relating to the disclosure of leniency documents may be addressed by excluding their use in civil litigation without the consent of the leniency applicant. It is also worth considering limiting the liability of the immunity recipient to the harm caused to the direct and indirect purchasers of its products or services.

In conclusion, the proposals currently on the table in the EU can be described as a cautious step in the right direction. Even if all these proposals were implemented in their most ambitious version, we would be unlikely to see the role of private enforcement develop to the levels experienced in the United States in terms of the number of cases, size of damages awarded, or settlements. However, reforms at the European and national levels are much needed. The current underdevelopment of private actions detracts from the achievable level of deterrence and compliance and is leaving uncompensated substantial unlawful transfers (in the order of several billions of Euros per year) from buyers to sellers. It is hoped that the Commission and the Member States will proceed swiftly to implement reform packages addressing the areas of collective actions, indirect purchasers' standing and passing-on,[83] and coordination between public and private enforcement in the ways explained above.

83. It is an open question whether the clarification which is needed in the area of indirect purchasers' standing and passing-on should come through a legislative intervention or be left to the jurisprudence of the courts. We recognize strong arguments both ways and, while a legislative solution would probably be superior in terms of achieving legal certainty and uniformity throughout the EU, there may be merit in observing case law developments in the Member States and, possibly, in the ECJ on a reference for a preliminary ruling, before any legislative reforms.

IX. Annex: Summary of the White Paper Proposals and the OFT's Recommendations to Her Majesty's Government

A. THE WHITE PAPER PROPOSALS

In 2008, more than two years after the publication of a Green Paper on Damages for breach of the EC antitrust rules, the Commission published a White Paper on the same subject.

The White paper proposals may be summarized in the following way:

- **Standing.** The Commission notes the need to foster collective actions and suggests that both "representative actions" and "opt-in collective actions" be made available to any individual who has suffered harm caused by an infringement of EC antirust laws. The Commission proposes that only entities designated on a standing basis should be able to bring an action on behalf of identifiable victims. Entities designated on an ad hoc basis, that is, for the purpose of a given action only, should only be able to bring an action on behalf of their members or some of their members.

- **Disclosure.** The Commission proposes that across the EU a minimum level of disclosure of evidence should be ensured, suggesting that, inter alia, national courts should have powers to order parties and third parties to disclose 'precise categories of relevant evidence', subject to certain conditions to avoid overly broad and burdensome disclosure obligations. The Commission also proposes that national courts should have the power to impose sanctions for either destruction of relevant evidence or refusal to disclose such evidence.

- **Binding effect of decisions.** Final decisions by NCAs finding a breach of Articles 81 or 82 should be binding on national courts. Private parties may rely on them as a basis for a follow-on action. Currently, only decisions of the Commission are binding under Community law. In the United Kingdom, decisions of the OFT and the concurrent regulators are binding on the courts under the Competition Act 1998.

- **Fault requirement.** Member States' laws differ as to whether, in addition to establishing a breach of the competition laws, fault must be separately established to sustain a damage claim. The Commission proposes that, in Member States that require fault to be proven, once the victim has shown a breach of Article 81 or 82, the infringer should be liable for damages caused unless he demonstrates that the infringement was the result of a genuinely excusable error. This would not appear to change the position in England

and Wales, where tortious liability for breach of statutory duty does not require the claimant to prove the defendant's fault.

- **Damages.** The Commission suggests full compensation but not multiple damages.

- **Passing-on overcharges.** The Commission proposes that defendants should be entitled to raise the passing-on defense against a claim for compensation of the overcharge. The burden of proof should be on the defendant. The standard of proof should be the same as that which the claimant must meet. The Commission further suggests that indirect purchasers should be able to rely on a rebuttable presumption that the illegal overcharge was passed on to them in its entirety. Consolidation mechanisms are encouraged.

- **Limitation periods.** The Commission suggests that the limitation period should not start to run before the day on which the infringement ceases (for continuous or repeated infringements) and/or before the victim of the infringement can reasonably be expected to have knowledge of the infringement and of the harm it caused him. The Commission further suggests that at least two years be allowed for the commencement of a private action after the infringement decision on which the claimant relies has become final (i.e., after all court appeals of agency decisions have been exhausted). This is consistent with the current position under the U.K. Competition Act 1998.

- **Costs of damages actions.** The Commission faces a variety of cost allocation rules among the Member States, most of whom apply the "loser pays" principle. It would appear that the Commission is not proposing any binding Community measure in this area (the language used in the White Paper is: "… it would be useful for Member States to reflect on their cost rules …"). However, the Commission suggests that Member States adopt measures to foster settlements, set court fees so that they do not become a disproportionate disincentive to competition damage claims, and allow courts to issue cost allocation orders that derogate from the normal cost rules, that is, from the "loser pays principle."

- **Interaction between leniency programs and private actions.** The Commission proposes that corporate statements submitted by a leniency applicant should be protected against disclosure regardless of whether the leniency application is accepted, rejected, or leads to no decision by an agency. The Commission puts forward for further consideration the possibility of limiting the civil liability of the immunity recipient to claims by his direct and indirect contractual partners.

B. OFT'S RECOMMENDATIONS TO HER MAJESTY'S GOVERNMENT

In November 2007, the OFT published a set of Recommendations to Her Majesty's Government on Private Actions in Competition Law: Effective Redress for Consumers and Businesses. Following a public consultation, the OFT recommended:

- Allowing representative bodies to bring stand-alone and follow-on representative actions for damages and applications for injunctions on behalf of named consumers and businesses or on behalf of consumers and businesses at large.
- Introducing conditional fee agreements in representative actions which allow for an increase of greater than 100 percent on lawyers' fees.
- Codifying courts' discretion to cap parties' costs liabilities and to provide for the courts' discretion to give the claimant cost-protection in appropriate cases.
- Establishing a merits-based litigation fund.
- Requiring U.K. courts and tribunals to "have regard" to U.K. NCAs' decisions and guidance.
- Conferring a power on the Secretary of State to exclude leniency documents, appropriately defined, from use in litigation without the consent of the leniency applicant,
- Conferring a power on the Secretary of State to remove joint and several liability for immunity recipients in private actions in competition law so that they are only liable for the harm they caused (or not liable at all in exceptional circumstances).

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Alive and Clicking: Collusion Theories in Merger Analysis at the Federal Trade Commission

Malcolm B. Coate

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Alive and Kicking: Collusion Theories in Merger Analysis at the Federal Trade Commission

*Malcolm B. Coate**

This paper explores the use of collusion theories in merger analysis at the U.S. Federal Trade Commission ("FTC"). The 1992 *Merger Guidelines* ("Guidelines") focused more on unilateral effects concerns, relegating collusion analysis to a second-tier theory. That said, both structural and behavioral conditions conducive to establishing or maintaining an arrangement to restrict competition were listed in the *Guidelines* to structure collusion analysis. This paper undertakes a systematic review of 75 merger decisions to identify the conditions that increase the likelihood of a collusion finding. Standard structural concerns are readily identified, while behavioral factors defy characterization. The results of the analysis also support a Folk Theorem in which structural concerns are validated with some type of performance evidence. Further work finds that allegations of maverick conduct add little to the analysis, while the Bush administration appears to have been slightly more likely to identify a collusion problem than the Clinton administration.

*Malcolm Coate is an economist with the Federal Trade Commission. The analyses and conclusions set forth in this paper are those of the author and do not necessarily represent the views of the U.S. Federal Trade Commission. The author would like to thank Jeff Fischer and Seth Sacher for helpful comments.

I. Introduction

The 1992 revision of the *Merger Guidelines*[1] accepted the burden to move beyond a structural checklist and tell a logical story that links a competitive effect of concern to the consummation of a proposed merger.[2] Two lines of analysis were given, one based on a traditional collusion theory (re-branded as "coordinated interaction") and the other tied to anticompetitive activity that the merged firm could undertake independently ("unilateral effects"). This evolution was inevitable in light of the continual decline in the breadth of the *Philadelphia National Bank* ("PNB") structural presumption.[3] In 1990, the appellate court in *Baker Hughes* concluded that the PNB presumption could be offset with evidence on a wide range of pro-competitive considerations.[4] Once the respondent presented some evidence compatible with a pro-competitive outcome for the merger, the plaintiff had to prove a likely competitive concern stemming from the merger. Thus, to prevail on the merits, the plaintiff needed evidence. To structure this evidence, it needed a story.

A careful review of *Guidelines*-based enforcement would conclude that the new unilateral effects theory defined the enforcer's leading story.[5] For a unilateral effects theory, the government only had to introduce evidence on a unique similarity for the merged firm's products, given a limited number of rivals. If this information was lacking, simple market share evidence could establish a presumption. In effect, unilateral effects might end up as nothing more than a structuralist model underpinned with a veneer of economic authority.[6] Collusion

A CAREFUL REVIEW OF *GUIDELINES*-BASED ENFORCEMENT WOULD CONCLUDE THAT THE NEW UNILATERAL EFFECTS THEORY DEFINED THE ENFORCER'S LEADING STORY.

1. U.S. Department of Justice and Federal Trade Commission, *Horizontal Merger Guidelines*, Antitrust Trade Reg. Report 1559 (1992).

2. Paul Denis, *Advances of the 1992 Horizontal Merger Guidelines in the Analysis of Competitive Effect*, 38(3) Antitrust Bull. 479-515 (1993).

3. U.S. v. Philadelphia National Bank, 374 U.S. 321 (1963). For a useful commentary on the decline of the Philadelphia National Bank presumption, *see* Jonathan Baker, *Mavericks, Mergers, and Exclusion: Proving Coordinated Effects Under the Antitrust Laws*, 77 NYU L. Rev. 135-201 (2002).

4. U.S. v. Baker Hughes Inc., 731 F.Supp. 3 (D.D.C. 1990), aff'd 908 F.2d 981 (D.C. Cir. 1990).

5. *See, e.g.*, Robert Lande & James Langenfeld, *The Evolution of Federal Merger Policy*, 11(2) Antitrust 5-9 (1997); or Jonathan Baker, *Why Did the Antitrust Agencies Embrace Unilateral Effects?* 12(1) Geo. Mason L. Rev. 31-37 (2003).

6. *See, e.g.*, Robert D. Willig, *Merger Analysis, Industrial Organization Theory and Merger Guidelines*, *in* Brookings Papers: Microeconomics 281-332 (Brookings Institution, 1991); or Gregory J. Werden & Luke M. Froeb, *The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy*, 10(2) J.L. Econ. & Org. 407-26 (2004). These models link structure to performance by assuming away a wide range of real-world complications. The models may be appropriate in special

analysis was demoted to a second-tier theory, as that analysis was generally qualitative, often discussing factors summarized in Posner (1976) and (2001).[7] While the maverick model had long been mentioned as a viable empirical structure for the coordinated interaction analysis, the bulk of the *Guidelines* focused on generic models of coordination. Without a clear model of collusion, it was hard to know when one had "enough" evidence.

With roughly 15 years of experience under the 1992 *Merger Guidelines*, it is possible to study the implementation of the coordinated interaction policy with a systematic review of the relevant enforcement decisions undertaken for a set of mergers filed between 1993 and 2005. By limiting the study to coordinated interaction cases, it is possible to build on a 2007 paper by Davies, Olczak, and Coles, and use the enforcement decisions to create a model of coordinated interaction.[8] As long as the investigations with easy entry are deleted from the sample, the enforcement decision and the conclusion on the ease of collusion are the same. A number of structural variables are readily available to build a Posnerian model of collusion. Moreover, the structural model can be supplemented with a performance-based effects variable to test the importance of evidence related to various explicit or implicit natural experiments. If the natural experiment evidence matters, then theoretical analysis of ease of collusion may take a back seat to natural experiments in predicting merger effects. Finally, the model can be expanded to (1) determine if claims of maverick status are relevant and (2) identify any political influence on the analysis.

Overall, the results are broadly compatible with Posner's structural theory and support the importance of natural experiment evidence. A small positive effect for the Bush administration is identified, but no pure maverick effect appears to exist. In court, natural experiment evidence should assist a plaintiff in meeting its burden of proof, a task that has been difficult in recent unilateral effects cases.[9]

Section II of this paper provides an introduction to early oligopoly (collusion) theory, with a specific focus on Stigler's contribution. It also introduces concepts from modern game theory. Section III presents an overview of the impact of economics on the *Merger Guidelines*. A review of the FTC's case files highlights the

footnote 6 cont'd
case situations, but fall short of general models of the competitive process. For example, the models require the producing firm to post price, rather than negotiate on terms of trade. Likewise, the models presume the firms are locked into a specific product portfolio. For a critical overview, *see* Malcolm B. Coate, *Efficiencies in Merger Analysis: An Institutionalist View*, 13 GEO. MASON SUP. CT. ECON. REV. 189-240 (2005).

7. RICHARD A. POSNER, ANTITRUST LAW: AN ECONOMIC PERSPECTIVE (1976) (2nd ed. 2001).

8. Stephen Davies et al., Tacit Collusion, Firm Asymmetries and Numbers: Evidence from EC Merger Cases (unpublished manuscript) (2007).

9. *See, e.g.*, U.S. v. Sungard 172 F. Supp. 2d 172 (D.D.C. 2001); or U.S. v. Oracle, 331 F. Supp. 2d 1098 (N.D. Cal. 2004).

role of natural experiment evidence in the enforcement process. The basic modeling is presented in section IV, with the statistical results in section V. Section VI concludes.

II. Economics of Oligopolistic Collusion

Oligopoly has a long history, predating the formalization of economics. Schumpeter traced the term "oligopoly" to Saint Thomas Moore's 1516 book, *Utopia*.[10] Cournot structured the oligopoly concept by postulating firms simply assumed their rivals would hold output fixed, regardless of what the firm in question did.[11] This assumption allowed market equilibrium to be computed at a price between the monopoly and competitive level. Chamberlin linked the oligopoly market equilibrium to "recognized mutual dependence."[12] After first touching on the Cournot and Bertrand structures, Chamberlin posited oligopolists would assume their actions affect the responses of their rivals and that they take that conclusion into account in setting price. When this interdependence was completely recognized, profit-maximization behavior generated a monopoly outcome.[13] Uncertainty could generate less perfect recognition and thus lower prices, although this tacit coordination would generally allow oligopolistic (collusive) firms to raise price well above the competitive level. Market concentration, on its own, seemed problematic. Stigler's model of collusion showed how market interactions were really much more complex, with a wide range of factors affecting the likelihood of collusion. Modern game theory formalized the basic Stigler insights.[14] These developments are discussed below.

Stigler's 1964 analysis represented a huge innovation in collusion (oligopoly) theory as the model detailed various conditions that made interdependent pricing more or less likely.

A. CHICAGO ANALYSIS

Stigler's 1964 analysis represented a huge innovation in collusion (oligopoly) theory as the model detailed various conditions that made interdependent pricing

10. Joseph Schumpeter, History of Economic Analysis (1954); and Thomas Moore, Utopia (1516). Thomas Moore was sainted as he died a martyr for the Catholic Church in 1535. Reading a few pages of the English translation of the book, *Utopia*, clearly suggests that the book has little to do with neoclassical economic theory. Adam Smith's position as the first modern economist appears secure.

11. Augustin Cournot, Researches into the Mathematical Principles of the Theory of Wealth (1838) (Eng. trans. 1971).

12. Edward Chamberlin, The Theory of Monopolistic Competition (1933).

13. *Id.* at 54.

14. George Stigler, *A Theory of Oligopoly*, 72(1) J. Pol. Econ. 44-61 (1964).

more or less likely. After making the initial point that product homogeneity plays an important role in the development of a collusive pricing scheme, Stigler advanced the idea that cartel participants can track changes in sales patterns to detect (and thus deter) competitive pricing in the marketplace. Stigler's model suggests that detection of competitive conduct is easiest when information on prices and sales is readily available. It is also possible to infer competitive conduct from the totality of the evidence. Such an inference is more likely when the number of buyers served by each competitor is relatively large (numerous customers switching leads to inference of discount pricing, even though little market share is lost), the market is relatively stable (buyers grow or shrink slowly, so they are less likely to switch suppliers for reasons unrelated to discounts), and the industry is relatively static (few new buyers exist to disrupt historical business relationships). If competitive conduct is readily identified, it is less likely to occur.

Stigler's theory clearly identified the two considerations associated with coordinated interaction concerns. First, the incumbents must be able to converge to a joint course of conduct to elevate price above the competitive level. Second, the incumbents require a mechanism to detect (and then punish) deviations from the arrangement to increase the probability that all participants abide by the chosen course of conduct. Understanding the structure of the firms, the fundamentals of the market transaction, and the information available to competitors is shown to be necessary to model the ease of collusion. Finally, the model retains a dynamic flavor as conditions that upset the collusive equilibrium (e.g., entry, growth, and innovation) are thought to make persistence of non-competitive pricing less likely. These core ideas are repeated in more modern characterizations of the collusion problem.

B. POST-CHICAGO GAME THEORETIC ANALYSIS

Game theory offers a mathematical characterization for the oligopolistic interactions among firms.[15] Before providing an overview, it is necessary to introduce the models through which game theory represents the competitive process. The standard competitive baseline for both homogenous and differentiated goods is the one-shot Nash-Bertrand price-setting game. In a homogeneous market, with comparable cost conditions, the perfectly competitive equilibrium is generated as firms simply cannot raise price above the marginal cost, while in a differentiated goods market, firms unilaterally price above marginal costs to cover the fixed costs associated with the differentiation.

GAME THEORY OFFERS A MATHEMATICAL CHARACTERIZATION FOR THE OLIGOPOLISTIC INTERACTIONS AMONG FIRMS.

15. Carl Shapiro, *Theories of Oligopolistic Behavior, in* HANDBOOK OF INDUSTRIAL ORGANIZATION (Schmalensee & Willig, eds. 1989); and DENNIS CARLTON & JEFFREY PERLOFF, MODERN INDUSTRIAL ORGANIZATION (2000).

Supra-normal profits are eliminated by entry.[16] The Cournot regime represents a secondary structure relevant to situations in which the firms (with homogenous goods) compete by setting output levels. Under these conditions, firms would restrict output below the competitive level and force price up. Again, entry could eliminate the supra-normal profits.[17]

The possibility for collusion is usually modeled through the use of punishment strategies integrated into infinitely repeated games ("supergames") in which oligopolists compete in period after period. Technically, the strategies remain non-cooperative as each firm unilaterally chooses to implement the punishment tactics. However, the interactive process implicit in the supergame represents almost a textbook characterization of tacit collusion (or mutual dependence recognized), so economists cannot help but characterize these non-cooperative games as collusive. Supergames allow for an infinite number of equilibriums, and lead to the Folk Theorem of Oligopoly: collusive equilibriums are sustainable for some set of parameters. Of course, competitive equilibriums are also possible, leading to another characterization of the Folk Theorem: competitive equilibriums are sustainable for some set of parameters. While economists could add complexities to try to eliminate the plethora of equilibriums, the problem remains that games could also be restructured to generate any desired theoretical equilibrium.[18] Game theory illustrates just how difficult collusion is to prove.

GAME THEORY ILLUSTRATES JUST HOW DIFFICULT COLLUSION IS TO PROVE.

Game theoretical analysis remains useful, because the models highlight the discount rate that links the periods of the supergame together. Minimal discounting of the next period's returns (which implies rapid reactions to precompetitive price reductions) makes less than competitive equilibriums more viable. In

16. Mergers that fall short of monopoly may have no effect for the homogeneous goods market, but may allow a material price increase for a differentiated good. This non-cooperative price increase for differentiated goods markets represents the core of the unilateral effects concern introduced in the 1992 *Merger Guidelines*, although more detailed analysis must ensure (1) the model actually represents reality, (2) the price effect is material, and (3) repositioning of other rivals is unable to offset the price increase.

17. Mergers may generate price effects in Cournot games, although the model would rarely be useful, as most firms set price and not output. Without some institutional restriction on output expansion to match a rival's action, the Cournot structure is not viable. The Cournot game may be more useful as a collusion model, assuming some exogenous agreement on the "rules of the game" has created the artificial incentive for firms to hold output fixed. Given an agreement to fix output via the Cournot structure, a merger would tend to make the market less competitive. *See* Malcolm Coate & Mark Williams, *Generalized Critical Loss for Market Definition*, 22 RESEARCH IN L. & ECON. 41-58, note 4 (2007).

18. For an analysis of the problems with real world applications of Post-Chicago economics, *see* Malcolm Coate & Jeffrey Fischer, *Daubert,* Science, and Modern Game Theory: Implications for Merger Analysis (unpublished manuscript) (September 2008), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1268386.

effect, the ability of the collusive suppliers to respond quickly to competition means precompetitive behavior is less likely. While it is well understood that this speed of adjustment is related to the ability to quickly detect competitive conduct in a "spot" market, it is less obvious how to model speed of detection when customer-supplier commitments are relevant. If a firm can establish a long-term relationship with a large customer by cheating on a cartel, then it may be impossible for its rivals to respond quickly even if the competitive conduct can be detected immediately. In effect, cheating on the collusion may allow the independent firm to lock up new business for a long period of time (this implies the need to use a high discount rate in the mathematical model). Understanding how market processes work should enable an analysis of the extent of vertical customer-supplier relationships. While game theory leaves a role for structural checklists, it significantly increases the level of detail required to undertake competitive analysis.

III. Application of Economics to Merger Analysis

Over the years, economic theory has generated a number of insights for the merger review process. The 1968 *Guidelines* focused enforcement on very small changes in market share, but noted that a more detailed analysis should be undertaken when share appeared to be a poor predictor of competitive effect. More aggressive enforcement was warranted when the target firm was likely to be a disruptive force in the market (this "disruptive force" concept was later remarketed as the "maverick" firm). The 1982 *Guidelines* added a set of "other factors" relevant to oligopoly analysis.[19] Structure, conduct, and performance considerations were all mentioned and the discussion generally tracked Posner's oligopoly checklist. The 1992 *Guidelines* presented a more complex economic analysis that separated the discussion associated with reaching an agreement from the commentary on policing an agreement. The importance of a sophisticated understanding of information structures, along with knowledge of the basic institutional mechanisms of a market, was also stressed. However, the importance of performance evidence was limited to a comment on explicit price-fixing and a couple of footnotes.[20]

As the foundation for the 1982 *Guidelines*, Posner's checklist is addressed in sub-section A, while the 1992 *Guidelines* material is discussed in more detail in sub-section B. A final sub-section, which focuses on performance evidence, is

19. *See* U.S. Department of Justice, *Merger Guidelines*, Antitrust Trade Reg. Report 1069 (1982); and U.S. Department of Justice, *Merger Guidelines*, 2 Trade Reg. Report (CHH) ¶ 4510 (1982).

20. The 1992 *Merger Guidelines* note that market conditions are likely to be conducive to coordinated interaction when firms in the market have (1) engaged in express collusion and (2) salient characteristics of the market have not changed. Implicit performance evidence may also be addressed. *See*, *e.g.*, § 2.1 (focuses on consumer harm) and note 22 (mentions the use of normal course of business documents) of the *Guidelines*.

included to introduce the "Folk Theorem of Merger Enforcement." This concept, implicit in the staff applications of the *Guidelines*, suggests that structural collusion models should be tested with exogenous evidence.

A. THE 1982 *GUIDELINES* AND POSNER'S CHECKLIST

Richard Posner created a classic checklist of characteristics associated with oligopolistic interdependence as part of his ambitious attempt to expand the reach of the antitrust laws to encompass tacit collusion.[21] These structural conditions are listed below.[22]

- High market share: The Herfindahl statistic (defined by the sum of the square of the market shares held by the firms in the market) is a generally accepted proxy for impact of market share on the probability of less than competitive conduct. While higher values for the Herfindahl statistic tend to increase the likelihood and duration of competitive problems, the magnitude of the effect must be evaluated on the basis of industry-specific evidence. High values of the Herfindahl are correlated with relatively few significant competitors (firms required to participate in the cartel), but the Herfindahl is able to proxy the relative size of the firms.

- No fringe: Fringe firms are price takers and thus unlikely to participate in any arrangements to raise price. Collusion is more likely to evolve or persist, the smaller the fringe (and the more limited its ability to expand output).

- Inelastic demand at competitive price: The market elasticity measures the loss of sales associated with customers substituting away from the market. If a price increase leads to a small reduction in output (inelastic demand), then the significant firms need only to reduce their production slightly to force price up. Thus, coordinated interaction is more likely to occur.

21. In contrast to most scholars who consider pure tacit collusion to be legal (due to the lack of an agreement), Posner looked at the effects of the tacit collusion (usually higher prices) and found price-fixing. He proposed a more aggressive style of analysis in which the pricing in a less than competitive manner would be illegal. *See* Posner (1976), *supra* note 7. To promote this outcome, Posner listed a number of factors that make markets relatively more susceptible to tacit collusion and introduced conduct and performance factors that were potentially compatible with collusion. While this style of price-fixing analysis never had much support, it became the standard for merger analysis in the 1980s.

22. Two conditions (local markets and cooperative practices) are not listed here, because they seem more related to reaching an illegal price-fixing agreement than coordinated interaction. Posner also added one characteristic of conduct (antitrust record) to his list of conditions favorable to collusion and four examples of problematic conduct (exchange of price information, industry-wide resale price maintenance, base point pricing, and exclusionary conduct) as examples of economic evidence of less than competitive behavior. These conduct considerations were also mentioned in the 1982 *Guidelines*, along with the 1968 *Guidelines'* concept of the disruptive firm. Finally, Posner's analysis included a list of performance conditions suggestive of less than competitive behavior. Evidence associated with current less than competitive performance is useful to test the implications of the structural analysis.

- Entry takes a long time: Collusion to restrict competition is more likely to evolve and persist when entry takes a long time, as the potential returns to less than competitive behavior are higher.

- Buying side of market is unconcentrated: Arrangements to restrict competition are more likely to persist when the buyer side of the market is relatively atomistic. Large buyers may be able to threaten the stability of a cartel by shifting significant purchases to suppliers willing to price competitively.

- Standardized product: Firms generally find it easier to agree on the terms of coordination and ensure all significant rivals participate in the arrangement when the product is standardized. Also, standardization makes it easier to detect deviations from any collusive agreement.

- Non-durable product: Non-durable goods are not relevant for market competition in future periods, because customers cannot invest in maintenance to extend the life of the good. Thus, markets with non-durable goods are more likely to suffer from collusion than markets with durable goods.

- Principal firms sell at same level of distribution: Coordination interaction is simpler to establish and maintain when it is operationalized at one vertical level.

- Similar cost structures and production processes: Some form of collusion is more likely to evolve and persist when all the significant competitors share the same cost structure and technology.

- Demand is static or declining: Firms are more likely to sustain a policy of less than competitive behavior when the market is static or declining, because the oligopolists do not have to deal with a constant flow of new customers and products into the market.

- Prices can change quickly: The ability to adjust price in a timely manner makes punishment strategies more effective and hence tends to make coordinated interaction more successful.

- Sealed bidding: The use of sealed bidding makes it easier to identify competitive pricing, as the winning bid must be published. As secret price discounts are impossible, collusion is more likely to be sustained over time.[23]

The fundamental problem with a structural Posnerian merger analysis is the lack of a general theoretical analysis to facilitate the evaluation of the relevant factors. Empirical information on the structural factors can be tabulated and

23. The 1982 *Guidelines* generalize this point to focus on the supplier's ability to obtain detailed information on prices, outputs, or specific transactions. Sealed bidding is simply one example in which good information is available.

some broad-based observations derived, but balancing the impact of the factors to evaluate the likely competitive effect of a merger is purely subjective.[24] Moreover, it is unclear how to work the structural effect of the merger into the analysis. As Scheffman and Coleman note, checklists "are too crude to provide much assistance in determining whether a coordinated interaction theory is relevant."[25]

In its defense, Posner's collusion presentation moves beyond structure and includes market performance evidence. Generalizing Posner's price-fixing analysis to address merger enforcement would therefore trigger a search for performance evidence compatible with the structural competitive concern. Hence, a complete Posnerian study of a merger in an oligopolistic industry could generate useful results, as the implications of the structural analysis would be validated with performance evidence.

B. THE 1992 *MERGER GUIDELINES* CHARACTERIZATION

The 1992 *Merger Guidelines* address the limitation of the structural analysis first by sub-dividing the coordinated interaction issue into its constituent parts (predicting (1) whether a post-merger arrangement is likely to evolve and (2) if that arrangement is likely to persist), and then by insisting that the analysis provides an explanation of how prices could be elevated above the competitive level. By concentrating on the need to tell a story, the revised *Guidelines* are better able to focus the analysis on the relevant informational and institutional structures in the market. While informational issues underlie a number of the Posnerian conditions, the *Guidelines* stress the importance of information as a stand-alone structural characteristic. The institutional details of the competitive process within a market must also be evaluated to determine if post-merger collusion is likely. For example, generic information on the business conditions facing rivals may increase the probability of some form of agreement, while the availability of information on specific transactions or individual prices and output levels may make the detection of price discounting more likely, all else equal. On the other hand, customer-supplier relationships might moot the importance of information, because once the customer switches, the new vertical relationship is established. This relationship may possibly be immune to short-run offers of discounts.

24. The standard Posnerian checklist could be used to conclude a merger is not likely to enhance collusive pricing if the review identified few factors suggestive of concern. Alternatively, if enough assumptions are made, then a mathematical model could be parameterized and used to estimate the price effects of a merger. *See, e.g.*, Janusz Ordover et al., *Herfindahl Concentration and Mergers*, 95 Harv. L. Rev. 1857-73 (1982).

25. David Scheffman & Mary Coleman, *Quantitative Analysis of Potential Competitive Effects from a Merger*, 12(2) Geo. Mason L. Rev. 319-69, 327 (2003).

The basic Posnerian considerations generally point in the same direction for the two 1992 *Guidelines* questions.[26] For example, high market share is considered to make an agreement on price more likely to occur, because fewer firms need to be involved in the understanding. Likewise, an agreement is easier to monitor and police when it is only necessary to follow the actions of a few competitors. Arguments can also be made that certain conditions support one oligopoly task, while making the other less likely. Either effect could dominate, given specific market conditions. For example, Stigler's model shows agreements are easier to monitor in a static market. However, a more detailed collusion model could suggest price agreements are less likely to form in static markets, because the potential profit from collusion is lower.

However, a more detailed collusion model could suggest price agreements are less likely to form in static markets, because the potential profit from collusion is lower.

A few of Posner's factors are generalized by the *Guidelines*. Posner considers collusion more likely when demand is static or declining. The *Guidelines* expand this concept to address any dynamic change in the market. Maintaining a collusive agreement is simply more difficult when market conditions (e.g., demand curves, cost conditions, or innovation) are changing rapidly. Second, Posner observes collusion is more likely when prices can change rapidly. The *Guidelines* also generalize this issue to focus on the characteristics of the typical transaction. The speed associated with changing any detail in the representative transaction could also affect the ease of collusion.

While the *Guidelines*' structure offers insight into the issue of collusion, it must link the analytical structure to the merger in question to be useful for antitrust policy.[27] In the overview to the coordinated interaction section, the *Guidelines* state: "A merger may diminish competition by enabling the firms selling in the relevant market more likely, more successfully, or more completely to engage in coordinated interaction that harms consumers."[28] The term "more likely" implies

26. Three of Posner's characteristics are not explicitly mentioned in the *Guidelines*' competitive effects section, but remain relevant to the overall analysis. Inelasticity of demand is not noted, although the factor is important in the market definition analysis. Likewise, the relevance of ease of entry for the likelihood of coordinated interaction is not explained, although it is obviously covered in the entry section. Finally, the durable nature of the good is not highlighted as relevant to collusion, but is touched on in the entry discussion.

27. In a recent commentary, Dick (2003) suggested that the *Guidelines'* analysis should focus on two questions to address this concern. First, what constrains the suppliers' pre-merger incentive or ability to coordinate their actions? In effect, the merger review must discover what drives the pre-merger pricing decisions. This understanding leads to the second question: How will the proposed merger change the existing constraints on competition? This approach would also link the *Guidelines'* analysis to a specific economic theory. *See* Andrew Dick, *Coordinated Interaction: Pre-merger Constraints and Post-merger Effects*, 12(1) Geo. Mason L. Rev. 65-88 (2003).

28. *See* 1992 *Merger Guidelines*, at § 2.1.

the merger causes some type of regime shift in which the merger changes the market from competitive to collusive (the maverick model explicitly mentioned in the *Guidelines*' text is simply an example of a regime-shift model) and the phrase "more completely" suggests that some type of structuralist model is relevant (as the market is currently less than competitive, and the merger worsens the situation). "More successfully" implies some effect on the durability of the coordinated interaction process (regime shift becomes more likely to persist given the fixed probability it will occur or the structuralist effect becomes more long-lasting).

Coate and Ulrick discuss three styles of coordinated interaction analysis (maverick, general regime shift, and structuralism) that are found in FTC staff analysis.[29] Maverick analysis applies when one of the merging parties has a relatively unique and significant incentive to deviate from the terms of the collusive consensus. Under certain conditions, the maverick firm ensures that the market remains competitive and its loss through merger leads to some form of collusion. In the standard maverick theory, facts are used to identify and prove the specific characteristics of the maverick and its loss is then considered likely to trigger collusion.[30] Mathematical precision is possible if the compatible oligopolists are considered to act in a less than competitive manner, but coordinated pricing is not profitable in light of the competition from the maverick firm (in combination with the fringe entities). Parameterizing a complex model should allow the calculation of a competitive equilibrium.[31] Then the analyst could adjust the model to transform the maverick from an independent competitor to a cartel participant and compute the merger-related price increase. Baker (2002) and Baker and Shapiro (2007) appear to advance the maverick model as the only relevant model of collusion for antitrust policy.[32]

29. *See* Malcolm Coate & Shawn Ulrick, *Influential Evidence in Merger Enforcement at the Federal Trade Commission*, *in* Issues in Competition Law & Policy (W.D. Collins, ed. 2008). The three theories are also mentioned in Malcolm Coate, *Economic Models in Merger Analysis: A Case Study of the Merger Guidelines*, 2(1) Rev. L. & Econ. 53-84 (2006). Moreover, Scheffman & Coleman (2003), *supra* note 25, at 328-29, without benefit of the systematic review of the files, define three comparable theories (evidence of existing effective coordinated interaction, removal of a maverick, and removal of other impediments to coordination).

30. Baker & Shapiro (2007) also posit a more generic maverick model to be applied when specific facts are not available. This model simply assumes that the merger partners have a significant probability of being the maverick when the number of significant competitors is small. Thus, the merger is likely to lead to the loss of this hypothetical maverick. This second "theory" simply appeals to structuralism, as specific facts supporting the maverick hypothesis are not required. *See* Jonathan Baker & Carl Shapiro, Reinvigorating *Horizontal Merger Enforcement* (unpublished manuscript) (June 2007), *available at* http://faculty.haas.berkeley.edu/shapiro/mergerpolicy.pdf

31. The analysis could start with the classic Landes and Posner (1981) model and generalize the monopolist to represent the set of collusive firms and the fringe to include the maverick. The model would be calibrated to generate a competitive equilibrium when the maverick prices as a fringe firm. *See* William Landes & Richard Posner, *Market Power in Antitrust Cases*, 94(5) Harv. L. Rev. 937-96 (1981).

32. *See* Baker (2002), *supra* note 3; and Jonathan Baker & Carl Shapiro, *supra* note 30.

The staff also considers a general "regime-shift" model in which the pre-merger structure gives rise to a competitive outcome, while the post-merger structure is conducive to some form of collusion. Modeling the regime shift is limited only by the imagination of the merger analyst. For example, mergers that created or enhanced the power of a leading firm could be problematic as price leadership could facilitate both establishing and policing an agreement. Moreover, the concern with the leadership would be heightened if the leader actively supported conduct (e.g., product standardization or information exchanges) that appeared to reduce the costs of coordinated interaction. Mergers may also create more symmetry in the market. When few rivals exist, increasing the product, marketing, distribution, and cost symmetries in the market could make a collusive equilibrium more likely. Concerns would be enhanced when market structures ensure that rivals understand their mutual interdependence. More generic models of competition could also give rise to collusive concerns when the number of firms in the market is reduced. These regime shift models can also be quantified by modeling the shift from competition to some form of collusion.[33]

Structuralism is also anticipated in the *Guidelines*, as the pre-merger structure could support a small collusive surcharge, while the post-merger structure could enable a higher collusive price. Structural evidence may suggest that at least weak coordinated interaction is likely pre-merger. Post-merger, the structure will become much more compatible with less than competitive behavior and the market price may rise. While theorists consider Cournot to represent unilateral behavior, the single shot Cournot game could be used to give a mathematical veneer to a structural coordinated interaction model, as prices rise with a reduction in the number of independent competitors, holding costs constant. The analyst need only parameterize the model for the current market conditions and compute the effect associated with deleting a rival. Interested analysts could generalize the simple model to allow for fringe competition, differentiation, or cost asymmetries.

BUILDING A THEORETICAL FOUNDATION FOR COORDINATED EFFECTS CONCERNS APPEARS TO MOVE THE ANALYSIS BEYOND THE CHECKLIST STAGE AND CREATES A ROAD MAP FOR ECONOMIC ANALYSIS.

Building a theoretical foundation for coordinated effects concerns appears to move the analysis beyond the checklist stage and creates a road map for economic analysis. The analyst can review the record of the investigation and collect evidence to parameterize the relevant theory in much the same way as unilateral effects analysis. Scheffman and Coleman detail a number of studies that can be undertaken as part of the analysis.[34] Facts must determine which of the many

33. Fisher et al. (1988), for example, used a Cournot structure for the collusive regime, while a Bertrand structure illustrated the pre-merger competition. *See* Alan Fisher et al., *Price Effects of Horizontal Mergers*, 77 CAL. L. REV. 777-827 (1988).

34. *See* Scheffman & Coleman (2003), *supra* note 25.

oligopoly theories are appropriate for the specific merger. In effect, this analysis can appeal to the timeless Friedman commentary on methodology in economics.[35] The analyst can infer the market behaves "as if" competition follows a model of collusion whenever the collusion model generates testable implications for competition in the particular market that are not falsified by the evidence. If multiple models survive the testing process, the choice between the models is based on a balancing of simplicity and fruitfulness considerations. Mathematical derivation from stylized facts may be helpful, but it is not necessary. In the next subsection, details on how this testing process appears to have played out are given.

C. FOLK THEOREM OF MERGER ENFORCEMENT

In reviewing FTC enforcement activity, it is clear that the *Guidelines*' analyses define testable hypotheses for the competitive effects of mergers.[36] The structural analysis explains the effect of the merger-related change in structure on the competitive environment. Competitive concerns are raised when the merger is likely to generate an adverse effect on consumers in a relevant market. While the bulk of the coordinated effects analyses remain qualitative, mathematical collusion models can be designed to predict less than competitive outcomes that can then be balanced against efficiencies. Any type of economic analysis actually generates a testable hypothesis for the effect of the merger.

From reviewing the case files, it is clear that a "Folk Theorem of Merger Enforcement" exists. Simply put, this theorem observes that whenever pre-merger evidence suggests a causal relationship between structure and performance exists, then a merger materially affecting structure is likely to substantially lessen competition. Theory is needed to give context to the evidence and evidence is needed to test the implications of the theory. As Friedman observes, economic science is hypothesis testing. Of course, testing does not guarantee success in court, because the defendants might also advance a validated economic theory suggestive of continued competition. The legal process sorts out the valid evidence and reaches a decision on the merits.

35. MILTON FRIEDMAN, ESSAYS IN POSITIVE ECONOMICS 3-43 (1953).

36. Scheffman et al. (2003) note that evidence on customer concerns and hot documents have always been used to support inference of coordinated interaction in the modern *Guidelines* era. Natural experiments were also noted as relevant to the study of likely competitive effects. *See* David Scheffman et al., *Twenty Years of Merger Guidelines Enforcement at the FTC: An Economic Perspective*, 71(1) ANTITRUST L.J. 277-331, 304 (2003). The FTC-DOJ merger commentaries also detail situations in which evidence is used to support inferences of less than competitive behavior stemming from a merger. *See* U.S. Federal Trade Commission & U.S. Department of Justice, Commentary on the Horizontal Merger Guidelines (2006), at 22-23, *available at* http://www.ftc.gov/os/2006/03/CommentaryontheHorizontalMergerGuidelinesMarch2006.pdf. *Also see* Malcolm Coate, *Empirical Analysis of Merger Enforcement under the 1992 Merger Guidelines*, 27 REV. INDUS. ORG. 279-301 (2005); and Coate & Ulrick (2008), *supra* note 29.

In another article, three types of evidence used by FTC staff for "testing" coordinated interaction theories are discussed.[37] In the best case scenario, evidence of natural experiments is found in which a structural change, comparable to the merger in question, generated anticompetitive effects. This evidence is likely to be more available for the very explicit theories of concern (maverick and structuralist), because the search for evidence can be focused. Natural experiment evidence may also be inferred from evidence on hot document or customer complaints. In this instance, the idea is that the hot document or customer complaint is based on a firm's or customer's recollection of a natural experiment (or series of natural experiments), that leads to the conclusion memorialized in the specific evidence. The reviews of the FTC analyses show some form of evidence is regularly found in the investigations, although actual evidence is not a necessary condition for an enforcement action.[38] In the next section, econometric analysis checks for the link between natural experiment evidence and coordinated interaction findings.

IV. An Enforcement Model for Coordinated Interaction

The background on economic theory will serve to structure the search for an empirical model of coordinated interaction. Statistical analysis should highlight relationships between structural characteristics and FTC enforcement policy (this paper's proxy for a collusion concern), as well as identify the impact of the natural experiment effects evidence or any other explanatory variable. One limitation must be discussed. The 1992 *Guidelines* propose a case-specific study of coordinated interaction, with the analyst required to obtain data on information structures and institutional realities. This style of analysis is not easily quantified and thus must be left out of this study. However, this search for informational structures and institutional realities is (1) based on structural considerations and (2) would be expected to affect the collection of effects evidence. Thus, the formal *Guidelines* style of analysis may be implicit in the modeling. In the following two subsections, the data collection process is reviewed first, followed by a discussion of model specification.

37. Coate (2005), *id.* It is also possible to consider natural experiment evidence that supports a pro-competitive theory of the merger. Initial analysis shows this pro-competitive evidence variable does not have a significant effect on the enforcement decision.

38. In addition to explicit or implicit natural experiments, it is possible to test theories of competitive concern with general economic evidence. The anticompetitive effects associated with merger to monopoly are the best example of such an analysis, because economic science has systematically found monopolies behave in a less than competitive manner. The structure-conduct-performance model also provided the scientific basis for merger enforcement until the general version of the theory was falsified. *See* Coate and Fischer (2008), *supra* note 18. A recent study by Kovacic et al. (2005) suggests that mergers to duopoly are problematic, at least in homogeneous (chemical) industries. *See* William Kovacic et al., Lessons for Competition Policy from the Vitamins Cartel (unpublished manuscript) (Sep. 2005), *available at* http://ssrn.com/abstract=818744.

A. DATA ON COLLUSION ANALYSIS IN MERGER ENFORCEMENT

The merger enforcement decision at the FTC has been studied in a number of recent papers.[39] While these papers estimate the probability of a merger challenge, the basic data can also be used to evaluate the likelihood of a subsidiary finding that the merger makes collusion "more likely, more successful, or more complete." This subsidiary finding will drive the enforcement decision if (1) coordinated interaction is the appropriate theory, (2) barriers to entry are present, and (3) efficiencies are integrated into the analysis to account for the overall effect (if any) of cost savings. Thus, to transform a dataset focused on the enforcement decision into a dataset addressing the likelihood of coordinated interaction, it is only necessary to delete all the unilateral effects cases, remove the matters in which ease of entry is dispositive on the issue of competitive concern, and incorporate evidence on efficiencies into the model. Of course, such an analysis is only able to identify the interpretation of coordinated interaction that appears in FTC enforcement decisions. To the extent the agency's interpretation is not consistent with an economic evaluation of the coordinated interaction concerns, the analysis may draw incorrect conclusions.[40]

The data collection process started with the 166 merger investigations identified in Bergman et al. exhibiting between one and three markets potentially affected by the merger and added twenty-one new matters reviewed in 2004 and 2005 to the dataset.[41] The 108 unilateral effects cases were deleted to focus purely on the collusion investigations. A further 19 files were deleted as the staff attorneys concluded entry was easy. This left a total of 60 collusion cases. To increase the sample, additional markets associated with the 187 matters were coded, whenever (1) the theory of concern was oligopoly and (2) the legal staff found barriers to entry. By looking at every market studied in the 187 investigations, it was possible to increase the sample to 76 investigations. One case had to be deleted, because the decision to close was intertwined with the failing-firm analysis. This left a sample of 75 collusion merger investigations undertaken during the 1993 to 2005 period.

39. *See* Mats Bergman et al., Comparing Merger Policies: the European Union versus the United States (unpublished manuscript) (March 2007), *available at* http://ssrn.com/abstract=975102; Malcolm Coate & Shawn Ulrick, *Transparency at the Federal Trade Commission: The Horizontal Merger Review Process*, 73(2) Antitrust L.J. 531-570 (2006); and Coate (2005), *supra* note 36. For an overview of all the recent papers using FTC data, *see* Malcolm Coate, An Overview of Transparency at the Federal Trade Commission: Generalities and Innovations in Merger Analysis (unpublished manuscript) (September 2008), *available at* http://ssrn.com/abstraxt=1111687.

40. Of course, the study would remain useful as an evaluation of the internal review structure even if it could be shown that some enforcement decisions were not consistent with standard economic theory.

41. *See* Bergman et al. (2007), *supra* note 39.

Figure 1 defines the structural and evidence variables collected from the FTC files and provides the ranges and summary statistics.[42] In addition to the standard information on the Herfindahl (HHI), the detailed review of the files identified

Figure 1 Overview of the Explanatory Variables

Variable	Definition	Range	Mean Enforced	Mean Closed
Herfindahl (HHI)	Herfindahl Index computed by summing the square of market share held by each firm in the post-merger market	1437-7008	3727*	2990
Significant Rivals	Number of pre-merger "significant" rivals in market affected by merger	3/10	4.02*	5.85
Fringe Share	Market share held by firms not considered significant in the analysis.	0/43.6	9.20*	12.3
Share Ratio	Ratio of share of largest post-merger firm to share of second largest entity	1/5.8	1.91	1.76
Homogenous Good	Indicator variable for homogeneous goods finding in attorney data	0/1	.521	.370
Buyer Sophistication	Indicator variable to identify cases in which buyer power was found by either attorneys or economists	0/1	.188*	.370
Inelastic Demand	Indicator variables for market in the hospital, oil, or drug industry.	0/1	.417	.370
Vertical Issues	Indicator variable to identify cases in which vertical aspects of merger were considered by attorneys or economists	0/1	.125*	.370
Efficiencies	Indicator variable for efficiency finding by either attorneys or economists	0/1	.438*	.741
Evidence	Index of anti-competitive findings associated with customer complaints, hot documents or historical natural experiments by either attorneys or economists; 0 implies no such findings, 3 means all three factors reported.	0-3	1.17*	.407
Maverick Firms	Indicator variable for Maverick firm finding in the attorney files	0/1	.270	.111
Administration (Bush)	Indicator variable for control of FTC by Chairman appointed after June 2001	0/1	.333	.185
Cases	Number of matters reviewed		48	27

*The sample mean for the enforced cases is significantly different from the mean for closed cases.

42. The raw structural information (market concentration, theories of concern, and product homogeneity) are based on the attorney analyses. More complex variables (not necessarily addressed in every attorney memo) are based on findings by either attorneys or economists.

the number of firms in the market with an ability to materially affect the outcome of the competitive process (defined as significant rivals and measured prior to the merger) as well as raw market share data. This share information allowed the calculation of two additional explanatory variables: fringe share and leadership ratio. The fringe variable was computed by summing the shares attributed to the significant rivals and then subtracting that number from one. It ranged from 0 to over 40 percent. The leadership proxy was calculated by dividing the share of the leading firm by the share of the second largest firm in the market. This variable ranged from one (for numbers equivalent situations in which no firm leads the market) to over five. It was calculated for the post-merger environment to incorporate the change in structure caused by the merger.

Five binary variables were coded to capture insights associated with market structure. A homogeneous goods index was derived based on a staff finding of relative similarity for the products in the market.[43] If the staff reported the specific good was customized to the buyer's specification, this was also coded as homogenous when the bulk of the firms in the market were able to meet consumer demand on relatively equal terms. Buyer sophistication signaled a staff observation that customers had some ability to negotiate with their suppliers.[44] Having a large buyer implied sophistication, but no clear buyer share cutoff existed. The vertical variable identified mergers in which the staff investigation identified some vertical relationship affected by the merger.[45] For example, if a large upstream firm with a horizontal presence in a downstream market bought a competitor, the vertical relationship would be found. A fourth variable attempts to proxy the elasticity of demand with information on the industry involved in the merger.[46] Finally, an efficiency index was assembled.[47] It takes on the value one whenever either the attorneys or economists report evidence of merger-specific efficiencies. As the FTC often stops short of formally endorsing the efficiency claims, language suggestive of merger-specific savings was taken as relevant.[48] Figure 1 reports that the enforced matters were statistically likely to exhibit different findings for sophistication, vertical issues, and efficiencies.

43. *See* Coate (2006), *supra* note 29.

44. Buyer sophistication was borrowed from Coate & Ulrick (2008), *supra* note 29. In effect, the variable flagged markets in which the staff recognized that the institutional arrangements associated with market competition allow complex bilateral negotiations.

45. The vertical variable is lifted from Bergman et al. (2007), *supra* note 39.

46. Hospital, oil, and drug industry matters are assumed to exhibit relatively inelastic demand.

47. *See* Coate (2005), *supra* note 36.

48. The level of verification memorialized in the file varied from case to case, thus conclusionary language was used to code the index.

Performance evidence associated with natural experiments was also collected.[49] Simply put, each file was reviewed for hot documents, validated customer complaints, and economic effects compatible with the theory of concern. The evidence on hot documents and customer concern was relatively easy to obtain from the files.[50] FTC staff generally highlighted these "legal" findings and explained their importance relative to the rest of the information in the file. Thus, it was possible to separate the relevant from the irrelevant information.[51] The economic effect variable required more creativity, because economic facts must be interpreted in light of a theory of concern. In this instance, the analysis first identified the theory of concern at the core of the investigation and then evaluated any natural experiment supportive of the predictions of that theory. For example, economic evidence that the acquired firm had behaved as a maverick in the past and protected competition would be considered supportive of the implications of a maverick theory of violation.[52]

Two other variables ("maverick" and "administration") were recovered from the files. The maverick variable flags the cases in which the staff simply reported a claim of maverick status for one of the merging parties, but failed to present natural experiment evidence supportive of the maverick model. Thus, this indicator flags the cases in which the investigation identified unsubstantiated allegations of maverick status. The second variable, administration, identified the cases filed after June 2001. These matters were all decided under a chairman appointed by U.S. President George W. Bush.

Figure 2 presents some initial information on the data. As a first analysis, the shares of all of the significant competitors were identified and analyzed. The first row summarizes results for 36 matters in which one of the two merging firms holds the largest share in the market. By definition, these firms would obtain an even larger lead on their rivals if the merger was consummated. However, it is important to note that the merger would not allow the new firm to dominate the market, because that concern would have triggered a unilateral effects analysis. The second row focuses on transactions in which a larger leading firm was created by the merger. Here, another 22 matters are flagged, although in some mat-

49. *See* Coate (2005), *supra* note 36.

50. This is detailed in Coate & Ulrick (2006), *supra* note 39.

51. Customers complain about all sorts of things, sometimes related to the likely competitive effect of the merger on competition and other times linked to the effect of the merger on their business. FTC staff reviews identify the complaints associated with a loss of competition. Likewise, a range of documents can be identified as "hot." The review only flags the claims when the document links an adverse effect on competition to the consummation of the merger.

52. Posner's list of performance characteristics could be useful to show the market is currently performing in a less than competitive manner, a result supportive of a structural model of concern. *See* Posner (1976), *supra* note 7.

Figure 2 Types of Collusion Investigations with Barriers to Entry[a] (percentage in parenthesis is enforcement rate)

Market Conditions	Matters	Homogeneous Sophisticated	Homogeneous Not Sophisticated	Not Homogeneous
One Merger Partner is Industry Leader	36 (66.7%)	6 (50%)	13 (92.3%)	17 (52.9%)
Together Merged Firm Becomes Industry Leader	22 (63.6%)	0 (NA)	10 (70.0%)	12 (58.3%)
Together Merged Firm Becomes Second Firm	8 (62.5%)	1 (0%)	0 (NA)	7 (71.4%)
Number Equivalent Analysis	5 (60.0%)	0 (NA)	3 (100%)	2 (0%)
Together Merged Firm Becomes Third Firm	4 (50.0%)	1 (0%)	1 (0%)	2 (100%)
Total	75 (64.0%)	8 (37.5%)	27 (81.5%)	40 (57.5%)

[a]Number of Cases for each structural regime listed in Figure.

ters, the post-merger share would barely exceed that of the previous market leader. The next row provides information for mergers that create a larger number two firm. Only eight cases are found. The fourth row reports on five numbers-equivalent cases. In these matters, the staff weights all the significant competitors equally; hence, a merger would reduce the number of players by one, but have no other effect on structure. The last row counts the mergers that create stronger number three competitors, where a total of four cases are noted.

In theory, structural analysis should differ for each type of case, probably leading to significant differences in enforcement probabilities. In fact, this does not occur as the enforcement rates vary only slightly (50 percent to 66.7 percent) over the sample. The rest of figure explores this result, disaggregating the cases by a combination of homogeneous goods and buyer sophistication status. Although the sample of homogeneous goods with sophisticated buyers is small, these matters show relatively low enforcement (37.5 percent), especially when compared to the rate associated with homogeneous goods with unsophisticated buyers (significantly higher at 81.5 percent). Moreover, the homogeneous goods markets without sophisticated buyers are statistically more likely to end in enforcement action than the differentiated goods.[53] In light of these complex interrelationships in the data, econometric analysis is required to sort out the regularities in the data.

53. Interestingly, buyer sophistication does not appear to affect enforcement probability in the differentiated goods sample.

B. MODELING MERGER ENFORCEMENT

In theory, it would be desirable to separately model what appear to be the FTC's three collusion theories (maverick, regime shift, and structuralist). However, this approach is precluded by data limitations. Instead, it is necessary to aggregate all the data together and explore two general models, one that focuses purely on structural considerations and the other that adds an evidence variable to the analysis. If evidence matters, then the variable would take on a significant positive sign. If the structural variables also matter, then they will retain their statistical significance (and signs). Within this approach, it is possible to model market concentration in more detail, investigate the scope of customer sophistication, and explore the impact of the maverick theory. Finally, it is possible to see if the Bush administration changed the decision-making process for coordinated interaction investigations.

FINALLY, IT IS POSSIBLE TO SEE IF THE BUSH ADMINISTRATION CHANGED THE DECISION-MAKING PROCESS FOR COORDINATED INTERACTION INVESTIGATIONS.

The basic structural model focuses on a concentration index (e.g., the Herfindahl index) and five control variables (homogeneous good, buyer sophistication, inelastic demand, vertical considerations, and merger-related efficiencies). The market share data available in the files allows the analysis to move beyond the Herfindahl index and explore market structure in more detail. Three variables are considered. First, the number of significant competitors is included, because the coordinated interaction would require cooperation from all significant players. Second, the size of the fringe could matter, because a substantial fringe might be able to undermine the collusion. Third, the potential for leadership might be important, because a relatively large firm could set the terms for coordination. Leadership is defined as the post-merger ratio of the share of the leading firm to its largest competitor. These three variables are used to replace the Herfindahl index in some of the specifications.

The five control variables all represent standard structural considerations. First, it is generally considered easier to coordinate price when the product is relatively homogeneous. Therefore, a positive sign is expected for this index. Buyer sophistication tends to make coordinated interaction less likely, as sophisticated buyers are able to negotiate with the various competitors in the market and possibly undermine collusive prices. In this instance, a negative sign is expected. Inelastic demand is generally thought to make collusion more likely as output restrictions needed to support higher prices would be relatively low. A positive sign is likely. The vertical indicator identifies markets in which firms interact at various vertical levels. In general, this observation would suggest that less room exists for collusive behavior. Moreover, the merger supposes a change in this vertical relationship, a change that would tend to reduce vertical transaction costs. These vertical efficiencies also make cartelization less likely. Overall, a finding of vertical ramifications implies a negative relationship with a finding of coordinated interaction.

The final variable addresses efficiency considerations. Findings of efficiencies imply an effect on the cost structure of some of the market competitors. Cost differentials can affect the probability of coordinated interaction in two ways. First, cost differences make it more difficult to coordinate on price, hence collusion is less likely. Second, even if some coordination can occur, the price could be lower. Both considerations point to a negative effect on coordinated interaction. One generalization of the basic model is considered, as customer sophistication is interacted with the homogeneous good variable. Basically, sophistication would be expected to have a much greater effect when the market is homogeneous and customers could more easily pit suppliers against each other and undermine any collusive agreement.

The structural model is complemented with the evidence index. This variable serves to identify the degree of exogenous support for the relevant theory of coordinated interaction. Logically, the more evidence supporting a competitive concern, the more likely a collusion finding will be made. Hence, a positive coefficient is expected for the evidence variable. The other structural variables may retain their significance, or become statistically indistinguishable from zero, depending on whether evidence supplements or trumps market structure.

Next, an indicator for a claim of maverick status is added to the model. If the maverick status mattered, it would make a finding of collusion more likely whenever a maverick firm is identified.[54] Note the analysis is only testing for the importance of a maverick allegation, because natural experiment evidence related to a maverick effect is already included in the model through the evidence index.

Finally, an indicator for Bush administration control of the FTC is included in the model. The shift parameter indicates whether the Bush administration revitalized coordinated effects analysis.[55] This data is unable to determine if the administration was more aggressive overall, because it would require a joint study of entry and coordinated effects analysis.

V. Estimation of the Models

The statistical analysis is undertaken in a series of twelve probit regressions, presented in Figures 3, 4, and 5.[56] The discussion in this section will track the visu-

54. *See* Baker (2002), *supra* note 3.

55. *See* Darren Tucker & Bilal Sayyed, *The Merger Guidelines Commentary: Practical Guidance and Missed Opportunities*, ANTITRUST SOURCE 1-13, 8 (May 2006).

56. Both the continuous and discrete variables are transformed with the natural logarithm function to allow for more nonlinear effects (one is added to the evidence index to enable the transformation). A clustered errors technique is used to address the fact that some mergers are represented by two or three analyses of different markets of concern.

al presentation of the models, although the bulk of the analysis will be focused on the last model in each figure. Examples of the impact of various structures on the probability of a concern will be given.

Figure 3 presents a standard Herfindahl-based model of coordinated interaction.[57] In all the specifications, the Herfindahl index is positively related to a concern, suggesting market shares matter. In model 3-1, customer sophistication, inelastic demand, and vertical issues also affect the collusion finding, with the expected signs. Efficiencies exhibit the expected negative effect, but tests slight-

Figure 3 Standard Oligopoly Model[a]

	3-1 (Structure)	3-2 (Structure)[b]	3-3 (Evidence)	3-4 (Evidence)[b]
HHI	2.756*** (5.11)	2.864*** (4.32)	2.588*** (4.55)	2.665*** (3.66)
Homogenous Good	.3813 (.92)	1.054** (2.18)	.4563 (1.02)	1.358** (2.13)
Buyer Sophistication or Sophistication * Homogeneous	–1.480*** (–2.85)	–2.10** (–2.51)	–1.452*** (–2.77)	–2.522*** (–2.81)
Inelastic Demand	.8416* (1.91)	.3398 (.83)	1.058** (2.23)	.5889 (1.32)
Vertical Issues	–.9773** (–1.98)	–.5562 (–1.12)	–1.388*** (–3.34)	–1.105*** (–2.57)
Efficiencies	–.5945 (–1.43)	–.7982* (–1.95)	–.6931 (–1.34)	–.9962* (–1.70)
Evidence	–	–	1.253*** (2.86)	1.603*** (2.80)
Constant	–20.97*** (–4.89)	–22.05*** (–4.17)	–20.08*** (–4.51)	–20.94*** (–3.61)
Predictions (percentage)	80.0	82.7	84.0	86.7
Pseudo-R-square	.4296	.4505	.5015	.5489
Log Pseudo-likelihood	–27.95	–26.93	–24.43	–22.11

a. t-statistic in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. b. Buyer sophistication is interacted with homogeneous good index.

57. The change in Herfindahl along with the interaction of the change and the Herfindahl can be included in the model, but proved to be statistically insignificant and thus are not reported.

ly below conventional levels of statistical significance. A finding of homogeneous goods increases the chance of enforcement, but the effect is not significant. Looking back at Figure 2, this result is not surprising, because customer sophistication appears to interact with product homogeneity. Re-estimating the model with the buyer sophistication variable interacted with homogeneity (model 3-2) generates the expected effect.[58] On its own, a homogeneous good facilitates coordinated interaction. However, when customer sophistication is relevant, a competitive problem is less likely, as sophisticated consumers have an ability to protect themselves from collusive overcharges. In this second regression, the other control variables retain their signs, but now only the efficiency variable is statistically different from zero.

The next two regressions (models 3-3 and 3-4) repeat the two initial specifications, but add the evidence variable to the regressions. Here, evidence suggestive of a competitive concern has the expected positive effect on the probability of a coordinated interaction finding. This result serves to confirm the importance of testing structural oligopoly models with natural experiment-related evidence. The structural variables retain their signs and all but one achieves statistical significance in model 3-4.[59] The final model correctly predicts 87.5 percent of the 48 collusion findings, and 85.2 percent of the 27 no-effect conclusions.

The magnitude of the coefficients cannot be directly compared across the models. Instead, any comparison must evaluate the standard normal function given values for all of the other variables in the model. For example, consider an efficient merger in a homogeneous goods market. Using model 3-2, the probability of a collusion concern moves from 15 to 87 percent, as the Herfindahl increases from 1,400 to 3,000. Switching the focus to model 3-4 generates marginal reductions in the enforcement probability when no evidence is available, but a single finding of evidence causes the probability to jump to a range of 44 to 97 percent. Other information can significantly change these probabilities. Holding the structure and evidence variables at fixed values, the direction of the effect for the five remaining structural variables can be computed by just summing up the relevant coefficients. For example, in the last situation mentioned, findings of buyer power and inelasticity will create a probability of collusion ranging from almost 0 to 87 percent. Overall, it is clear that both structural findings and evidence matters.

Figure 4 repeats the analysis in Figure 3, but replaces the Herfindahl index with three parameters designed to offer a more detailed structural analysis. The

58. An alternative specification in which the buyer sophistication index was interacted with both the homogenous and differentiated good variables was also estimated. The interaction of buyer sophistication and differentiation was not significantly different from zero in any of the models, and it was removed from the model for expositional ease to obtain the specifications presented in the text.

59. The five structural indictor variables test jointly significant. (The Chi-square statistics are 19.26, 12.65, for models 3-3 and 3-4, respectively. Both models' results are above the cutoff for five degrees of freedom of 11.07.)

count on the number of significant competitors is inversely related to the likelihood of coordinated interaction in all the specifications. The ratio of the share of the leading firm to its closest competitor takes on the expected positive sign, but its significance level is marginal in the first two specifications. Fringe share is significant in only one specification.[60] It is possible to nest the models in

Figure 4 **Complex Oligopoly Model**[a]

	4-1 (Structure)	4-2 (Structure)[b]	4-3 (Evidence)	4-4 (Evidence)[b]
Significant Rivals	–4.320*** (–3.83)	–5.787*** (–5.29)	–4.857*** (–3.70)	–8.100*** (–3.92)
Share Ratio	.5422 (1.62)	.5529 (1.57)	.7673** (1.96)	.9379** (2.01)
Fringe	–.3339 (–1.58)	–.0890 (–.44)	–.4607* (–1.95)	–.2319 (–.65)
Homogenous Good	.6109 (1.34)	1.939*** (3.46)	.6445 (1.30)	3.209*** (3.29)
Buyer Sophistication or Sophistication * Homogeneous	–1.547*** (–2.77)	–3.227*** (–3.48)	–1.893*** (–2.90)	–5.488*** (–3.77)
Inelastic Demand	1.002* (1.93)	1.222** (2.12)	1.352** (2.50)	1.953** (2.29)
Vertical Issues	–1.321** (–2.28)	–1.109* (–1.74)	–1.959*** (–3.47)	–2.295*** (–3.53)
Efficiencies	–.5485 (–1.32)	–.7694 (–1.54)	–.9883* (–1.95)	–1.867** (–2.51)
Evidence	–	–	1.661*** (3.14)	2.773*** (2.80)
Constant	7.075*** (4.10)	8.362*** (4.95)	7.532*** (3.69)	11.30*** (3.56)
Predictions	81.33	84.0	85.33	89.3
Pseudo-R-square	.4735	.5530	.5670	.6902
Pseudo Log-likelihood	–25.80	–21.91	–21.21	–15.18

a. t-statistic in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. b. Buyer sophistication is interacted with homogeneous good index.

60. Interacting fringe share with product homogeneity does not improve the results. Possibly fringe firms face expansion constraints in the real world, or customers require partnership relationships with their core suppliers even when the good is homogeneous. Thus, industry-specific facts may limit the importance of the fringe.

Figures 3 and 4 by adding the Herfindahl index to the Figure 4 specification.[61] In the (unreported) regressions, the Herfindahl index never becomes statistically significant, while the joint hypothesis setting the coefficients of three new concentration variables to zero cannot be rejected for the first two specifications. Thus, it is impossible to distinguish between the first two sets of share-based parameters. Once the evidence variable is added, it is possible to conclude that, when taken together, significant rivals, ratio of shares of leading firms, and fringe share are statistically different from zero.[62] Thus, when data is available, the user has a reason to prefer models 4-3 and 4-4 over models 3-3 and 3-4.

Adding the evidence variable in models 4-3 and 4-4 increases the level of significance of the other control variables. The results for buyer sophistication match those found in Figure 3, in which the sophistication effect is basically focused in homogeneous goods industries. Elasticity and vertical ramifications remain significant in the new specifications, while the pure efficiency effect becomes significant. Thus, in all the specifications, the structural variables remain important when the model is generalized to address explicit or implicit natural experiment evidence.[63] Model 4-4 correctly predicts 91.7 percent of the findings of collusion and 85.2 percent of the no-effect matters.

Predictions for the probability of a collusion finding would generate similar results to those discussed above, although now the structural parameters would focus on significant competitors: the leadership share ratio and fringe share. As noted above, models 4-3 and 4-4 appear preferable to the simple Herfindahl models 3-3 and 3-4. Of course, for any particular merger, it is straightforward to compute a fitted value for any probit equation and use the standard normal function to generate the probability of a collusion finding that ranges from zero to one. While the models in Figure 4 do a better job of predicting than those in Figure 3, the difference is small (1 to 2.5 percentage points). A more sophisticated analysis would look at each prediction. The review of the fitted probabilities shows the more complex model (model 4-4) predicts the correct outcome within ten percentage points of the actual outcome (over 90 percent for concern and under 10 percent for no concern) in 69.3 percent of the cases. In contrast, the standard Herfindahl model (model 3-4) only achieves this success in 52 percent of the transactions. Overall, the more detailed model appears to perform better,

61. As the three new structural variables explain 84 percent of the variance in the Herfindahl, multicollinearity may limit the results of the testing.

62. The Chi-square statistics needed to reject a zero effect for rivals, share ratio and fringe share are 14.4 and 13.86 for models 4-3 and 4-4, respectively. Both are greater than the relevant Chi-square cutoff of 7.81. For models 4-1 and 4-2, the test statistics are insignificant, at 4.56 and 4.51, respectively.

63. Given that the t-statistics already highlight the significance of the five structural indictor variables, it is not surprising that the joint Chi-square test also generates highly significant results (17.7 for model 4-3 and 19.22 for model 4-4).

Figure 5 Oligopoly Model with Maverick or Administration Variables[a]

	5-1 (Maverick)	5-2 (Admin)	5.3 (Maverick)	5-4 (Admin)
HHI	2.630*** (3.61)	2.521*** (3.58)		
Significant Rivals			–7.943*** (–3.90)	–8.535*** (–4.13)
Share Ratio			.9416** (2.00)	1.134** (2.38)
Fringe			–.2420 (–.65)	–.04312 (–.16)
Homogenous Good	1.376** (2.09)	1.283* (1.84)	3.170*** (3.23)	3.398*** (3.40)
Sophistication * Homogeneous	–2.519*** (–2.77)	–2.763*** (–2.98)	–5.414*** (–3.73)	–6.588*** (–3.97)
Inelastic Demand	.6989 (1.43)	.7376 (1.48)	1.938** (2.31)	2.877*** (3.07)
Vertical Issues	–1.104*** (–2.60)	–1.544*** (–2.98)	–2.258*** (–3.46)	–3.012*** (–3.74)
Efficiencies	–.9937* (–1.71)	–.7700 (–1.33)	–1.843** (–2.48)	–1.844** (–2.41)
Evidence	1.668*** (2.79)	1.773*** (2.84)	2.773*** (2.72)	3.098*** (2.94)
Maverick Firms	.7195 (1.00)	–	.3225 (.38)	–
Administration	–	.9384 (1.62)	–	1.446** (2.04)
Constant	–20.80*** (–3.55)	–20.18*** (–3.53)	11.02*** (3.42)	10.69*** (3.90)
Predictions	84.0	82.7	88.0	89.3
Pseudo-R-square	.5572	.5741	.6911	.7265
Pseudo Log-likelihood	–21.70	–20.87	–15.14	–13.41

a. t-statistic in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

although the model requires a complex understanding of market structure. Without access to all the market share data, the complex model cannot be used.

Figure 5 explores two special considerations, one that turns out interesting and one that does not. The models in model 3-4 and 4-4 were recycled into Figure 5, with models 5-1 and 5-3 focus on maverick firm findings, while models 5-2 and 5-4 search for a change in understanding of oligopoly associated with the Bush administration.[64]

The maverick results are anticlimactic, as the index associated with maverick-based analysis is insignificant (although the coefficient takes on the expected positive sign). The results on the other variables are robust, suggesting that exclusion of this effect does not impact the analysis. While maverick-based analysis remains a theory of collusion and finding natural experiment evidence on real-life maverick behavior would generally support a competitive concern, speculation on maverick status adds nothing to the likelihood of a coordinated interaction finding. These statistical conclusions are not compatible with the Baker-Shapiro hypothesis that the maverick model is the single theory of collusive oligopoly.

The results associated with the Bush administration variable are much more interesting. In model 5-2, the dummy variable exhibits a positive effect and the test statistic approaches conventional levels of significance. While the coefficients on the other variable jump around a little, the results do not strongly support inclusion of the administration effect. In contrast, model 5-4 identifies a significant administration effect, suggesting that the Bush enforcers were more likely to infer coordinated interaction. While the magnitude of the coefficient is smaller than those associated with the other binary variables, the effect is still substantial. Adding this variable allows the model to explain 93.8 percent of the collusion findings and 88.9 percent of the matters in which the theory was rejected.

FINALLY, WELL-DONE COORDINATED INTERACTION ANALYSES ARE SIMPLY MORE LIKELY TO PREVAIL IN COURT, BECAUSE THESE STUDIES WILL PROVIDE A LINK TO REAL-WORLD EVIDENCE.

Further analysis addressed the question of how big a shift in policy is suggested by the data. A simple split of the sample showed an enforcement probability of 59.3 percent prior to June 2001 and 76.2 percent after. Without statistical decomposition, it is impossible to tell how much difference is due to the change in understanding and how much is related to the specific sample.[65] Data limitations (there are only 21 Bush administration cases) preclude this analysis. However, it is possible to simulate the pre-Bush situation that would have occurred had the Bush administration merger been filed prior to June 2001. Focusing on the 16 collusion findings, the data suggest that the Bush effect is responsible for 3 of the conclusions. Of course,

64. For the 1993-2005 sample, the Bush administration investigated 45.9 percent of their cases with a collusion theory, while the earlier administrations studied 41.8 percent of their cases with collusion analysis. The similarity in these two figures suggests the Bush administration did not systematically reclassify matters from unilateral to collusion.

65. For an example of decomposition, *see* Bergman et al. (2007), *supra* note 39.

it is impossible to draw any overall conclusions without an analysis of entry barriers. If the Bush administration also made it easier to reach an ease of entry finding, the two effects could cancel out.

VI. Conclusions

FTC enforcement policy allows the analyst to draw insights into the agency's best understanding of coordinated interaction in oligopolistic industries. By focusing on the sub-sample of collusion cases and excluding the matters in which entry is easy, the enforcement decision effectively proxies a finding on ease of collusion. Statistical results are broadly consistent with economic theory. Concentration-related variables like the Herfindahl, a count of the number of competitors, the homogeneity of the market, and the sophistication of the customer base for homogeneous goods all have strong and expected effects on the outcome. Proxies for inelasticity, vertical relationships, and changes in cost caused by efficiencies also have expected effects in some specifications. A leadership variable appears to contribute to the analysis, while no consistent effect for the size of the fringe can be found.

This structural model retains some explanatory power when a variable associated with exogenous evidence (natural experiments, validated customer concerns, and hot documents) is added to the model. This test of the Folk Theorem of Merger Enforcement (if evidence supports a structural problem, then it is reasonable to infer a competitive concern from a relevant change in structure) confirms the importance of the natural experiment evidence. Other results note a maverick theory of violation does not add to the concern associated with the structure and the Bush administration was more aggressive in its analysis of coordinated interaction.

A number of practical implications are obvious. First, market definition must remain the first step in merger analysis. Coordinated interaction only makes sense if the rivals that the merged firm is expected to coordinate with can be identified. Second, market shares can be integrated into the analysis in a sophisticated manner when the detailed data is available. Third, natural experiment evidence can be very useful in confirming the implications of a structural model. Thus, the *Guidelines*' focus on market institutions remains highly relevant, because understanding the competitive dynamics of the market is likely to aid the search for natural experiments. Fourth, customer concerns and hot documents should be analyzed in great detail in an attempt to isolate the natural experiments that underpin these concerns. While some facts may be lost to history, the reconstruction of even a qualitative natural experiment could serve to confirm a coordinated effects theory. Fifth, it appears customer sophistication serves to reduce the likelihood of collusion. This observation also implies a need to understand how the market of interest actually performs. Finally, well-done coordinated interaction analyses are simply more likely to prevail in court, because these studies will provide a link to real-world evidence. While it is possible to ground unilateral effects analyses in fact, it is also possible to become captured by the deductive logic that builds from the profit-maximization assumption to the theoretical conclusion on a price increase. In court, assumptions are not proof.

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Tacit versus Overt Collusion Firm Asymmetries and Numbers: What's the Evidence?

Stephen Davies & Matthew Olczak

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Tacit versus Overt Collusion Firm Asymmetries and Numbers: What's the Evidence?

*Stephen Davies and Matthew Olczak**

It is conventional wisdom that collusion is more likely the fewer firms there are in a market and the more symmetric they are. This is often theoretically justified in terms of a repeated non-cooperative game. Although that model fits more easily with tacit than overt collusion, the impression sometimes given is that 'one model fits all'. Moreover, the empirical literature offers few stylized facts on the most simple of questions—how few are few and how symmetric is symmetric? This paper attempts to fill this gap while also exploring the interface of tacit and overt collusion, albeit in an indirect way. First, it identifies the empirical model of tacit collusion that the European Commission appears to have employed in coordinated effects merger cases—apparently only fairly symmetric duopolies fit the bill. Second, it shows that, intriguingly, the same story emerges from the quite different experimental literature on tacit collusion. This offers a stark contrast with the findings for a sample of prosecuted cartels; on average, these involve six members (often more) and size asymmetries among members are often considerable. The indirect nature of this 'evidence' cautions against definitive conclusions; nevertheless, the contrast offers little comfort for those who believe that the same model does, more or less, fit all.

*Stephen Davies is a professor in the School of Economics and member of the ESRC Centre for Competition Policy at the University of East Anglia. Matthew Olczak is a research associate at the ESRC Centre for Competition Policy at the University of East Anglia. The research on which this article draws has benefited considerably from helpful comments from many participants in seminars and conferences in Athens, Bonn, Budapest, the OFT, Reading, UEA, Valencia, and Warwick. Malcolm Coate and Morten Hviid have also provided very helpful comments, and Oindrila De has kindly made her data on EC cartels available. We would like to thank, but not implicate, all these people. The financial support of the Economic and Social Research Council is also gratefully acknowledged.

I. Introduction

This article explores a strangely under-documented topic in the empirical literature: Are there well-defined (i.e. observable and predictable) differences between the market structures which give rise to tacit collusion as opposed to overt collusion (cartels)?

It is certainly received wisdom that collusion is more likely to occur with fewer leading players in a market and the more symmetric the players are. This was recognised long ago in the traditional Structure-Conduct-Performance paradigm, and was subsequently formalized with the theory of repeated non-cooperative games. This wisdom is also shared by practitioners. Dick[1] explains that case law embraces the presumption that suppliers' *ability* to coordinate should be closely linked to their fewness in numbers, quoting from *FTC vs. PPG Industries*[2]: "[W]here rivals are few, firms will be able to coordinate their behaviour, *either by overt collusion or implicit understanding*, in order to restrict output and achieve profits above competitive levels." (italics added). He goes on[3] to provide a persuasive explanation of why asymmetries make collusion difficult.

However, in spite of this consensus, the empirical literature offers few stylized facts on the most simple of questions—how few is few, how symmetric is symmetric, and how, if at all, does this differ between the different forms of collusion? It is true that, for cartels at least, we have extensive case evidence on firm numbers, although less so on asymmetries; but for tacit collusion, remarkably little is known about either. Given that the concept is somewhat elusive and not always easy to tie down in practice, this is hardly surprising.

Unfortunately, if we turn to theory for answers, it is of little assistance. The repeated game is best thought of as a model of tacit collusion, but it is also often assumed to apply equally to cartels, (see added italics above) and there often seems to be an implied presumption that one model fits all. Harrington is surely justified in claiming[4] that "there is a gap between antitrust practice—which distinguishes explicit and tacit collusion—and economic theory—which (generally) does not."

More generally, this nexus of overt and tacit collusion raises a number of policy-relevant questions: "How far are tacit collusion and cartels seen as substi-

1. A. R. Dick, *Coordinated Interaction: Pre-Merger constraints and Post-Merger Effects*, 12(1) Geo. Mason L. Rev. 65, 71 (2003) [hereinafter "Dick"].

2. FTC v. PPG Indus., 798 F.2d 1500, 1503 (D.C. Cir. 1986).

3. *Id.* 72-9.

4. J. E. Harrington, Jr., The collusion chasm: reducing the gap between antitrust practice and industrial organizational theory, Slide 7, Csef-Iger Symposium on Economics and Institutions, [hereinafter "Harrington"] *available at* http://www.econ.jhu.edu/People/Harrington/Capri.ppt.

tutes?" Do firms only look to form cartels when the legally safer option of tacit collusion is unattainable? Is cartel formation sometimes provoked by a breakdown in tacit collusion? And once a cartel is busted, should we suppose that subsequent behavior will approximate 'competition', or is some sort of tacitly collusive equilibrium a more sensible counterfactual?

For the purpose of this paper, we will define 'market structure' very narrowly by the number of firms in the market and the asymmetries among them. Asymmetries will be defined in terms of market shares—we are concerned with looking for stylized facts applicable across markets—but we should stress that these merely serve as a reduced form indicator of what really matters, i.e. the underlying causes of those asymmetries (see the next section). This is a limitation on our analysis which we freely acknowledge.

PRECISELY BECAUSE TACIT COLLUSION IS DIFFICULT TO IDENTIFY AND MEASURE IN PRACTICE, OUR RESEARCH STRATEGY IS ALMOST INEVITABLY INDIRECT.

Precisely because tacit collusion is difficult to identify and measure in practice, our research strategy is almost inevitably indirect. We begin in section 4 by recounting our recent study of the merger control case decisions by a Competition Authority ("CA"), the European Commission, in which tacit collusion (coordinated effects) appears to have been an issue. We argue that this is probably the only way of assembling a fairly large body of cases, equivalent to existing databases on cartels. Although actual tacit collusion is generally not illegal (hence the absence of actual cases), merger control is one area of policy where CAs are obliged to assess the prospects that tacit collusion may arise. In most major jurisdictions, there is a reasonably large number of such cases, providing the scope for constructing a database sufficiently large to support econometric analysis designed to uncover stylized facts about the sorts of market structures that are associated with tacit collusion—at least as seen through the eyes of the CA. From a European perspective, this is not without interest given the controversies of recent years concerning the Commission's decisions in celebrated cases such as *Airtours* (1999)[5], the academic critique of the Nestle-Perrier case[6] and the 2004 revision to the European Merger Regulation.

Having approached tacit collusion indirectly from this perspective, in section 5 we turn to a sometimes neglected (at least in the mainstream) academic literature: experimental work on tacit collusion. Again, given the obvious difficulties in simulating real world markets in a sterile laboratory environment, empirical experimental research should only be viewed as an indirect source of facts. However, given the difficulties in applying more traditional econometric field

5. M.1424 Airtours/First Choice (1999).

6. O. Compte, F. Jenny, & P. Rey, *Capacity constraints, mergers and collusion*, EUR.ECON REV, 46(1), 1-29 (2002) [hereinafter "Compte, et al"].

analysis, we suggest that experiments, like coordinated effects merger cases, can offer important insights into the subject. In section 6, we move on to more familiar literature on the characteristics of prosecuted cartels. Here, the facts on firm numbers are already reasonably well documented, but we add some new findings on asymmetries which are emerging from our own ongoing research on the structure of a sample of EC cartels.

These three disparate sources offer some thought provoking contrasts: while 'tacit collusion' is typically found in only symmetric duopolies, cartels are usually characterized by more (sometimes far more) than just two players, and often display very pronounced asymmetries in the members' market shares. Section 7 speculates on some of the implications for future research in all three areas: merger analysis, experimental research, and cartels.

The next two sections first provide some preliminaries. Section 2 briefly surveys the standard theoretical expectations on collusion and market structure. Section 3 introduces a simple geometric device, which we employ throughout the article to aid exposition.

II. Terminology and the Standard Economists' Model of Collusion

At the outset, we should be clear on terminology. Motta[7] provides a very clear discussion of the economist's distinction between cartels and tacit collusion, which captures what we take to be the prevailing view—it is certainly ours. While collusion might be defined in economic theory as any market outcome in which prices are high (relative to those in the one-shot non-cooperative equilibrium), collusion should only be considered illegal (i.e. equivalent to a cartel) where firms explicitly coordinate their actions. Where there is no explicit coordination, collusion is tacit and not illegal by default definition. The term 'tacit collusion' is perhaps a little inappropriate—'tacit coordination' might be less open to misunderstanding—but common practice dictates that we retain 'tacit collusion' here. Of course, in particular cases, there will be debate about certain practices—are they explicit or tacit coordination—but that is not the subject of this paper[8].

> THUS TACIT COLLUSION NEED NOT, AND GENERALLY WILL NOT, ENTAIL EXPLICITLY AGREED STRATEGIES OR INFORMATION EXCHANGE.

7. M. MOTTA, COMPETITION POLICY: THEORY AND PRACTICE, Ch. 4, (2004).

8. For a discussion of this issue, *see* M. MOTTA, CARTELS IN THE EUROPEAN UNION: ECONOMICS, LAW, PRACTICE, *paper written for* Fifty years of the Treaty: Assessment and Perspectives of Competition Policy in Europe Conference, IESE Business School (2007).

Thus tacit collusion need not, and generally will not, entail explicitly agreed strategies or information exchange, and the spirit of what we have in mind is still captured perfectly by Chamberlin's (oft-quoted, by amongst others, Tirole[9]) words of 75 years ago:

> "If each (firm) seeks his maximum profit rationally and intelligently, he will realise that when there are only 2 sellers, his own move has a considerable effect upon his competitors, and that it makes it idle to suppose that they will accept without retaliation the losses he forces upon them. Since the result of a cut by any one is inevitably to decrease his own profit, no one will cut, and although the sellers are entirely independent, the equilibrium result is the same as though there were a monopolistic agreement between them."[10]

This anticipates, and fits comfortably with, the contemporary interpretation of tacit collusion as a potential equilibrium outcome from a dynamic non-cooperative game. That model is routinely recited in all self-respecting industrial organization textbooks, and is rigorously and concisely summarized by Ivaldi et al[11], writing for the European Commission.

Starting from the base case of an homogenous product symmetric duopoly with Nash reversion, Ivaldi et al derive a series of standard results on the market conditions under which tacit collusion is likely, including: transparent prices, frequent interaction, absence of barriers, and buyer power; but here we are most concerned with firm numbers and asymmetries. They show that the critical threshold for the discount factor increases (and collusion becomes less likely) as the number of firms increases. The intuition is that, with more firms, each firm gets a lower share of the pie from colluding, thus increasing the gains from cheating as well as reducing the attractiveness of long-term collusion. A second factor working in the same direction, but not covered in this model, is that the likelihood that firms are able to tacitly coordinate on a price is reduced the greater the number of firms involved. Turning to market share asymmetries, they show how collusion is most likely with perfect symmetry, but becomes increasingly less likely as the two firms' shares diverge. The intuition here is that a smaller market

9. J. Tirole, The Theory of Industrial Organization, 240 (1988).

10. E. H. Chamberlin, The Theory of Monopolistic Competition, 48 (1933).

11. M. Ivaldi, P. Rey, P. Seabright, & J. Tirole, The Economics of Tacit Collusion, (report for DG Comp, European Commission, 2003) *available at* http://ec.europa.eu/comm/competition/mergers/studies_reports/the_economics_of_tacit_collusion_en.pdf.

share reduces the profitability of sticking to the collusive price. However, there is an important caveat on asymmetries—as Ivaldi et al note, market shares are endogenous, and what really matters are the causes of the asymmetry. For example, if the asymmetry derives from a fundamental asymmetry between the firms in costs, then the high cost (low share) firm has more to gain from undercutting and less to fear from retaliation.

In recent years, the theoretical literature has explored various possibilities on the causes of asymmetry: Rothschild[12] on costs, Compte, et al[13] on capacity, Kühn[14] on the number of products, and Vasconcelos[15] on capacity/costs. Although the details of these models vary, the underlying mechanism always works through the asymmetry this causes in the firms' incentives to collude/punish/deviate.

Although the general message that emerges from this literature is that asymmetries reduce the likelihood of collusion, it is clear that any rigorous test of the theory should seek to identify the causes of asymmetry in particular markets. In the current paper, however, because we are more concerned with general stylized facts that might apply across industries, we are almost inevitably reliant on using observed market shares to deduce the degree of asymmetry.

> BUT ONE MIGHT ALSO ARGUE THAT THERE IS A FUNDAMENTAL DISCONNECT BETWEEN TREATING COLLUSION AS AN OUTCOME OF A NONCOOPERATIVE GAME AND THE ANTITRUST CONCEPT OF COLLUSION.

Finally, it should be stressed that models within this genre are presented as models of tacit collusion. However, as hinted earlier, it is not uncommon to use the same model to derive predictions and explanations relating more to cartels. Indeed, it is not difficult to find examples, even in the best textbooks and articles, where the authors(s) appear to use the terms cartels and tacitly collusive groups interchangeably. This might be explained simply by rather sloppy use of language, but one might also argue, along with Martin,[16] that there is a fundamental disconnect between treating collusion as an outcome of a noncooperative game and the antitrust concept of collusion.

12. R. Rothschild, *Cartel Stability When Costs are Heterogeneous*, Int'l J. Indus.Org., 17, 717-734 (1999).

13. Compte, et al *supra* note 6.

14. K.U. Kühn, The Co-Ordinated Effects of Mergers in Differentiated Products Markets (University of Michigan, Legal Working Paper Series 34, 2004).

15. H. Vasconcelos, *Tacit Collusion Cost Asymmetries and Mergers*, 36 (1) Rand J. Econ. 39-62 (2005).

16. S. Martin, *Competition Policy, Collusion and Tacit Collusion*, Int'l J. Indus.Org, 24, 1299-1301 (2006).

III. Depicting Market Structure: The Oligopoly Triangle

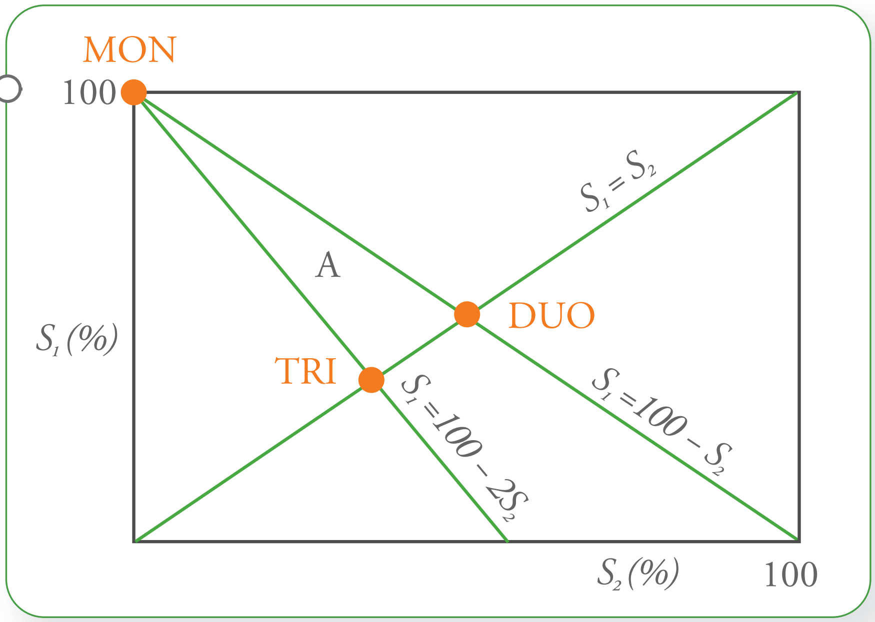
The exposition of the remainder of this paper is considerably eased by introducing the following graphical device as a way of summarizing and comparing market structures. For a given market in which there are N firms, this entails plotting the market share of the number one ranked firm against that of the number two ranked firm, as in figure 1. We explain the interpretation of this diagram first where N≤3, and then for N>3.

A. LITERAL TRIOPOLY

In a market with N=3, with firms ranked by the size of their market shares, S1, S2 and S3, the point (S2, S1) must provide a complete characterisation of the structure of that market.

Figure 1

The Oligopoly Triangle for Literal Triopoly



As shown in Box 1, by construction, the point (S2, S1) must lie within the A triangle, with the three corners identifying the three limiting market structures: perfect monopoly (MON), symmetric duopoly (DUO), and symmetric triopoly (TRI). Outside these extremes, location within the triangle represents: (i) the level of concentration (if measured by S1+S2, this is higher for points *closer* to the downward sloping diagonal (S1=100-S2)), (ii) the degree of size symmetry between S1 and S2 (the distance *from* the upward sloping diagonal, S1=S2, hereafter referred to as the symmetry diagonal), and (iii) the degree of size asymmetry between S2 and S3 (distance *along* the symmetry diagonal).

Box 1

Proof that (S_2, S_1) lies in the A triangle, with corners, MON- DUO-TRI

Box 1

Given that

(i) firms are ranked by size, $S_1 \geq S_2$, and the point must therefore lie on or above the symmetry diagonal, $S_1=S_2$

(i) the sum of shares can not exceed 100% , $S_1+S_2 \leq 100$, so the point can not lie outside the downward sloping diagonal, $S_1=100-S_2$

(i) firms are ranked by size, $S_2 \geq S3$. Thus, by trivial manipulation, $S_2 \geq 100–S_1-S_2$, i.e. $S_1 \geq 100–2S_2$. So the point must lie no lower than the line $S_1=100–2S_2$.

B. WITH MORE THAN THREE FIRMS

Interpretation of the triangle becomes less clear-cut when there are other firms in the market, but it remains true that the location of any point reveals both the level of two-firm concentration (S1+S2) and the degree of asymmetry between S1 and S2. Although there is now an indeterminacy on the relative sizes of S3 and all other smaller firms, we can add some further insights by inserting two additional lines to the diagram (figure 2).

Denoting the combined market shares of all other smaller 'fringe' firms by F, then the point will only remain within the original A triangle if S2≥ F+S3[17]. Hereafter, we refer to this as the 'literal triopoly' triangle: a sufficient, but not necessary, condition for location within A is that the market is a literal triopoly—other structures, with relatively small fringes (in the above sense) will also be located within A.

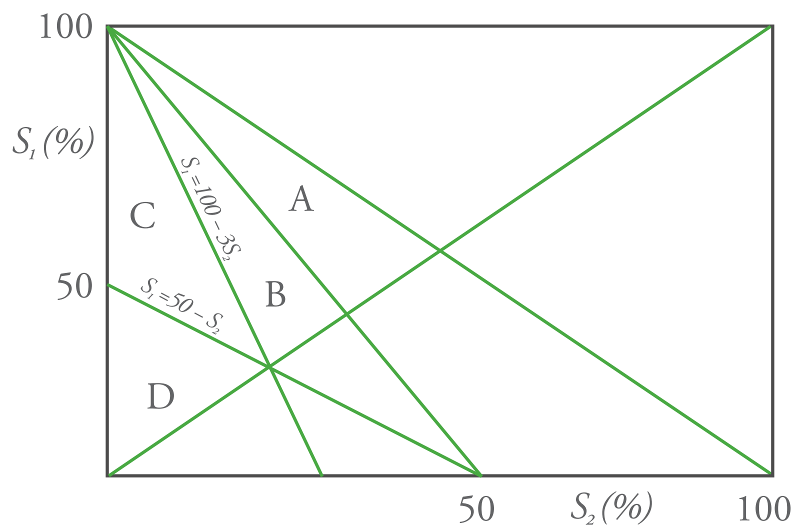
Second, for all other points lying below A, there must be a non-empty fringe F which is reasonably large. In general, as F becomes larger, we will tend to move towards the origin. In fact, we can be a little more precise. As proved in Box 2, all points in the D area close to the origin refer to markets where F is 'very large' in the sense that F≥S1; while all points in the C area refer to markets where F at least exceeds the size of S2. Note, however, that both statements refer to sufficient conditions, meaning that we can not exclude the possibility of F>S1 even

17. Proof: since S1+S2+S3+F=100, then S1≥100–S2 requires that 100–S2–S3–F≥100–2S2, i.e. S2≥F+S3.

in C or F>S2 even in B—it depends also on the size of F relative to S3. For the same reason, the interpretation of area B is even more indeterminate, although, in general, it is likely that only markets with relatively small F will qualify for inclusion in B.

Figure 2

The Oligopoly Triangle Beyond 3 Firms



Box 2

Box 2
Proof that triangles C and D denote approximate magnitudes of the fringe, F

Given that $S_1+S_2+S_3+F=100$,

(i) if $F \geq S_1$, then $100-S_1-S_2-S_3 \geq S_1$, and, since $S_2 \geq S_3$, sufficient condition for this is $100-2S_1-2S_2 \geq 0$ or $50-S_2 \geq S1$, i.e. in area D

(i) if $F \geq S_2$, then $100-S_1-S_2-S_3 \geq S_2$, and, since $S_2 \geq S_3$, sufficient condition for this is $100-S_1-3S_2 \geq 0$ or $100-3S_2 \geq S_1$, i.e. in areas C and D

In summary then, we suggest that this 'oligopoly triangle' provides a useful first-blush way of summarizing and comparing market structures between antitrust cases. The rest of this article puts the triangle through its paces in three applications, each designed to explore the conditions under which collusion might occur. In each case, closeness to the symmetry diagonal will reveal the degree of symmetry between the two largest players and distance from the origin will give a guide to the size of the fringe, with a ranking of the areas D>C>B>A denoting descending order of the minimum size of F.

IV. An Analysis of Coordinated Effects Mergers as an Indirect Strategy for Observing Tacit Collusion

As explained in the introduction, the purpose of Davies, et al[18] was to explore the conditions under which tacit collusion might arise by looking through the eyes of an antitrust authority (in this case the European Commission) in the one area of policy where its decisions reveal its views on tacit collusion, namely those mergers which (may) have coordinated effects (collective dominance). Throughout, we use the terms 'collective dominance' and 'coordinated effects'[19] synonymously—both refer to mergers where it is anticipated that the firms remaining in the market post-merger (including the merged firm) would be likely to coordinate their actions. Clearly, no CA could ever allege that such coordination would amount to overt collusion—to block a merger on such an interpretation would be tantamount to asserting that firms would act illegally, post-merger! Rather, we take it as given that coordination in this context can only refer to 'tacit collusion'.

CLEARLY, NO CA COULD EVER ALLEGE THAT SUCH COORDINATION WOULD AMOUNT TO OVERT COLLUSION—TO BLOCK A MERGER ON SUCH AN INTERPRETATION WOULD BE TANTAMOUNT TO ASSERTING THAT FIRMS WOULD ACT ILLEGALLY, POST-MERGER!

At the heart of Davies, et al's paper is a very simple model of decision making by a CA. It assumes that, when deciding whether a given merger should be allowed to proceed or require remedies or prohibition, the CA considers all markets in which there are overlaps between the merging parties. For each market,

18. S.W. Davies, M. Olczak, & H. Coles, Tacit Collusion, Firm Asymmetries and Numbers: Evidence From EC Merger Cases (CCP, Working Paper No. 07, 7, 2008) [hereinafter "Davies et al, 2008"].

19. We employ the European terminology, Collective Dominance, merely because our data derive from European cases over the period when this was the European Commission's common parlance for coordinated effects. Since the revision to the merger regulation in 2004, 'coordinated effects' has become common terminology, even within Europe.

it chooses between nonintervention and declaring a theory of harm, and, if the latter, whether it is single or collective dominance (unilateral or coordinated effects). In coming to its decisions, the CA follows its own merger guidelines[20] and therefore takes into account: (i) the potential market shares and asymmetries of the leading players post-merger;[21] and (ii) a checklist of other market conditions, including barriers to entry, buyer power, spare capacity, and transparent prices —these are referred to as the X conditions. The CA is then assumed to employ a structural model (more precisely, a model of the structure of market shares) to decide between theories of harm, but this only comes into play if the checklist of X conditions is satisfied (e.g. high barriers, no buyer power). Although simple, this model faces a key empirical problem which must be solved if it is to be estimated on a sample of real world mergers. The checklist of market characteristics not only lists the status of near-necessary conditions but these characteristics are also difficult for the researcher to measure, or even proxy, in an objective manner. As explained presently, our solution is to identify the subset of mergers for which the X conditions are likely to be satisfied, and then estimate the structural model only for that subset.

A. THE SAMPLE MERGERS

In order to test this model, we assembled a sample of mergers for which there are good reasons to suppose that the Commission seriously contemplated collective dominance as a potential consequence of the merger. (In general, we can assume that single dominance is always contemplated in principle.) We drew from the full population of over 2,400 merger reports published by the Commission, 1990-2004. This is the period from the introduction of the European Merger Regulation ("ECMR") in 1990[22] up to its revision in 2004. We ended the period at 2004 in order to avoid potential complications from any structural break at the time of the revision.

Within the full population, all merger reports were word-searched for the use of one or more of the following phrases: *collective dominance*, *(tacit) collusion*, *joint dominance*, *oligopolistic dominance*, or *coordinated effects*. This identified 94 candidate mergers, but closer textual examination revealed that in 32 of these the above phrases were only used in a cursory manner—typically in a throwaway sin-

20. *See* http://www.usdoj.gov/atr/public/guidelines/hmg.htm for the U.S. guidelines, and http://eur-lex.europa.eu/LexUriServ/site/en/oj/2004/c_031/c_03120040205en00050018.pdf for the EU guidelines.

21. We follow the convention of all CAs by 'estimating' the hypothetical post-merger market share of the merging firms by simply adding their pre-merger shares.

22. These are all mergers with a European, as opposed to purely national, dimension, and which exceed specified turnover thresholds. See S.W. Davies & B.R. Lyons, Mergers and Merger Remedies in the EU: Assessing The Consequences For Competition (Edward Elgar, 2007) and S.W. Davies & M. Olczak, Assessing the Efficacy of Structural Merger Remedies: Choosing Between Theories of Harm? (CCP, Working Paper No. 08, 28, 2008) [hereinafter "Davies and Olczak 2008"] for more discussion of the ECMR and remedies.

gle sentence or short paragraph, revealing that the Commission had easily dismissed the possibility. This leaves 62 mergers in which the text of the Commission's report includes a non-trivial discussion of the possibility that the merger might lead to collective dominance in at least one of the markets involved in the merger. It should be stressed that this search process merely identified all mergers in which collective dominance was seriously considered as a potential problem in at least one market. As will be seen, in by no means all cases did the Commission eventually judge that the merger would lead to collective dominance.

Four key descriptive facts on this sample help set the scene:

1. Collective dominance evidently arises as an issue only very infrequently; in no year during this period does this sample account for more than 4 percent of the total of all mergers.[23]

2. Since we confine our attention to only those cases where there are market overlaps between the merging firms, all mergers were purely horizontal.

3. Nearly all of these mergers are multi-market, involving more than just one product market and, remembering the EU context, more than just one Member State. In total, the 62 mergers covered 456 different markets in which there were overlaps between the merging parties and for which there is useable data. Thus, the average merger covers seven different markets, but with a skewed distribution around the average. However, in nearly all mergers, the different markets covered are very closely related in product space. This is either because the merger relates to the same product market in different countries and/or even when there is more than one product market, they are closely related. In 54 of the mergers, all markets covered belonged to the same 4-digit industry, and seven involved only two 4-digit industries. (This becomes important below for our assumption of X-homogeneity.)

4. Post-merger, most markets have no more than two or (much less frequently) three major players. As an illustration, defining a 'significant' market share as more than 15 percent, then the sample includes: just one market in which there are five significant players; 12 quadropolies; 89 triopolies; 245 duopolies; and 97 monopolies. Of course, 15 percent is an arbitrary yardstick, but any plausible alternatives yield qualitatively similar conclusions.

23. From figures reported by M. Bergman, M. B. Coate, M. Jakobsson, & S.W. Ulrick *in* Comparing Merger Policies: The European Union Versus the United States (2007) (working paper on file with the authors), it would appear that coordinated effects is considered far more frequently as an issue by the FTC in the United States. Dick (*supra* note 1) reports that, between 1999-2003, the FTC successfully challenged 11 proposed mergers under a coordinated effects theory.

B. EC'S DECISIONS FITTED TO THE OLIGOPOLY TRIANGLE

The EC actually intervened in only 25 of these mergers: the merger was prohibited in four cases and allowed to proceed in 21 cases (subject to remedies in one or more markets). We argue that it is only in this sub-sample of mergers—where an intervention occurred in one or more markets—that it is possible to isolate the structural model of single and collective dominance by controlling for other important market characteristics (including barriers to entry, absence of buyer power, and price transparency) that are embodied in the necessary X conditions. This argument relies on an assumption referred to as *X-homogeneity*; all markets covered by a given merger are assumed to share the same X market characteristics. For example, if the market for large tin cans is characterized by high entry barriers in Germany, the same is likely to be true for small tin cans, as well as for tin cans in France. If this assumption holds, then the decision to intervene in some markets in a particular merger implies that the necessary X conditions have been satisfied—not only for those markets, but also for all other markets covered by the same merger. It then follows that intervention in some markets, but not others, can be explained by structural conditions rather than by X market characteristics. Of course, this can only be an approximation to reality, but Davies, et al[24] present detailed discussion and empirical evidence in support.

For this reason, we include only these 25 mergers, covering 222 different markets, for an in-depth empirical analysis of the EC's decisions: 29 involved intervention for collective dominance ("CD"), 89 for single dominance ("SD")[25], while in the remaining 104 no intervention was deemed necessary. Note then that, in each of these mergers, the Commission reveals that it has considered the possibilities of both SD and CD—typically for different markets in the same merger—and that it is common to find, for a given merger, noninterventions and interventions, as well as different types of interventions across markets.

Figure 3 plots the potential post-merger structures, differentiated by decisions, within the oligopoly triangle. In figure 3(a), for the sake of clarity, the scatters are not shown but represented using head counts of the number of interventions relative to the total number of markets in each of the four areas. The probability of intervention is highest (nearly three-quarters) in the literal triopoly area A, and very low in the large-fringe area D (only 8 percent). The intervention rate is now less than one per cent.

24. Davies et al 2008, *supra* note 18 at 36-8.

25. The collective dominance decisions also include 15 markets in which a group of firms were 'structurally linked' in some way (usually shareholdings). In these cases the EC essentially views these firms as a single entity, and we combine the shares of the linked firms and count them as SD decisions.

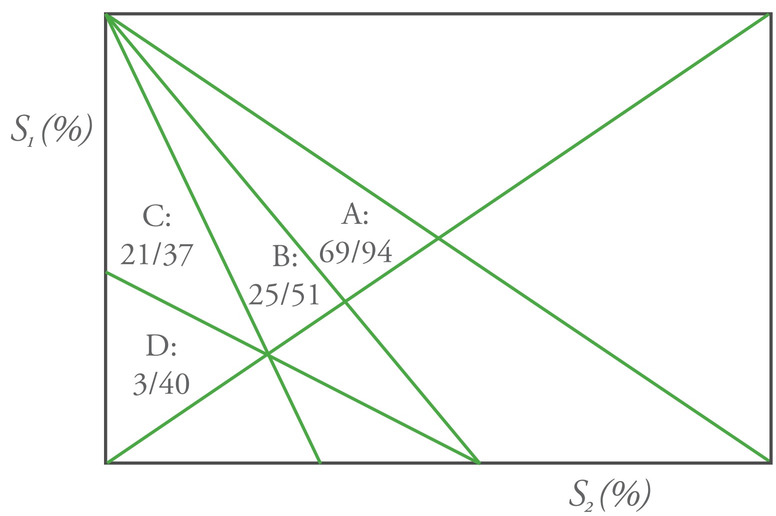


Figure 3a

Fitting the Sample Mergers and Markets to the Triangle

(a) Number of Interventions by Area

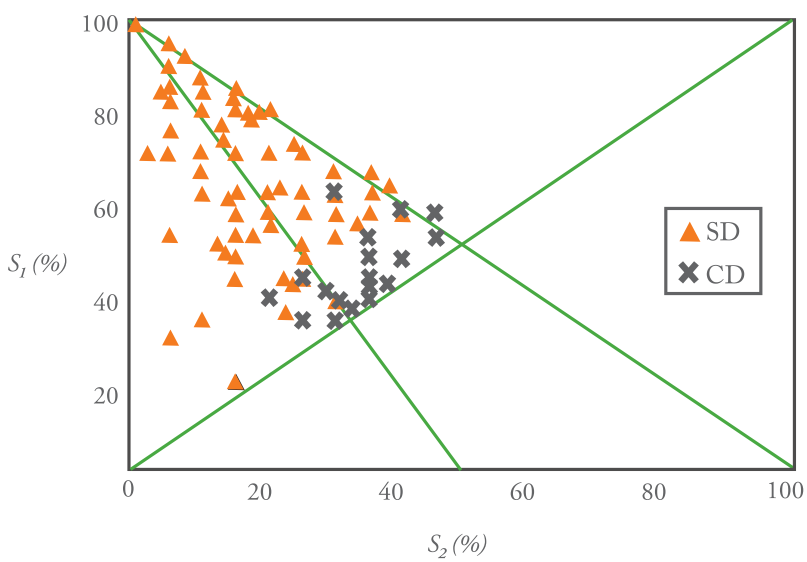


Figure 3b

Fitting the Sample Mergers and Markets to the Triangle

(b) By Types of Interventions by Area

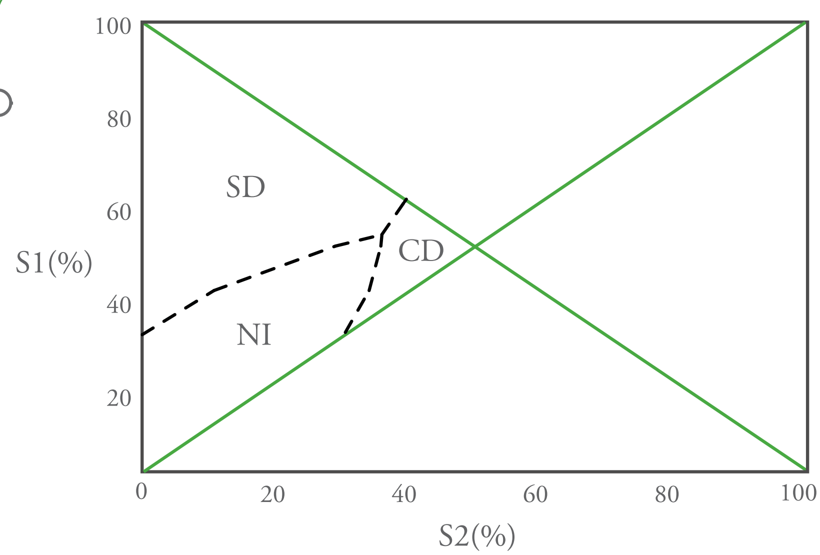
Figure 3(b) shows the full scatter of points, but for the intervention markets only (i.e. now excluding markets without intervention). This clearly demonstrates that CD only occurs within a narrow band just above the symmetry diagonal. It is most common in A, and to a lesser extent B, but extremely rare otherwise. It appears the European Commission requires, as a necessary condition for collective dominance, that the joint share of the two largest firms be high and that their shares be fairly symmetric. On the other hand, single dominance decisions occur with greater incidence in all areas except D, and typically with pronounced asymmetry between S1 and S2, especially in area A.

C. ECONOMETRIC MODEL

In Davies et al, this is formalized using a multinomial logit econometric estimator, in which there are three outcomes (NI, SD, and CD) and there are two simple market share explanatory variables—the sum and ratio of the market shares of the two largest firms: SUM (S1+S2) and RATIO (S2/S1). These two variables are both strongly significant at the 99 percent level in both the equations for SD and CD. They also have the expected signs, indicating that interventions are more likely in concentrated markets (high SUM) and, for CD, in symmetric markets (high RATIO) but for SD in asymmetric markets (low RATIO). The model successfully explains 79 percent of all 222 decisions. Figure 4 displays the predicted decisions graphically, and Table 1 shows the different possible outcomes implied, depending upon the size of the number one and two ranked firms.

Figure 4

Predicted Theories of Harm as Revealed in EC Merger Decisions (Davies, et al, 2008)



Thus, when the number one ranked firm has a very large post-merger market share (>65 percent), the model predicts that the Commission will always decide SD. However, at lower values for S1, the decision also depends crucially on the size of S2 post-merger. For example, at S1=55 percent, while the Commission will always judge the structure to entail dominance, this will be single dominance if S2 is relatively small, but collective dominance if S2 is relatively large. Perhaps most interesting is where S1=45 percent—here all three outcomes can occur, depending on S2: where S2 is 'large', the EC opts for a CD decision, where S2 is 'small', it opts for SD, but for intermediate S2, it opts for NI. This implies that there are some cases where the number two firm is considered to be sufficiently

Table 1 **Possible Outcomes at Different Sizes for S1 and S2**

	Different outcomes according to size of S2		
S1 (%)	NI	CD	SD
45	$18<S2<35$	$35<S2$	$S2<18$
55		$37<S2$	$S2<37$
65			For all S2

large to counteract the otherwise dominant position of the leader, but not sufficiently large to result in collective dominance (i.e. the resulting size asymmetry rules out tacit collusion.)

As a short policy postscript, we can report how things have changed beyond the above time period. After the Commission revised its Merger Regulation in May 2004 (up to mid 2007), there were 13 mergers which satisfy our criteria for inclusion in the above sample: non-trivial discussion of coordinated effects (which has now displaced collective dominance as preferred terminology) and a remedy imposed in one or more markets. Strikingly, of the 274 markets covered by these mergers, in only two has the Commission justified an intervention citing the possibility of coordinated effects.[26] Thus, while the proportion of all mergers matching the criteria for inclusion remains in the region of two percent, indicating an unchanged willingness to contemplate and discuss coordinated effects, the probability that the judgment will actually invoke coordinated markets in any market is now less than one percent.

Moreover, even in these two cases, the decisions are equivocal:

> "The Commission does not rule out the possibility that the merger, besides producing non-coordinated effects … may also lead to a weakening of competitive pressure as a result of coordinated effects"[27]

and

26. In five others, it intervened on the basis of structural links—a not dissimilar proportion to that found pre-2004.

27. Commission Decision, Case COMP/M.3916, T-Mobile Austria/Tele.ring, (Apr. 26, 2006) para 127.

"The [merger] would with high probability lead to non-coordinated effects . . . [and] the Commission has serious doubts that [it] may even lead to . . . coordinated effects"[28]

In other words, not only is the Commission now wary of invoking coordinated effects, but also, rather strangely and in stark contrast to its practice up to 2004, it couples coordinated effects with unilateral effects.

In other words, not only is the Commission now wary of invoking coordinated effects, but also, rather strangely and in stark contrast to its practice up to 2004, it couples coordinated effects with unilateral effects. One might interpret this either as a reluctance to come off the fence, or as a belief that both effects might occur simultaneously. The post 2004 period is the subject of our ongoing research.[29]

V. Experimental Literature

A cynic, when faced with the above results and asked "How have they advanced the understanding of tacit collusion and collusion in general?" might quite justifiably respond: "Nothing, unless one can trust in the ability of a CA to correctly identify such markets." But putting aside disbelief for the sake of the argument, our results on firm numbers and size asymmetries provide some intriguing parallels and contrasts with findings from the adjacent experimental and cartel literatures. We first consider the experimental literature, and here a consensus appears to be emerging.

The generic advantage of experimental as compared to real-world, fieldwork-based econometric, empirical work is that one can control for potentially confounding factors. In the context of mergers, for example, Fonseca and Normann[30] suggest that "economic conditions, cost gains derived from the merger, barriers to entry or industry maturity" might all obscure any underlying "strategic effects of mergers on pricing behaviour." In the context of tacit collusion, this advantage is arguably even more pronounced given the difficulties in unambiguously identifying tacit collusion in real world settings. For the experimenter, this problem is side-stepped by defining a tacitly collusive outcome as any in which

28. Commission Decision, Case COMP/M.4141, Linde/BOC (June 6, 2006), para 179-180.

29. Davies, et al 2008, *supra* note 18.

30. M. A. Fonseca & H. T. Normann, *Mergers, Asymmetries and Collusion: Experimental Evidence*, 118 Econ. J. 387- 400 (2008) [hereinafter "Fonseca"].

"prices [are] above Nash prices"[31], where the former are experimentally observed and the latter set by the conditions of the experiment.

To date, experimentalists have devoted far more attention to firm numbers than to asymmetries. Huck et al conducted a meta analysis on 19 previous studies, 1963-2003, which used Cournot experiments. (See also Engel[32] for a wider, but more loosely focused meta analysis). These studies involved between two and five symmetric firms and satisfied certain requisites including: no communication among participants, fixed groups interacting repeatedly, homogeneous products, and usually linear demand and costs. Measuring collusion (inversely) by the ratio of the experimental 'industry' output to the analytical Cournot-Nash outcome, they found a statistically significant (at a five percent level) inverse correlation between firm numbers and collusion. However, on average, it is only in two-firm markets that actual output is less than the Cournot level, leading to their headline finding: "Collusion sometimes occurs in duopolies (but) is very rare in markets with more than two firms."[33]

Huck, et al also conduct some experiments of their own, within a more unified framework than is possible in a meta analysis. These corroborate the meta analysis—collusion sometimes occurs when there are only two firms, but never in markets with four or more. Even in three firm markets, the average outcome is close to the Nash equilibrium. Thus, their message is clearly captured by their title: "Two are few and four are many."[34] This title is clearly a deliberate implicit reference to Selten's seminal argument that "four are few and six are many." (Although he coined this in a slightly different context, Selten's paper appears to have been a major stimulus to much of the experimental literature.)

The experimental literature on asymmetries is much thinner. Huck, et al report only two in the Cournot setting, the most relevant for current purposes being Mason, et al.[35] In their experiments, outputs were found to be significantly higher (and thus prices lower) where firms have asymmetric, rather than symmetric, costs. Their explanation appeals to the greater difficulties in coordination where firms are dissimilar.

---

31. S. Huck, H. T. Normann, & J. Oechssler, *Two Are Few and Four Are Many: Number Effects in Experimental Oligopolies,* 53 J. Econ. Behav. & Org. 436, fn.4 (2004) [hereinafter "Huck"].

32. C. Engel, How Much Collusion? *A Meta-Analysis on Oligopoly Experiments*, 3/4 J. Competition L. Econ. 491-549 (2007).

33. Huck *supra* note 31 *at* 440.

34. *See* R. Selten, *A Simple Model of Imperfect Competition, Where Four Are Few and Six Are Many*, 2 Int'l J. Game Theory 141-201 (1973).

35. C. F. Mason, O. R. Phillips, & C. Nowell, *Duopoly Behaviour in Asymmetric Markets: An Experimental Evaluation*, 74 Rev. Econ. & Stat. 662-670 (1992).

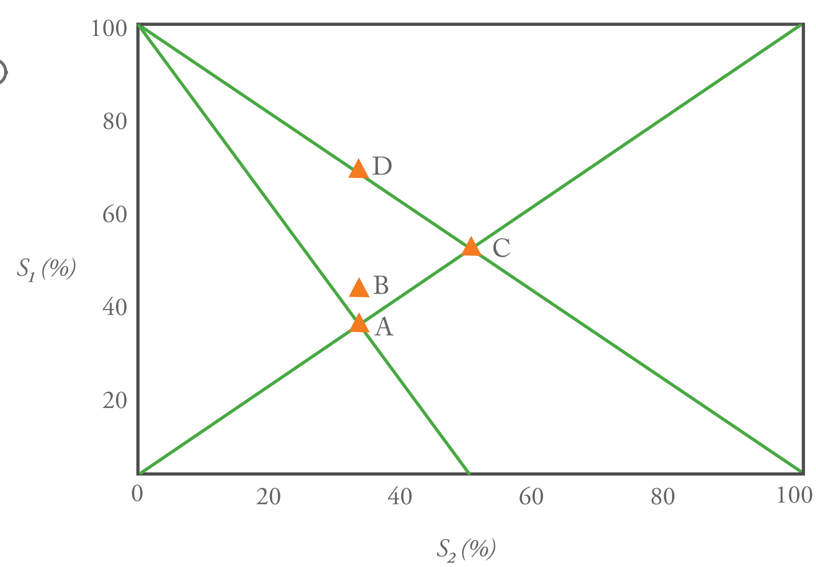
More recently, asymmetries have been explored by Fonseca and Normann[36] in an experimental setting which closely follows the Bertrand-Edgeworth model employed in Compte, et al[37], referred to in section 2. Here, firms set prices (as opposed to setting quantity in Cournot) but subject to potential capacity constraints. The range of alternative market structures considered is admittedly limited, either just two or three firms, and in each case with either symmetric or asymmetric capacities, but this allows for easy interpretation in terms of our oligopoly triangle (figure 5).

Table 2 Alternative Market Structures in Fonseca and Normann's Experiments

		Firm Capacities		
	Structure	S1	S2	S3
A	3 Symmetric	134	134	134
B	2 Asymmetric	160	134	108
C	2 Symmetric	201	201	
D	2 Asymmetric	268	134	

Figure 5

Fonseca and Normann's Market Structures in the Triangle



36. Fonseca *supra* note 30.

37. Compte, et al, *supra* note 6.

Their experiments are repeated, with fixed groups, over 30 period sessions, thereafter subject to random stopping. In all cases, demand and total capacity are identical. Prices are only compared from period 11 onwards, to allow for learning effects within a repeated game. They report the following ranked weighted mean prices across the four treatments:

$p_C > p_D > p_A > p_B$

In other words, price is highest with two equal sized firms (C) and lowest with three unequal sized firms (B). These results are consistent with what we might loosely refer to as the 'standard' predictions on the likelihood of collusion. Thus,

- Reductions in firm numbers lead to increased prices, both in the symmetric and asymmetric cases, i.e. $p_C > p_A$ and $p_D > p_B$
- Asymmetry leads to reduced prices, holding numbers constant at either 2 or 3, i.e. $p_C > p_D$ and $p_A > p_B$

As it happens, this ranking is not quite as predicted by the very specific model of Compte et al[38], and that should be the subject of future research. But this need not distract us here from the key conclusion for our purposes—asymmetry appears to have a definite collusion-dampening role, even where there are just two or three firms.

Of course, as is true for any area of economics, the experimental methodology is not without its limitations. In the current context, it can be argued that it is very difficult to emulate the real world conditions under which tacit collusion can occur with experimental subjects who are often largely inexperienced (in Fonseca and Normann, students from the home university). Thus we should be extremely wary in concluding that tacit collusion in the real world is unlikely with more than two players, and/or with asymmetries. Arguably, real world factors, such as mutual trust and familiarity, fostered over quite long periods of time, are at the heart of tacit collusion. But these factors are very difficult to simulate in a laboratory environment, especially with the fairly trivial prizes given even in the best-funded laboratory. While it is true that the experiments reported above were repeated over many simulated time periods (allowing for learning) with subjects not randomly matched (typically cooperation is never observed experimentally with random matching), a future research agenda must surely

38. In Compte, et al, the one shot equilibrium would imply, alternatively, $p_D > p_C > p_B > p_A$, because the mixed strategy Nash equilibrium average price is decreasing in the capacity of the smaller firms. This is because, with greater capacity for the smaller firms, the largest firm is more likely to opt for a lower price. On the other hand, in their collusive equilibrium, the critical discount rate increases with the capacity share of the largest firm. If this translates into a lower collusive price, then this predicts: $p_A > p_B > p_C > p_D$. The intuition here is that because the larger is the capacity share of the largest firm, the less severe is the punishment, which makes collusion harder to sustain.

include a deeper investigation of communication among participants[39]. Certainly, we know from other areas of experimental work that the probability of cooperative outcomes increases rapidly as greater communication is allowed (Huck, et al[40]). Just how much communication, and of what type, is contentious, but this is at the heart of the antitrust debate about what constitutes proof of overt, as opposed to tacit, collusion.

JUST HOW MUCH COMMUNICATION, AND OF WHAT TYPE, IS CONTENTIOUS, BUT THIS IS AT THE HEART OF THE ANTITRUST DEBATE ABOUT WHAT CONSTITUTES PROOF OF OVERT, AS OPPOSED TO TACIT, COLLUSION.

VI. Cartels

Unsurprisingly, the empirical literature on cartels is far more extensive than that on tacit collusion since cartel cases are unambiguous and often publicly documented by competition agencies and others. However, caution is needed here, too, because we only observe detected cartels, and sample selection bias can not be ruled out with respect to market structures. Nevertheless, this literature tells a very different story from that of the two previous sections.

First, on firm numbers in general, the evidence may suggest that "cartels are more likely if concentration is large and/or there are relatively few firms in the market/industry" (Grout and Sondregger).[41] However, exceptions are sufficiently frequent for Levenstein and Suslow[42] to refer to "the lack of a clear empirical relationship." They offer as possible explanations: sample selection bias (detected cartels may not be a random sample of the population), the potential for a counteracting reverse causality (the softer competition implied by cartels may allow more firms to survive), and, most tellingly for present purposes, the possibility that small numbers markets may be able to tacitly collude as an alternative to cartel formation. A rough indication of the typical number of firms involved in some of the most prominent existing studies of cartel studies is shown by Table 3.[43]

Second, rather surprisingly, the empirical cartel literature appears to have paid much less systematic attention to asymmetries. However, a casual reading of CA

39. A related experimental literature is already emerging which contemplates the possibility of participants switching from overt to tacit collusion in the face of a leniency program, see for example J. Hinloopen & A. Soetevent, From Overt to Tacit Collusion: Experimental Evidence on the Adverse Effects of Corporate Leniency Programs (2008) (mimeo *available at* the University of Amsterdam).

40. Huck *supra* note 31, at 438 Summary 1.

41. P. Grout & S. Sonderegger, Predicting Cartels (Office of Fair Trading, Economic Discussion Paper OFT773, 15, 2005) [hereinafter "Grout"] *available at* http://www.oft.gov.uk/shared_oft/reports/comp_policy/oft773.pdf.

42. M.C. Levenstein & V.Y. Suslow, *What Determines Cartel Success?,* XLIV J. Econ. Literature 58 (2006).

43. Derived from *Id.*, Tables 4 and 5. De's database is referred to in the text.

Table 3 Number of participants in Cartels

Numbers	Hay & Kelley	Fraas & Greer	Posner	Levenstein & Suslow	De (EC Cartels)
Mean	7.25	16.7	29.1		6.1
Median	7	8	6 to 10	8	5
Fewer than 10	79%	60%	64%	63%	85%

reports on real world cartels suggests that size asymmetries among cartel members can sometimes be quite pronounced. From their sample of 24 EC case studies, Grout and Sondregger[44] suggest that "we clearly observe a considerable heterogeneity in the market shares held by cartel members," citing four examples in particular: Citric Acid, Methionine, Far Eastern Trade Tariff and Surcharge Agreement (EC), and Ferry Operators-Currency Surcharges.

Some of our own ongoing work in progress[45] provides a useful overview on both firm numbers and asymmetries. The database relates to the 41 successfully prosecuted EC cartels, 1990-2006, for which it has been possible to extract the required data on market shares, and excludes all cases involving associations of firms. As can be seen from the last column of Table 3, the median number of firms in this sample was 5 and the mean 6. Both are slightly lower than the comparable statistics from the previous studies shown in the table, but nevertheless confirm that cartels typically entail larger firm numbers than in either our own work on mergers or in the experimental literature: only eight cartels involved just two firms, three involved three, 22 involved four to seven firms, and 12 involved eight or more firms.

However, our findings on the extent of size asymmetries within this sample are more intriguing.

However, our findings on the extent of size asymmetries within this sample are more intriguing.

Again, the oligopoly triangle (figures 6) provides a useful quick guide, with the axes, as before, depicting the relative sizes of the two largest players within each cartel. Here, it is important to stress that, for this purpose, we show market shares as percentages of the total size (typically sales) of the cartel. To the extent that cartel members do not account for the entire market, this overstates firms' market shares. In this sample, the median cartel accounts for about 90 percent of the market, but in some cases it is much lower: coverage is less than 70 percent

44. Grout *supra* note 41.

45. This is in collaboration with Oindrila De, who has been responsible for the careful reading of the large number of EC case documents necessary to generate the estimates drawn on here.

in seven cases. Obviously, if expressed as genuine market shares, a number of the observations would be moved closer to the origin.

For simplicity and comparability with the earlier figure (3a), figure 6(a) first merely reports the counts across the four areas of the diagram. This is in marked contrast with the earlier story of figure 3: less than one third of cartels lie within the 'literal oligopoly' A area, and about a quarter are located in the unconcentrated C and D areas: clearly, explicit collusion can occur within relatively unconcentrated groups of firms.

Figure 6a

Size Asymmetries within EC Cartels

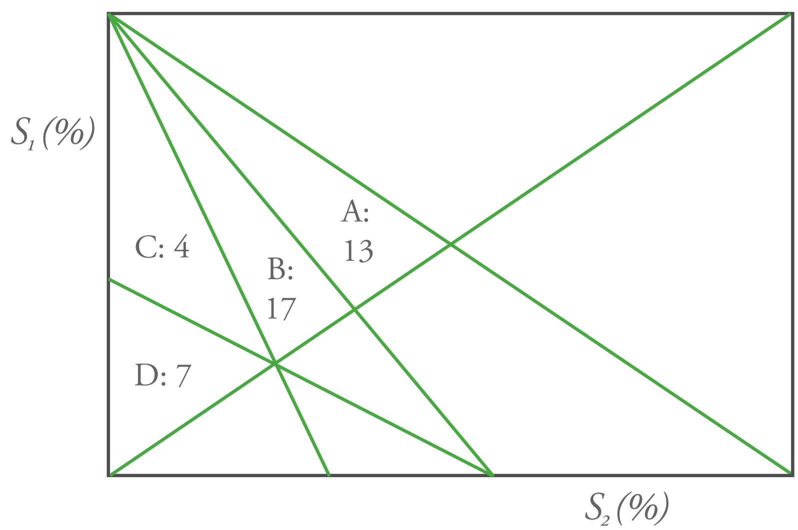
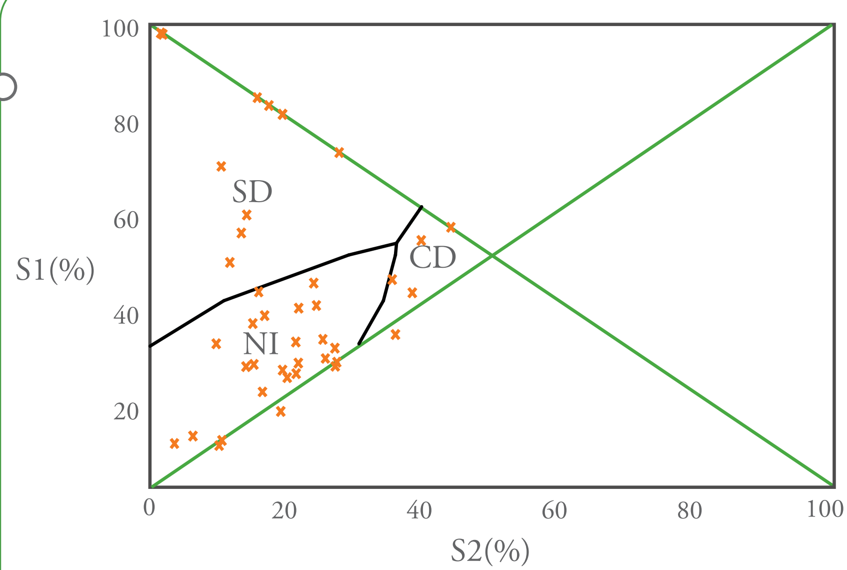


Figure 6b

Cartel Scatter with Merger Predictions Superimposed



The contrast is sharpened in figure 6(b) which shows the full scatter, with the fitted curves from our earlier merger analysis superimposed. This focuses more precisely on asymmetries, and it unambiguously establishes that size inequalities are far more pronounced than those found in our earlier coordinated effect mergers. Only five cartels lie within the region of collectively dominant mergers; 11 present structures which would have been judged to be singly dominant had they been the outcome of a merger, and the remaining 25 (i.e. 60 percent) would not have been judged to involve either collective or single dominance.

Generally, although it is apparent that a sizeable proportion of cartels do present rough symmetry between S1 and S2 (lying fairly close to the symmetry diagonal), a sizeable proportion do not and this calls into question whether symmetry of market shares is a pervasive feature of real world cartels.

Indeed, it is tempting to speculate from this figure that we can identify three broad types of cartel structure:

- "Tacit-collusive compatible"—in the very limited sense implied by this paper, i.e. consistent with structures identified by the EC as conducive to coordinated effects mergers. Such cartels amount to broadly symmetric concentrated duopolies.
- "Dominant leader"—with the largest firm accounting for roughly 50 percent of the cartels' sales, and its largest rival typically much smaller (say 20 percent or less).
- "Unconcentrated"—with neither of the largest firms accounting for much more than 40 percent or 30 percent, and usually much less. This type is fairly heterogeneous, including five or six cartels which might be categorized as triopoly or quadropoly, but the other 20 entailing very significant fringes.

It would be imprudent to push this typology too far—there is undoubtedly some fuzziness at the edges of the three types, and there are, no doubt, imprecisions in the raw market share data on which it is based. However, we believe the typology provides a very convenient framework within which to draw some of our main implications and conclusions.

VII. Implications and Conclusions

Our purpose has been to confront our previous findings on the market structural characteristics of coordinated effects EC merger cases, as a proxy for tacit collusion, with what is known from two quite separate empirical literatures—on the one hand, experimental research on tacit collusion, and, on the other hand, the observed market structures of some real world cartels. Underlying this purpose is a desire to assess the empirical similarities between explicit and tacit collusion in the light of a general practice among economists to assume that both phenomena can be understood by what is essentially the same model—the repeated game.

So what have we learned, and what does this suggest for future research? Putting aside some very important caveats for the moment, our results suggest the following:

- EC coordinated effect merger decisions and our review of the experimental evidence suggests that tacit collusion is rare with more than two firms, and without symmetry.
- EC cartel cases suggest that explicit collusion very often involves more than two firms, a 'typical' number might be five or six, but very often it is much more. Size asymmetries are often quite pronounced within cartels.

This can be developed by drawing on the tentative typology of the previous section as follows. In answer to one of our opening questions, are tacit and overt collusion substitutes? The answer may be yes, but only for a small subset of cartels—the tacit collusive compatible subset that involves just two, roughly equal players. For this subset, a further set of questions follow naturally: why did the firms involved opt for an illegal cartel, when a similar outcome might have been attainable tacitly? Did they emerge historically in the wake of a break-down in a tacit collusion? (This would be consistent with Harrington's[46] evidence that prices tend to fall prior to the formation of a cartel.) Analogously, once broken, is overt replaced by tacit collusion?

ANALOGOUSLY, ONCE BROKEN, IS OVERT REPLACED BY TACIT COLLUSION?

Turning to the dominant firm subset again, was it the dominant firm who was the ringleader? If yes, then why did it choose to instigate a cartel, rather than relying on the cause of its dominance (perhaps a cost advantage or a first-mover advantage) to ensure compliance of its smaller rivals in a non-aggressive tacit understanding? Moreover, are the internal mechanisms employed in this type of cartel distinctively different from those observed in other cartels?

Finally, for the unconcentrated subset, the salient question is: How were these cartels able to form and survive, given that they exhibited neither of the generally expected characteristics of fewness and similarity?

These questions are as real for the academician as they are for the policy practitioner. Probably the most fundamental distinction between overt and tacit collusion is that meaningful explicit communication is possible in the former but not the latter. From the existence of the unconcentrated subset, it would appear that communication may often considerably extend the feasible boundaries for

46. Harrington *supra* note 4 at slide 44.

an aspiring cartel with respect to both fewness and asymmetries, but how is this reconciled with existing theory? The experimental papers we have reviewed above are meticulous in not allowing any communication among participants. Is it possible in future work to relax this restriction in a way that simulates the sort of informal information exchange which, while falling short of hard evidence recognized by the Court, might yet extend the boundary of structures within which tacitly collusive outcomes emerge? Returning to our own work on mergers with coordinated effects, is it simply that the CA (in our case, the European Commission) has been over-cautious in employing the coordinated effects theory of harm in its decisions?

We believe that each of the questions posed above merits further research, and therefore endorse Harrington's second sentence below:

> "…Having drawn this distinction between explicit and tacit collusion, I am disappointed to say that, due to inadequacies in the underlying theory, the ensuing analysis will largely ignore it. Nevertheless, it is important to keep this distinction in the back of our mind and hopefully it'll move to the front of our collective mind in future research."[47]

We hope that the current paper will help nudge the topic closer towards the front of the agenda.

47. J. E. Harrington, Jr., *Detecting Cartels*, *in* HANDBOOK IN ANTITRUST ECONOMICS 4-5 (P. Buccirossi ed., 2008).

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

The Lawful Acquisition and Exercise of Monopoly Power and Its Implications for the Objectives of Antitrust

David S. Evans & Keith N. Hylton

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

The Lawful Acquisition and Exercise of Monopoly Power and Its Implications for the Objectives of Antitrust

*David S. Evans & Keith N. Hylton**

The antitrust laws of the United States have, from their inception, allowed firms to acquire significant market power, to charge prices that reflect that market power, and to enjoy supra-competitive returns. This article shows that this policy, which was established by the U.S. Congress and affirmed repeatedly by the U.S. courts, reflects a tradeoff between the dynamic benefits that society realizes from allowing firms to secure significant rewards, including monopoly profits, from making risky investments and engaging in innovation; and the static costs that society incurs when firms with significant market power raise prices and curtail output. That tradeoff results in antitrust laws that allow competition in the market and for the market, even if that rivalry results in a single firm emerging as a monopoly, but also prevent firms from engaging in practices that go out of bounds. The antitrust laws ultimately regulate the "boundaries" of the "game of competition." Three implications follow: First, the antitrust laws and intellectual property laws are based on similar policy tradeoffs between static and dynamic effects. Second, the antitrust rules have, all along, been based on this tradeoff and not on the effects of business practices on static consumer welfare in relevant antitrust markets. Third, one unintended consequence of the increased role of economics in antitrust analysis is to overemphasize the static considerations which are almost the sole focus of the economics literature considered by courts and competition authorities.

*David S. Evans is Executive Director, Jevons Institute for Competition Law and Economics, Visiting Professor, University College London; and Lecturer, University of Chicago Law School. Keith N. Hylton is Professor, Boston University Law School. They would like to thank Howard Chang and Richard Schmalensee for helpful comments and suggestions and Sokol Vako and Parit Sripakdeevong for research support.

I. Introduction

The antitrust laws of the United States have never prohibited a firm from having a monopoly as such or from enjoying the fruits of monopoly except in special circumstances. This observation is not new.[1] But its consequences for the objectives of antitrust, for the role of static versus dynamic competition in antitrust law, and for the debate over the tension between antitrust and intellectual property law are profound and underappreciated in the literature.

This article draws out the implications of the bedrock principle that neither monopoly nor its profits are unlawful. We highlight two. First, the U.S. antitrust laws recognize the role of "competition for the market" as a major source of innovation and monopoly profits[2] as the desirable rewards for entrepreneurship. Second, over the long run, the antitrust laws balance the benefits and costs of static and dynamic competition in the overall economy.

These two propositions pull some important additional implications in their wake. One is that there is no fundamental tension between the policies of antitrust law and intellectual property law; both balance the benefits and costs of static and dynamic competition for the economy as a whole. Another is that one cannot reliably appeal to the consumer-welfare objectives of the antitrust laws to rationalize legal tests based on examining short-run effects on price and output in relevant antitrust markets, although there may well be practical and operational reasons for doing so in the larger framework of antitrust analysis.

The article is organized as follows. Section II briefly summarizes the nature of the antitrust laws. As other authors have recognized, the antitrust laws are based on the presumption that society benefits from the competitive game among firms.[3] The antitrust laws provide some limited rules to prevent firms from playing this game in ways that could be harmful ultimately.

Section III documents that antitrust policy presumes that it is lawful to have a monopoly and to enjoy the fruits of that monopoly. It then draws out the implications of this principle for the tradeoff between static and dynamic efficiency and to the application of the antitrust laws for developing the "rules of the game."

1. *See e.g.,* United States v. Standard Oil, 221 U.S. 1, 62 (1911) ("[T]he statute … *by the omission of any direct prohibition against monopoly* in the concrete … indicates a consciousness that the freedom of the individual right to contract … was the most efficient means for the prevention of monopoly") (emphasis added). *See also* Philip Areeda & Herbert Hovenkamp, Fundamentals of Antitrust Law (3rd ed. 2004).

2. We use the term monopoly for convenience. It should be understood throughout as referring to firms that have significant market power under U.S. law or a dominant position under EC law.

3. *See* Massimo Motta, Competition Policy: Theory and Practice 17-19 (2004). For a discussion of how competition leads to welfare see *id.* at 39-53.

Section IV considers the antitrust laws of the European Community. Although the European Community ("EC") Treaty provides for the regulation of monopoly prices, the European Commission and the Community Courts have been reluctant over the last 50 years to invoke these powers. It is lawful in the European Community to have a monopoly and, by and large, to earn monopoly profits. Most countries follow U.S. or EC competition law[4] and therefore presume that monopolies and monopoly pricing are lawful per se.

Section V shows that antitrust and intellectual property policy share the same basic objectives. Tension arises mainly because they deal with the tradeoff between static and dynamic competition from different constitutional, legislative, and case law perspectives.

Section VI argues that one can think of antitrust law as following a two-step process. In the first step, antitrust policy considers the effect of practices on long-run economy-wide consumer welfare to assess where to draw the boundaries and thus which practices are clearly lawful or not. In the second step, antitrust policy considers whether particular practices near those boundaries are lawful or not based on a fact-intensive inquiry. The traditional competitive effects analysis of examining the impact of a practice on price and output in a relevant antitrust market provides a method for assessing these close calls.

Section VII argues that the increasing use of economic analysis in competition policy tends to shift the focus away from dynamic competition because most of the economic literature, dating back to the original Chicago work, is based on mathematical models of static competition. There is "static competition bias" that affects how economists analyze antitrust problems. This section also argues that the current industrial organization literature provides limited insights into the tradeoff between static and dynamic competition that is at the heart of how the courts (properly) think about the design of competition rules.

Section VIII makes some brief concluding observations. The article argues that the antitrust laws were designed to promote long-run economic welfare in the economy and have long recognized the importance of allowing firms to obtain monopolies and its rewards for achieving that objective. That has led the courts to establish both boundaries for the game of competition and rules for assessing whether these boundaries have been crossed. This article should not be read as arguing that recognizing the importance of dynamic considerations, in the foundations of antitrust law, necessarily provides a basis for moving those boundaries or modifying those rules in either the United States or European Community. But it does caution against relying on static economic analysis in determining where those boundaries should lie and in devising rules to assess whether those boundaries have been crossed.

4. Einer Elhauge & Damien Geradin, Global Competition Law and Economics (2007).

II. Competition Rules

The Supreme Court significantly shaped the antitrust laws of the United States during the first quarter century after the passage of the Sherman Antitrust Act of 1890.[5] This culminated in several classic decisions. *Trans-Missouri*,[6] in 1898, established that judicial interpretations of the Sherman Act would not be based on the common law of contracts in restraint of trade. *Standard Oil*,[7] in 1911, adopted the rule of reason test while *Chicago Board of Trade*,[8] in 1918, articulated the process for applying the rule of reason test.[9] *United States Steel*,[10] in 1921, clarified the limits of Sherman Act Section 2 in its application to monopolies. Some thought that the courts had taken too lenient a view on anticompetitive practices in the first two decades following the Sherman Act. That view led to the 1914 passage of the Clayton Act which proscribed particular practices including price discrimination, exclusive dealing, and tying under certain circumstances.

From their 1890 inception at the federal level, the antitrust laws soon evolved into a process for lightly regulating the competitive process.[11] Certain kinds of concerted action such as price fixing were prohibited. Other business behavior could be unlawful if it could be demonstrated that the firm had significant market power and engaged in practices that were seen as restricting competition. As a practical matter though, most businesses, including very large and powerful ones, could engage in an almost limitless range of practices that did not run afoul of the antitrust laws to make profits, fend off competitors, and increase their market shares.

One can see the role of the antitrust laws in the American economy in several ways. From 1890 to 1997, the U.S. Department of Justice ("DOJ") filed 1,355 civil antitrust cases, or about 13 per year.[12] One estimate suggests only about 20 percent of DOJ cases were for monopolization or exclusionary practices claims; the remainder concerned merger and horizontal per se claims.[13] That implies that

5. Many authors have examined the objectives of the antitrust laws by examining the history antecedents, the economic environment, and the legislative debate that led to the passage of the Sherman Antitrust Act of 1890. See the collection of papers THE POLITICAL ECONOMY OF THE SHERMAN ACT: THE FIRST ONE HUNDRED YEARS (Thomas E. Sullivan ed. 1991).

6. United States v. Trans-Missouri Freight Assn., 166 U.S. 290 (1897).

7. Standard Oil Co. v. United States, 221 U.S. 1 (1911).

8. Chicago Board of Trade v. United States, 246 U.S. 231 (1918).

9. The rule of reason and per se distinctions were further refined in *Trenton Potteries* case of 1927. United States v. Trenton Potteries Co., 273 U.S. 392 (1927). The boundaries of these rules were not clarified until United States v. Socony-Vacuum Oil Co., 310 U.S. 150 (1940). *See, e.g.,* KEITH N. HYLTON, ANTITRUST LAW: ECONOMIC THEORY AND COMMON LAW EVOLUTION 90-111 (2003).

10. United States v. United States Steel Corp., 251 U.S. 417 (1920).

11. John Vickers, *Abuse of Market Power*, 115 ECON. J. F224 (2005).

12. Joseph Gallo et. al., *Department of Justice Antitrust Enforcement, 1955–1997: An Empirical Study* 17 REV. INDUS. ORG. 75, 90 (2000). A further 82 were classified as having criminal as well as civil components.

roughly two cases a year involved monopolization or exclusionary practices claims. The same study found only about one-third to one-quarter of DOJ cases were filed against Fortune 500 firms.[14] The Federal Trade Commission ("FTC") filed 1,061 cases between 1915 and 1969, for an average of about 19 per year.[15] The U.S. antitrust enforcement agencies have engaged in relatively modest enforcement activities when viewed over long periods of time.

THE LIKELIHOOD THAT ANY PARTICULAR BUSINESS PRACTICE USED BY A FIRM WITH SIGNIFICANT MARKET POWER IS CHALLENGED IS ALMOST CERTAINLY QUITE SMALL.

The number of private antitrust cases that were filed varied from 452 to 1528 in the 29 year period from 1971-1999.[16] By way of comparison, the number of corporate tax returns varied from about 360,000 in 1926 to around 497,000 in 1947 to about 4.7 million in 1997.[17] Business size distribution roughly follows the 80:20 rule,[18] in which case the number of businesses that accounted for 80 percent of output varied from 72,000 to 814,000 between 1926 and 1999. If we assume that the antitrust cases were filed only against the firms in the top quintile, the number of private antitrust suits per business ranged from a high of about 1 in 293 firms in 1977 to a low of about 1 in 1,770 in 1997. It is important to keep in mind while considering these statistics that these antitrust cases only pertained to certain business practices that the companies sued engaged in. The likelihood that any particular business practice used by a firm with significant market power is challenged is almost certainly quite small.[19]

---

13. Detailed breakdowns by type of violation have been compiled for the 1955 to 1997 period. About 38 percent of civil cases filed by the DOJ during this period were for horizontal per se claims (e.g., price fixing, bid rigging, and market/territory/customer allocation schemes) and about 42 percent were for merger violation claims. About 8 percent of cases were monopolization claims and about 12 percent were exclusionary practices claims (e.g., predatory pricing, price discrimination, tying, and exclusive dealing). *Id.* at 95.

14. Between 1955 and 1997, for which more detailed data have been compiled, cases against Fortune 500 firms accounted for 454 of 1,348 cases based on 1 tabulation, and 631 of 2,689 cases by a second tabulation. *Id.* at 78. These tabulations include both civil and criminal cases. Data on cases against Fortune 500 firms are not available separately for civil versus criminal filings.

15. Richard A. Posner, *A Statistical Study of Antitrust Enforcement*, 13 J. L. & Econ. 365, 369 (1970). To our knowledge, additional details on the types of cases and defendants, or for other time periods, have not been compiled for FTC cases.

16. Richard Posner, Antitrust Law 46 (2nd ed. 2001). Many of these private antitrust cases were against the same defendant over the same issue and many of these involved price fixing. *See id.* at 47. Note that Posner also reports data for the 1960-1964 period; the minimum and maximum number of private antitrust cases in this period are 228 and 2005, respectively. 1739 out of the 2005 cases that occurred in 1962 were against electrical-equipment conspirators.

17. U.S. Census Bureau, Statistical Abstract of the United States (various years).

18. Robert Axtell, *Firm Sizes: Facts, Formulae, Fables and Fantasies* (Center on Social and Economic Dynamics Working Paper No. 44, 2006), arguing that size distribution of U.S. firms approximate a Pareto distribution.

19. We also recognize that the antitrust laws can have a significant effect in deterring business practices because of the fear of antitrust liability, which would not be captured in the number of cases filed. Such effects are inherently difficult to quantify. We believe the point remains that firms have a lot of latitude in choosing business practices that do not trigger antitrust scrutiny.

To presage the theme of the next section, the antitrust laws did not preclude the existence of large corporations that dominated their industries after 1890 although they certainly reined in some of the excesses of the latter 19th century. We have seen no statistics but in our experience many groupings of products that would ordinarily be defined as a relevant antitrust market have at least one firm with significant market power.[20]

The antitrust laws provide for a sort of referee process for the game of competition.[21] The focus is on tactics rather than outcomes. The federal enforcement agencies and private litigants can challenge the tactics taken by a business, such as exclusive dealing, and try to prove to the courts that those tactics should not be allowed. The courts can impose fines and penalties for businesses whose actions have gone out of bounds. While businesses whose actions have been condemned may see a heavy hand, as Standard Oil and AT&T did at opposite ends of the 20th century, the antitrust laws have made relatively modest intrusion into laissez-faire competition.

That is what antitrust is. It is worth emphasizing what it is not.

Antitrust law is not similar to public utility regulation designed to prevent certain companies that are deemed to have monopolies from charging excessive prices or earning too much profit. In fact, none of the U.S. antitrust statutes provides for any direct regulation of the prices charged by, or profits earned, by monopolies. U.S. courts are highly averse to using the antitrust laws to regulate prices even as a remedy for violating the antitrust laws.[22]

Antitrust law only concerns certain business actions that fall within its ambit.[23] It is only for these actions that courts will inquire into their effect on consumer

20. There would appear to have been periods where aggregate concentration in the U.S. economy, and/or of the relative importance of the largest firms in the U.S. economy, have increased, but, based on available data, the pattern is not systematic. Measuring concentration at an aggregate level is difficult. The available data are typically reported for markets that do not conform to antitrust markets. In addition, there are a number of other data and measurement shortcomings, such as the growing importance of exports for U.S. firms. For more details on this, see Lawrence J. White, *What's Been Happening To Aggregate Concentration in the United States? (And Should We Care?)*, NYU Economics Working Papers, Working Paper No. 02-03 (2001); F.M. Scherer & David Ross, Industrial Market Structure and Economic Performance 59-62 (3rd ed. 1990).

21. See e.g., Thurman Arnold, *Antitrust Law Enforcement, Past and Future*, 7 Law & Contemp. Probs. 5 (1940).

22. One classic statement of this aversion was Judge Wyzanski's discussion of the remedy imposed in United States v. United Shoe Machinery Corp., 110 F. Supp. 295 (D. Mass. 1953). Wyzanski expressed reluctance to regulate United's pricing because such an effort would turn United "into a public utility, and the Court into a public utility commission." *Id.* at 349. Wyzanski also noted that an injunction against United Shoe's price discrimination could not be enforced.

23. *See* Herbert Hovenkamp, Black Letter on Antitrust (1993). Hovenkamp dedicates a chapter to each type of business action that is subjected to antitrust law.

welfare. Thus, a firm with significant market power can raise prices, refuse to adhere to standards, cease production of goods and services, and engage in many other tactics that could be shown to reduce consumer welfare in the short- or long run in relevant antitrust markets or in the economy overall. It is at best shorthand, and not really correct, to say that "the purpose of antitrust is to maximize consumer welfare" except in the long-run economy-wide sense that we describe below. In practice, consumer welfare may provide the tiebreaker for those practices that the courts agree should be subject to antitrust scrutiny at all.[24]

That fact emphasizes the distinction between antitrust economics and antitrust law. Modern economic models can establish whether certain business practices can reduce consumer or social welfare in the short run under certain assumptions. These models can also be used to examine whether certain practices reduce consumer or social welfare in the factual context of a case. Modern economic models do not generally provide the courts with much help, however, for assessing whether a practice should be subject to antitrust scrutiny. Indeed, the same basic models that show that cartel price fixing reduces social welfare also show that monopoly pricing reduces social welfare. These models therefore over-identify anticompetitive practices.[25] The discipline of economics helps inform the application of antitrust analysis by the antitrust authorities and the courts. The antitrust laws themselves are based on a series of judgments made by the various branches of government, and especially the courts, concerning the role that the antitrust laws and institutions should play in regulating the market economy.

That leaves the question of why the United States has adopted this particular approach for regulating the competitive process and what series of judgments lie, at least implicitly, behind this approach.

24. It is well recognized that the courts do not seek to prohibit monopoly pricing or other exercises of monopoly power. See, e.g., Verizon v. Trinko, LLP, 540 U.S. 398 (2004). But courts do commonly attempt to assess the effect on consumer welfare of those practices that are subject to review. See, e.g., John E. Lopatka & William H. Page, *'Obvious' Consumer Harm in Antitrust Policy: The Chicago School, the Post-Chicago School and the Courts, in* Post-Chicago Developments in Antitrust Law 129, 129-132 (Antonio Cucinotta et. al., eds.2002)

25. Economists typically rely on factors outside their formal models to rationalize judicial decisions that have made some practices but not others subject to the antitrust laws. These factors include error costs, judicial costs, and effects on the incentives to innovate. *See* David S. Evans, *Economics and the Design of Competition Law, in* Issues in Competition Law and Policy 99 (D. W. Collins ed., 2008); Ronald A. Cass & Keith N. Hylton, *Preserving Competition: Economic Analysis, Legal Standards and Microsoft*, 8 Geo. Mason L. Rev. 1 (1999); Frank Easterbrook, *Predatory Strategies and Counterstrategies,* 47 U. Chi. L. Rev. 263 (1981).

III. The Objective and Premise of U.S. Antitrust Law

Previous works on the objectives of the antitrust laws have taken one of two approaches. A number of authors have tried to ascertain the "objective function" of antitrust from the legislative history of the Sherman Act.[26] Robert Bork, in perhaps the most influential work of this genre, argued that Congress intended the Sherman Act to maximize consumer welfare.[27] Some scholars have also relied on the legislative history to argue that Congress had other objectives in mind such as the protection of small businesses.[28] Other authors have concentrated on examining what the objectives of the antitrust laws should be. Older debates have surrounded whether the antitrust laws should focus entirely on consumer welfare rather than redistribution of wealth and other possible objectives. More recent discussions have focused on whether antitrust should seek to maximize consumer or total welfare.[29]

A. A REVEALED PREFERENCE APPROACH TO THE OBJECTIVES OF THE ANTITRUST LAW

We take a different approach based on what economists call revealed preference.[30] Suppose, for the same price and length of time, a consumer can go to an

26. An objective function refers to what decision makers are seeking to maximize. Economists assume that consumers maximize a utility function which is based on their preferences for different goods and services subject to their budget constraints. Economists assume that businesses are maximizing a profit function. Economists ordinarily assume that a benevolent social planner would maximize social welfare.

27. Bork writes "… the policy the courts were intended to apply is the maximization of wealth or consumer want satisfaction. This requires courts to distinguish between agreements or activities that increase wealth through efficiency and those that decrease it through restriction of output." *See* Robert H. Bork, *Legislative Intent and the Policy of the Sherman Act*, 9 J. L. & Econ. 7 (1966). We agree with Bork that the legislative intent of the Sherman Act was broadly to advance consumer welfare. Bork's analysis, including his quotes from Senator Sherman, illustrates some of the confusion in the subsequent literature. It mixes statements and concepts that correspond to classic static welfare maximization with those that correspond to dynamic total welfare maximization. For example, in a typical passage Bork notes that "[C]ongress was very concerned that the law should not interfere with business efficiency. This concern, which was repeatedly stressed, was so strong that it led Congress to agree that monopoly itself was lawful if it was gained and maintained only by superior efficiency." *Id.* at 12.

28. Lande argues that wealth transfer was the original objective. *See* Robert H. Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust: The Economic Efficiency Interpretation Challenged,* 34 Hastings L. J. 65 (1982). Hovenkamp argues that the protection of small businesses was a key objective. *See* Herbert Hovenkamp, *Antitrust's Protected Classes*, 88 Mich. L. Rev. 1 (1989).

29. *See e.g.,* Joe Farrell & Michael Katz, *The Economics of Welfare Standards in Antitrust*, 2 Competition Pol'y Int'l 3 (2006); Ken Heyer, *Welfare Standards and Merger Analysis: Why Not the Best?* 2 Competition Pol'y Int'l 29 (2006).

30. For a text that covers revealed preference, see Andreu mas-collel et. al., Microeconomic Theory 14 (1995).

opera or have dinner followed by a movie. The consumer chooses dinner and the movie. The consumer has revealed something about the underlying utility function she is seeking to maximize, subject to her budget constraint: the combination of dinner and the movie dominates opera. In the case of antitrust law we examine what choices the courts and other branches of government have made. From those choices we infer something about the objective function that those policymakers are maximizing.

The following broad choices have emerged from the U.S. antitrust laws:

- It is lawful for a firm to have significant market power.[31]
- It is lawful for a firm to engage in a multitude of practices that help it acquire significant market power.
- It is unlawful for a firm to engage in certain practices that help it acquire or maintain significant market power.[32]
- It is lawful for a firm to engage in a multitude of practices that enable it to maintain significant market power including holding on to a monopoly.
- It is unlawful for a firm to collude with other firms over setting prices and other market parameters.
- It is normally unlawful to acquire significant market power through a merger, acquisition, or joint venture.

These choices reveal several aspects of the underlying purpose of the antitrust laws.

First, the proscribed and permitted activities are not consistent with the view that the antitrust laws are seeking to maximize static consumer or social welfare in a relevant antitrust market. We know from the basic monopoly welfare loss triangle shown in Figure 1 that greater market power results in consumers paying higher prices, obtaining less output, and receiving less consumer surplus than they would with lesser market power. Greater market power also results in lower social surplus since the exercise of market power results in units of output not being produced for which the value of the output to consumers is greater than the cost to society of producing that output. Yet the antitrust laws provide businesses with wide latitude for acquiring and exercising significant market power.

31. Significant market power includes the extreme case of having a monopoly.

32. Over time the courts have changed their views on whether certain practices should be treated under a per se rule or the rule of reason and have made some practices that could have been the basis for antitrust liability either per se lawful or presumptively lawful.

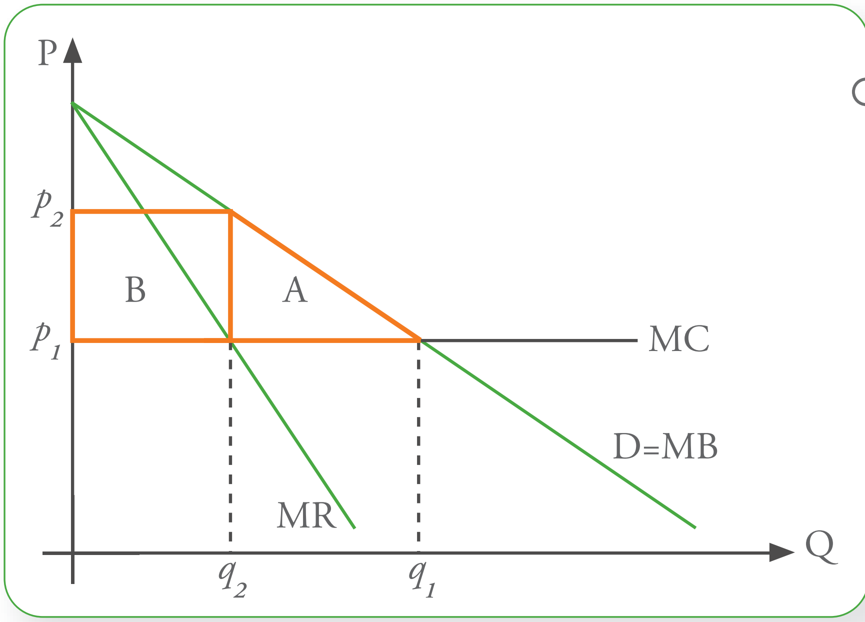


Figure 1

Monopoly pricing and consumer welfare

Second, the proscribed and permitted activities are not consistent with the view that the antitrust laws are seeking to maximize dynamic consumer or social welfare in a relevant antitrust market—in the sense of fostering a process of Schumpeterian creative destruction in that market.[33] Firms can exercise significant market power over long periods of time. They can do so even if they obtained that market power through luck or government-backed barriers to entry. The antitrust laws provide no facility for restraining dominant firms from charging high prices and earning significant profits. Firms with significant market power can also engage in a variety of actions that help them maintain that power such as advertising, various loyalty schemes, and obtaining patents.[34] They can, in practice, erect numerous barriers or benefit from ones that occur naturally, such as network effects, which deter entry.

Third, the proscribed and permitted activities are not consistent with other objectives that have been ascribed to the antitrust laws. They provide only limited relief to small businesses. Through many lawful means, larger firms can increase their market shares and in the course of doing so put smaller firms out

33. For an in-depth discussion of Schumpeter and Creative Destruction, *see* Thomas K. McCraw, *Joseph Schumpeter on Competition*, 4(2) COMPETITION POL'Y INT'L (2008).

34. In California Computer Products v. International Business Machines, 613 F.2d 727 (9th Cir. 1979), the court ruled that limiting monopolist right to engage in R&D would harm technological progress. In SCM Corp. v. Xerox Corp., 507 F.2d 358 (2d Cir. 1974), the court ruled that accumulating patents, no matter how many, is not itself illegal. *See also*, California Dental Association v. FTC 526 U.S. 756 (1999) holding that prohibitions to advertise were not a form of cartelization; Continental T.V., Inc. v. GTE Sylvania, Inc., 433 U.S. 36 (1977) discussing exclusive territories; State Oil Co. v. Kahn, 522 U.S. 3 (1997), discussing resale price maintenance.

of business. Small businesses can seek protection only if these larger firms engage in a relatively limited number of practices that have been deemed anticompetitive. More generally, the antitrust laws do not pursue a populist objective function. They do not allow the redistribution of income from firms with significant market power to other parts of society. Nor do they provide a forceful tool for preventing the significant agglomeration of significant economic—and, perhaps, political—power.

Before we describe the objective function that we argue is behind the antitrust laws, it is helpful to take a brief detour into the political debate that led to the passage of the Sherman Act and influenced its early evolution.

B. MONOPOLY POWER AND THE EARLY HISTORY OF ANTITRUST

There is no dispute that the Sherman Act was enacted in response to public concerns over the rapid rise of very large firms and certain practices that those firms engaged in with respect to their rivals and to other businesses.[35] There were diverse views, however, on the extent to which the consolidation of American industry was a problem and how the country should deal with it.[36]

The Democratic party of the time took the position that there is no good monopoly. Williams Jennings Bryan, the Democratic nominee for the Presidency in 1890, said,

> "I do not divide monopolies in private hands into good monopolies and bad monopolies. There is no good monopoly in private hands. There can be no good monopoly in private hands until the Almighty sends us angels to preside over the monopoly.[37]
>
> The Democratic platform of 1900 asserted:
>
> Private monopolies are indefensible and intolerable. . . . They are the most efficient means yet devised for appropriating the fruits of industry to the benefit of the few at the expense of the many, and unless their insatiate greed is checked, all wealth will be aggregated in a few hands and the Republic destroyed."[38]

35. *See, e.g.*, PHILIP AREEDA ET. AL., ANTITRUST ANALYSIS: PROBLEMS, TEXT, CASES 34-35 (2004).

36. Historian Richard Hofstadter provides a helpful summary which we draw upon. *See* Richard Hofstadter, *What Happened to the Antitrust Movement, in* THE POLITICAL ECONOMY OF THE SHERMAN ACT: THE FIRST ONE HUNDRED YEARS 20 (Thomas E. Sullivan ed. 1991).

37. Williams Jennings Bryan, *The Man before the Dollar: Society Not Enthralled to the Institution Solely Because the Institution Exists: The Remedy of Congressional License* in CHICAGO CONFERENCE ON TRUSTS: SPEECHES, DEBATES, RESOLUTIONS, LIST OF THE DELEGATES, COMMITTEES, ECT., 497 (1900).

38. TRIBUNE ASSOCIATION, THE TRIBUNE ALMANAC AND POLITICAL REGISTER 1901, 65 (1901).

Advocates for the powerful trusts took the opposite view, though they were comparatively reticent to speak in the face of hostile public opinion.[39] The near absence in Congress of strong vocal opposition to the Sherman Act may have reflected a perception on the part of opponents that the statute would be innocuous while dampening demands for more radical efforts to regulate the trusts.[40] The strongest statement against the principle of the Sherman Act was offered by Senator Platt of Connecticut:[41]

> "Unrestricted competition is brutal warfare, and injurious to the whole country . . . The true theory of this matter is that prices, no matter who is the producer or what the article, should be such as will render a fair return to all persons engaged in its production, a fair profit on capital, on labor, and on everything else that enters into its production . . . I believe that every man in business . . . has a right, a legal and moral right, to obtain a fair profit upon his business and his work; and if he is driven by fierce competition to a spot where his business is unremunerative, I believe it is his right to combine for the purpose of raising prices until they shall be fair and remunerative."[42]

Both extreme views were rejected when it came to adopting an antitrust policy. Instead, Congress passed legislation that put more teeth into the common law treatment of monopoly. The common law had historically refused to enforce contracts that were unreasonable restraints of trade (the classic case is *Davenant v. Hurdis* in which the tailor guild required that half of all cloth finishing for its members must be done by its members) and prohibited monopolies that had been acquired in certain ways (the classic case being the Queen's grant of a monopoly in playing cards in *Darcy v. Allen*).[43] Instead of just dissolving illegal-

39. WILLIAM LETWIN, LAW AND ECONOMIC POLICY IN AMERICA: THE EVOLUTION OF THE SHERMAN ANTITRUST ACT, 54-55 (1981).

40. HANS THORELLI, THE FEDERAL ANTITRUST POLICY: ORIGINATION OF AN AMERICAN TRADITION 215 (1955).

41. *Id.* at 198. The most vocal critic of Senator Sherman's proposed antitrust statute was Senator James George, *see* LETWIN, *supra* note 39, at 89. However, George did not object to the principle of the Sherman Act. George attacked the statute as unconstitutional and ineffective, though Bork's account suggests that George also believed that the trusts sometimes hurt small businesses by offering superior products, or lower prices attained through scale economies. Bork, *supra* note 27, at 17; *see also*, LETWIN, *supra* note 39, at 89-90 (describing George's critique of the Sherman's bill for its inability to distinguish desirable combinations from undesirable combinations, its unconstitutionality, and its ineffectiveness.)

42. THORELLI *supra* note 40, at 198.

43. 1 Clode, Early History of the Guild of Merchant Taylors 393-94 (1888); 11 Coke 84, 77 Eng. Rep. 1260 (K.B. 1603); *see also* DENNIS W. CARLTON & JEFFREY M. PERLOFF, MODERN INDUSTRIAL ORGANIZATION 632 (4th ed. 2004).

ly acquired monopolies and refusing to enforce restraints on trade, Congress provided for a system of criminal punishment that later evolved into a system of competition-based torts.[44]

William Letwin, in his classic work[45] on the origins of the Sherman Act, argues that this approach can be seen as recognizing that both competition and monopoly had their place in the economic system.[46]

> "The economists thought that both competition and combination[47] should play their parts in the economy. The lawyers saw that the common law permitted combination in some instances and prohibited it in others. Congressmen seized on this hidden agreement, and set out to construct a statute which by the use of common-law principles would eliminate excesses but allow 'healthy' competition and combination to flourish side by side."

Robert Bork has argued that Congress intended that the Sherman Act would outlaw practices that harmed consumer welfare.[48] He seems to have in mind static consumer welfare which falls when firms reduce output below the efficient level.[49] That strikes us as an overly simplistic interpretation and one that is not consistent with either the actual law or its subsequent implementation. Any firm that has market power restricts output below the level that an economic engineer seeking to maximize consumer welfare would set. Monopolies cause the greatest loss in consumer welfare—all else being equal. There is no economic reason why anyone seeking to maximize static consumer welfare would prohibit

44. The Sherman Act is a criminal statute. The right to bring a private action may have been implied, but did not become clear until the passage of the Clayton Act in 1914. Section 4 of the Clayton Act provides for private actions for treble damages. *See, e.g.,* Hylton, *supra* note 9, at 47-60.

45. Letwin, *supra* note 39, at 85.

46. At least some of the leading economists of the day were dubious about the whole antitrust enterprise. Richard Ely, who was the founder of the American Economic Association and the leader of a group of economists who rebelled against the laissez-faire tradition, seems to have recognized the loss of efficiencies in breaking up combinations such as the railroads and the need for direct regulation of prices. *See* Richard Ely et. al., Outlines of Economics 153 (2nd ed. 1912).

47. The word "combination" was used at the time to refer to firms that had become large through internal growth as well as through mergers.

48. *See* Bork, *supra* note 27.

49. Bork tends to equate anticompetitive practices as ones that reduce output and, although this is sometimes vague, in relevant antitrust markets.

cartels from engaging in price fixing that may lead to a monopoly price but allow monopolies to set a price that leads to a similar welfare loss.[50] One can attempt to reconcile this stark distinction by appealing to a multitude of factors, including the dynamic ones considered below. But these explanations lead inevitably to frameworks in which static consumer welfare maximization is, at best, one element. And these factors are usually brought in as *deus ex machina* to reconcile what are facially inconsistent results.

Since the passage of the Sherman Act there have been periodic attempts to revisit the extent to which the antitrust laws should deal with the "monopoly problem." The most famous, as well as the most successful, is the legislative package enacted in 1914, the Clayton and FTC Acts. The Clayton Act directed courts to apply a more rigid legal test—a type of per se rule—to tying, exclusive dealing, and price discrimination.[51] The FTC Act created the Federal Trade Commission and gave it power to prosecute "unfair methods of competition" which might be difficult to pursue under the Sherman Act because of the evidentiary requirements.[52] Both statutes sought to tighten the constraints on monopoly firms. The Clayton Act, as originally interpreted, did so by removing certain practices from the rule of reason framework established in *Standard Oil*. The FTC Act tightened constraints by creating an alternative enforcer that could pursue the anticompetitive conduct that was potentially immune because of the demanding evidentiary requirements of the Sherman Act. Both statutes have been interpreted more recently in a fashion that harmonizes them with the Sherman Act. More importantly, though, both statutes and the common law surrounding them have stayed well within the boundaries of the Sherman Act by taking a light hand with the monopoly problem.

More serious efforts to revisit the regulation of monopolies have failed to be enacted as law. For most of the first half of the Sherman Act's life, there were repeated attempts in Congress to enact federal incorporation statutes that would impose strict competition-based regulations on large corporations.[53] The federal incorporation statutes would have provided a direct route to preventing firms with monopoly power from either exploiting or enhancing that power through methods that would not violate the antitrust laws. The last federal incorporation

50. That was especially the case for the early years of the antitrust laws. The independent railroads that formed combinations were early targets. Without judging the issue, one can easily come up with reasons that these combinations increased consumer welfare, including by permitting coordination of traffic over a network or disciplining inefficient price wars resulting from railroads having high fixed sunk cost investments and low marginal costs. ELY, *supra* note 46, states that breaking up these combinations had unfortunate consequences.

51. For a general description, see HYLTON, *supra* note 9, at 47-48.

52. *Id.*

53. Daniel A. Crane, *Antitrust Antifederalism* 96 CAL. L. REV. 1 (2008)

attempt was the failed 1937 Borah-Mahoney bill that would have required corporations operating in interstate commerce to be licensed by the FTC.[54]

In response to recommendations of the White House Task Force on Antitrust Policy (Neal Report), in 1971 Congress considered a statute that would require the restructuring of oligopolistic industries, and, in 1973, another statute that would require dissolution of monopolies.[55] As recently as 1979, the National Commission for the Review of the Antitrust Laws and Procedures proposed that the Sherman Act be amended to permit the government to seek dissolution in the absence of a finding of monopolization under Section 2.[56]

Therefore, there has been a consensus between the judicial and legislative branches of government, for more than a century, that whatever evils monopoly may bring, society would be worse off regulating or preventing firms from seeking, obtaining, and exercising monopoly power.

Therefore, there has been a consensus between the judicial and legislative branches of government, for more than a century, that whatever evils monopoly may bring, society would be worse off regulating or preventing firms from seeking, obtaining, and exercising monopoly power.

There have also been periods in which the courts or antitrust enforcement agencies have taken a more hands-off approach. Posner's statistical study suggests that the DOJ was relatively quiet on antitrust matters from roughly 1910 to the late 1930s.[57] The Reagan administration introduced a shift in priorities away from monopolization cases that has continued in subsequent Republican administrations. But this variation has happened along a line that was drawn far away from regulating the outcomes of the competitive struggle among businesses, including ones that lead to monopoly.

C. WHAT THE COURTS HAVE SAID ABOUT MONOPOLY

To see how the courts have viewed firms with significant market power it is helpful to start with a decision that appears midway in the history of U.S. antitrust and is often viewed as one of the least friendly to firms sitting on enormous market shares: Judge Learned Hand's famous opinion in *U.S. v. Alcoa*.[58] The lower

54. *Id.* at.23-25.

55. *See, e.g.*, Areeda et. al., *supra* note 35, at 418.

56. Crane, *supra* note 53, at 26.

57. Posner, *supra* note 15, at 368.

58. Alcoa was decided by the 2nd Circuit Court of Appeals, on which Hand sat, because too many of the members of the Supreme Court had to recuse themselves. U.S. v. Aluminum Co. of America, 148 F.2d 416 (2nd Cir. 1945).

court had ruled against the government on the grounds that, although it had shown that Alcoa had a monopoly, it had failed to prove that Alcoa had engaged in anticompetitive conduct. Hand's opinion overturned the lower court and found that Alcoa had violated Sections 1 and 2 of the Sherman Act. He embraced the view that the purpose of the Sherman Act, and other government policy, "was to perpetuate and preserve, for its own sake and in spite of possible cost, an organization of industry in small units which can effectively compete with each other." He emphasized that Alcoa's sheer size enhanced its ability to engage in abuse.

Judge Hand also accepted, however, that there was nothing wrong with monopoly as the outcome of the competitive process. His views on this are worth quoting in full rather than just the famous last line:

> "Nevertheless, it is unquestionably true that from the very outset the courts have at least kept in reserve the possibility that the origin of a monopoly may be critical in determining its legality; and for this they had warrant in some of the congressional debates which accompanied the passage of the Act. This notion has usually been expressed by saying that size does not determine guilt; that there must be some "exclusion" of competitors; that the growth must be something else than "natural" or "normal"; that there must be a "wrongful intent," or some other specific intent; or that some "unduly" coercive means must be used. At times there has been emphasis upon the use of the active verb, "monopolize," as the judge noted in the case at bar. What engendered these compunctions is reasonably plain; persons may unwittingly find themselves in possession of a monopoly, automatically so to say: that is, without having intended either to put an end to existing competition, or to prevent competition from arising when none had existed; they may become monopolists by force of accident. Since the Act makes "monopolizing" a crime, as well as a civil wrong, it would be not only unfair, but presumably contrary to the intent of Congress, to include such instances. A market may, for example, be so limited that it is impossible to produce at all and meet the cost of production except by a plant large enough to supply the whole demand. Or there may be changes in taste or in cost which drive out all but one purveyor. A single producer may be the survivor out of a group of active competitors, merely by virtue of his superior skill, foresight and industry. In such cases a strong argument can be made that, although the result may expose the public to the evils of monopoly, the Act does not mean to condemn the resultant of those very forces which it is its prime object to foster: *finis opus coronat*. The successful competitor, having been urged to compete, must not be turned upon when he wins."

As Hand summarizes the state of antitrust jurisprudence in 1945 there is nothing unlawful about obtaining monopolies by "superior skill, foresight and industry." The monopoly is the "end that crowns the work" (*finis opus coronat*). Nor is there anything troubling if a firm gets the monopoly through "accident." This view echoes Supreme Court decisions that stretch back through the previous half century of antitrust. The most prominent pre-*Alcoa* monopolization decisions, *Standard Oil* and *U.S. Steel*, stress the distinction between lawful and unlawful methods of gaining monopoly power.[59] Indeed, the law was much more protective of monopolization efforts before *Alcoa*, because the courts required evidence of "specific intent" to monopolize.[60] Judge Hand's key change in the law of monopolization was to scrap the specific intent requirement. This was justified in his view because a monopolist, merely by setting his price at the monopoly level, causes the same harm to consumers as cartels do.

Antitrust law has moved far way from many of the anti-big business views expressed by Judge Hand in *Alcoa*. However, his analysis of why monopolies that win the competition for the market through superior skill, foresight, and industry have not violated the antitrust laws merely because of their success has become the standard treatment. All subsequent Section 2 decisions have embraced this view. Indeed, the fundamental test for monopolization, adopted after the *Alcoa* decision, requires the possession of monopoly power and "the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historical accident . . ."[61] The variation observed in the post-*Alcoa* case law is not over whether lawful monopolization exists, but precisely how to define the boundary between lawful and unlawful monopolization. *Alcoa* opened the door for courts to define a much larger set of activities as unlawful than would have been permissible under the pre-*Alcoa* law. But, for the most part, courts have been conservative in accepting *Alcoa*'s invitation. They have looked for practices that seem to raise a special risk of maintaining monopoly—such as the lock-in contracts condemned by Judge Wyzanski in *United Shoe*.[62] And more recently courts have come close, in the areas of predatory pricing (*Brooke Group*)[63] and essential facilities (*Trinko*),[64] to returning to the specific intent requirement of the pre-*Alcoa* law.[65]

59. *See, e.g.,* Areeda et. al., *supra* note 35, at 368-372; Hylton, *supra* note 9, at 186-188.

60. *See, e.g.,* Hylton, *supra* note 9, at 187-192.

61. United States v. Grinnell Corp., 384 U.S. 563, 570-1 (1966).

62. United States v. United Shoe Machinery Corp., 110 F. Supp. 295 (D. Mass. 1953), Aff'd Per Curiam, 347 U.S. 521 (1954).

63. Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993).

64. Verizon Communications v. Law Offices of Curtis v. Trinko, LLP, 540 U.S. 398 (2004).

65. *See, e.g.,* Hylton, *supra* note 9, at 202-219.

D. COMPETITION RULES

The antitrust laws are based on an objective and a premise.

The objective is economic progress broadly defined or, in the language of economics, long-run economy-wide consumer welfare. We believe the choices made by the legislatures and the courts are consistent with focusing on maximizing the performance of the economy, for the benefit of consumers, over long periods of time. We cannot conceive of their revealed preferences being consistent with any other objective function.

The premise is that the competitive process can generally be relied on to maximize long-run economy-wide consumer welfare. The pursuit of the crown of monopoly has been accepted by the courts and implicitly by the legislature as an important aspect of the competitive process. So much so, in fact, that the courts and legislature do not even want to distinguish between a monopoly that arrives through "accident" and one that arrives through superior skill.

In light of this objective and premise, the courts tend to proscribe business practices only when they become confident that these practices interfere with economic progress. That involves identifying situations in which the costs that consumers incur from the exercise of market power in relevant markets are substantial and outweigh the dynamic social benefits that the economy receives from allowing firms to receive monopoly profits as a reward for successful market competition. This tradeoff is between local costs (i.e. from those incurred in relevant antitrust markets) and global benefits (i.e. from stimulating investment and innovation in the overall economy).[66] Hardcore cartels are prohibited because the courts—and the U.S. Congress in passing the Sherman Act—have judged that the monopoly profits from cartels do not provide dynamic economy-wide social benefits that could offset the consumer welfare loss in relevant markets.[67] This global versus local tradeoff is central to our thesis and subsumes the more traditional static versus dynamic efficiency tradeoff.

66. To be precise, the tradeoffs are between the effect of prohibiting practices on consumer welfare loss in relevant antitrust markets including the deterrence effects of prohibiting those practices and the effect of those prohibitions on the incentives for making risky investments that could increase long-run consumer welfare in a variety of ways.

67. This judgment seems right to us but is not based on rigorous economic theory or empirical work. The prospect of sharing in cartel profits could induce entry and innovation in many of the same ways as the prospect of obtaining unilateral monopoly profits do. Similarly, one could argue that cartels may be necessary in a high-fixed cost oligopolistic industry subject to ruinous competition; *see* MICHAEL WHINSTON, LECTURES ON ANTITRUST ECONOMICS 16 (2006). This is an example of one of many aspects of antitrust in which modern economics rather incompletely informs the policy judgments that necessarily lay at the heart of antitrust law—a subject that we come to later in this article.

IV. Monopoly in European Community Competition Law

Our conclusion that the U.S. antitrust laws have a "revealed preference" for an objective function that maximizes long-run economy-wide social welfare applies with some qualification to EC competition law as well.

We focus our attention on Article 82 of the EC Treaty which pertains to abuses of a dominant position.[68] Under EC case law a firm has a dominant position if it "can hinder the maintenance of effective competition on the relevant market by allowing it to behave in an appreciable extent independently of its competitors and customers and ultimately of consumers."[69] As a practical matter, firms are usually found dominant if they have market shares of 50 percent or more and sometimes as low as 40 percent.[70] One can consider a dominant firm as one that has significant market power. The European Commission investigates and determines whether a firm has abused a dominant position. Its decisions can then be appealed to the European Court of First Instance and the European Court of Justice.[71]

Article 82 provides for two sorts of abuses.[72] The first are exclusionary abuses which are similar to those found in U.S. case law. The major difference is that the European Community treats most of these abuses under an essentially per se rule. A firm has committed an abuse if it is dominant and has engaged in the pro-

68. The European Community's antitrust laws are based on two articles of the Treaty of Rome that established the European Community in 1957. These articles were renumbered in subsequent treaties. Article 81 concerns concerted practices and is similar to Sherman Section 1 except insofar as Article 81(3) provides an explicit examination of efficiency rationales for horizontal agreements. The European Community's treatment of mergers and coordinated practices are similar to those in the United States, at least for the purposes of our discussion. For an introduction to EC competition law generally, *see* BELLAMY AND CHILD: EUROPEAN COMMUNITY LAW OF COMPETITION: (Peter Roth & Vivien Rose eds., 6th ed. 2008).

69. Case 322/81, Michelin v. Comm'n, 1983 E.C.R. 03461.

70. In T-219/99, British Airways v. Commission, 2003 E.C.R. II-05917, ¶¶ 211, 225, British Airways was found dominant in the context of Article 82 with a share which had declined from 46 percent to just under 40 percent during the period of abuse. The finding relied heavily, though, on the fact that the rest of the market was very fragmented. Subsequently, in Case COMP/38.233, Wanadoo Interactive, 2003, the Commission concluded in paragraph 227 that Wanadoo did hold a dominant position, albeit it only had a market share of 39 percent. The Commission reached this finding based both on the size and strength of Wanadoo's main competitors, who all had market shares between 6.5 percent and 16 percent.

71. The EC Member States have their own competition laws which are not covered in this section. The EC competition laws regulate business practices that involve multiple member states. For more detail, *see* ROTH & ROSE, *supra* note 68.

72. Neither of these two categories of abuses makes it unlawful for a firm to engage in practices that help it obtain a dominant position or, to use the Sherman Act phrase, "to monopolize."

scribed practice. Some practices that are seldom prohibited in the United States because plaintiffs bear a stiff burden under a rule-of-reason analysis remain problematic in the European Community. Moreover, the European Commission and the European courts tend to focus on whether the dominant firm has placed its competitors at an "unfair advantage."[73] From the standpoint of a dominant firm conducting business, the differences regarding exclusionary practices between the United States and the European Community are, however, matters of degree as well as both secular and cyclical trends in antitrust thinking.

The second type of abuse is "exploitative" which has no U.S. counterpart. In listing possible abuses of a dominant position Article 82 includes "directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions. . . ."[74] Thus Article 82 has a specific provision that bars firms that have a dominant position from charging "high prices." The European courts have found that it is unlawful for a dominant firm to charge a price for a product or service that is excessive relative to its economic value where value is based on the cost of the product or service or the price of comparable goods.[75]

However, the European Commission has taken its discretion, as the prosecutor, of not pursuing "excessive pricing" cases generally.[76] In 1975 the Commission said that "measures to halt the abuse of dominant positions cannot be converted into systematic monitoring or prices." In 1994 the Commission affirmed that,

> "The existence of a dominant position is not itself against the rules of competition. Consumers can suffer from a dominant company exploiting this position, the most likely way being through prices higher than would be found if the market were subject to effective competition. However, the Commission in its decision-making practice does not normally control or condemn the high level of prices as such."[77]

73. This is based on influences from the ordo-liberal school. *See* David Evans & Christian Ahlborn, The Microsoft Judgment and its Implications for Competition Policy Towards Dominant Firms in Europe (Apr. 2008) (Working Paper, *available* at http://ssrn.com/abstract=1115867); Christian Ahlborn & Carsten Grave, *Walter Eucken and Ordoliberalism: An Introduction from a Consumer Welfare Perspective*, 2 Competition Pol'y Int'l 197 (2006).

74. Treaty Establishing the European Community, Nov. 10, 1997, 1997 O.J. (C 321) Art. 82.

75. Case 27/76, United Brands v Comm'n, 1978 E.C.R. 00207.

76. Roth & Rose, *supra* note 68, at 9-074.

77. The European Commission currently has an investigation against Qualcomm in which the main issue is whether Qualcomm's royalty rates are "excessive." At the time of this writing the Commission has not issued either a statement of objections or a decision against Qualcomm.

The Commission's most recent decision on excessive pricing—in which it dismissed two complaints against the Port of Helsingborg by ferry operators—has indicated little enthusiasm for regulating the prices of dominant firms.[78]

Perhaps the clearest evidence that the Commission does not prevent dominant firms from enjoying the fruits of their market power is its approach during its lengthy Microsoft investigation. Despite finding that Microsoft has a near monopoly over computer operating systems the Commission focused on such Microsoft practices as refusal to supply and tying rather than on Microsoft's prices.[79] Moreover, the Commission has not pursued excessive pricing claims against numerous dominant firms that are undoubtedly charging prices that exceed the cost of provision.[80]

Overall the European Community has more stringent rules of the game for firms that achieve significant market power than does the United States. The European Community has per se rules where the United States has rule of reason; it finds practices unlawful under its per se analysis that would not be found unlawful under a rule of reason analysis with similar facts in the United States; and it allows for the possibility of restraining high prices. Moreover, for all intents and purposes, the European Commission has had the final word on abuses under Article 82. In the last 20 years the European Court of Justice has rejected a decision by the Commission concerning an Article 82 abuse on a substantive point only once, partially, out of 15 cases.[81]

However, the European Community provides for weaker enforcement of the antitrust laws than does the United States. There has been no mechanism for private enforcement of the competition laws for EC-wide offenses. Private actions remain relatively uncommon and difficult to pursue in most of the member states. Ordinarily, plaintiffs can only recover actual damages. Some European countries have begun to embrace class actions of some form and the European Community is considering the role of private actions going forward. The leading proposals for class actions have specifically rejected awarding multiples of damages.[82] As a

78. Case COMP/A.36.570/D3, Scandlines v. Port of Helsingborg, 2004 ¶¶ 52,53.

79. The exception was, in seeking to enforce its remedies, the Commission asserted that Microsoft's royalty rates for certain licenses were excessive, but even here the main concern was that the royalties would continue to exclude market rivals for server operating systems and that the proposed royalties came from unlawfully acquired dominance.

80. The main exception to this statement is that the Commission has pursued excessive pricing cases against some of the formerly state-owned monopolies but it has done so in part as the de facto regulator for these sectors.

81. Evans & Ahlborn, *supra* note 73, at 25.

82. *Commission White Paper on Damages Actions for Breach of the EC antitrust rules* (April 2008), *available at* http://ec.europa.eu/comm/competition/antitrust/actionsdamages/files_white_paper/whitepaper_en.pdf; UK Office of Fair Trading, *Private actions in competition law: effective redress for consumers and business,* Discussion paper, *available* at http://www.oft.gov.uk/shared_oft/reports/comp_policy/oft916.pdf.

result, the European Community has higher standards of behavior for dominant firms but weaker enforcement while the United States has lower standards but stronger enforcement.

As in the United States, the reality is that most dominant firms, and even monopoly ones, can engage in many activities that help them obtain significant market power and exploit that power. The European Commission has issued 17 decisions that find an Article 82 abuse between 1998 and 2007 for an average of about two decisions per year.[83] Firms generally face few constraints in acquiring dominant positions and securing the benefits of those positions through various business practices. The hand seems heavy for those companies that are touched by the EC's competition laws which can seem inflexible and harsh on successful firms. As a practical matter, though, the European Community follows the United States in regulating the boundaries of the game of competition but giving firms wide latitude within those bounds. Companies that win the competitive struggle in the European Community can generally expect to enjoy the fruits of their efforts: *finis opus coronat*.[84]

As a result, the European Community has higher standards of behavior for dominant firms but weaker enforcement while the United States has lower standards but stronger enforcement.

V. Reconciling Antitrust and Intellectual Property Law

As with the antitrust laws, the intellectual property ("IP") laws provide firms with some guarantees that they will receive the prize of monopoly profits in return for winning at the competitive game.

In most industrialized countries, however, many creations of the human mind receive no property protection at all.[85] Basic mathematical and scientific research results go into the common pool of knowledge. Albert Einstein obtained protection for his methods of refrigeration but nothing for his work on the general theory of relativity. Arguably brilliant business insights such as creating an international chain of coffeehouses or placing advertising on search results pages receive no protection.

83. The EC member states each has a competition authority and these authorities also issue the equivalent of decisions on abuse of dominance for domestic matters. The United Kingdom's Office of Fair Trade is one of the most active authorities. It issued 90 decisions between 2001 and 2007 regarding violations of Articles 81 and 82.

84. There are some warning signs that this may not continue, which we discuss below.

85. For an excellent survey of intellectual property policy *see* Adam B. Jaffe & Josh Lerner, Innovation and Its Discontents—How Our Broken Patent System Is Endangering Innovation and Progress, and What to Do about It (2004).

When they are granted, intellectual property rights come with restrictions. Companies can keep secret whatever recipes, methods, or insights they have. While trade secrets law prevents the theft of those secrets, these laws do not prevent others from reverse engineering or independently discovering the secret. To gain protections, inventors can seek a patent in some circumstances but only in return for disclosing the invention—thereby adding to the pool of knowledge—and only for a limited period of time. For written, spoken, and visual works inventors can obtain a copyright which provides significant protection from others replicating the works but also provides for fair use.

Debates have occurred in many countries on whether intellectual property protection has gone too far or not far enough.[86] But the broad consensus in industrialized countries for the last two centuries has been that, when entrepreneurs must invest in activities that have an uncertain payoff, they need to be able to expect to receive a reward for their efforts. Patents, copyrights, trademarks, and trade secrets establish limited property rights that enable entrepreneurs to receive rewards for successful products and services. At the same time, however, there has been a broad consensus against establishing property rights over results that require little effort to produce or ones that are in some sense too important for scientific progress to limit.

Intellectual property policy in the industrialized world therefore balances the losses from restricting output in markets against the benefits from providing incentives for investment and innovation.[87] On the one hand, it recognizes the importance of ex post monopoly profits in stimulating innovative effort. That is the main motivation for granting rights at all. On the other hand, it is sensitive to the inefficiencies that would result from limiting the dissemination of knowledge and the output of products and services based on that knowledge. While there are legitimate debates over whether there is too much intellectual property protection, it is important to recognize that a vast portion of "innovative efforts" that could be given protection are not. In addition to the limitations on scope and duration observed in patent and copyright statutes, the case law in both fields adheres to a general principle against awarding property rights for abstract ideas, formulae, or processes that could be embodied in many different types of innovation or expression.[88] These restrictions place sharp limits on the static welfare costs that could result from the key intellectual property statutes.

86. *Id*; *see also* James Bessen & Michael J. Meurer, Patent Failure: How Judges, Lawyers, and Bureaucrats put Innovators at Risk (2008).

87. *See* Nancy Gallini & Susan Scotchmer, *Intellectual Property: When Is It the Best Incentive System? in* Innovation Policy and the Economy (Adam B. Jaffe et.al., eds. 2002).

88. *See, e.g.,* O'Reilly v. Morse, 15 Howard (56 U.S.) 62, 112-113 (1853) (denying patent protection to processes that could cover both known and unknown applications); Gottshalk v. Benson, 409 U.S. 63, 67 (1972) (denying patent for software based on general mathematical algorithm); Mackay Co. v. Radio Corp., 306 U.S. 86, 94 (1939) (scientific truths and their mathematical expressions not patentable).

Antitrust law and intellectual property law serve very different policy purposes.[89] The former is designed to regulate the game of competition, while the latter is designed to establish the proper bounds of property rights over products of the mind. Nevertheless, they are based on the same fundamental recognition that profits from securing significant market power serve as a reward for expending effort on things that will ultimately benefit society and that securing this effort is worth the price of deviations from the static competitive outcome.

Some observers have suggested that there is a fundamental tension between antitrust law and intellectual property law. The more simplistic analyses claim that antitrust law is about preventing monopolies while intellectual property law is about creating monopolies.[90] As we have seen, that is quite wrong. Antitrust law does not seek to deter the formation of monopolies based on physical or intellectual property or based on knowledge that is not subject to any property protection. It does not seek to regulate the prices charged or the output produced by firms that secure significant market, including monopoly, power. Nor does it seek to dismantle or erode monopolies once secured. Vast fortunes have been made, in full view of the antitrust laws, by companies that have secured their positions through accident, super skill, foresight, or industry. Intellectual property law does not create monopolies with abandon. For a limited amount of the creations of the human mind it establishes property rights that may result in the owner obtaining and maintaining significant market power.

> WE ARE NOT SUGGESTING THAT THERE IS NO TENSION BETWEEN ANTITRUST LAWS AND INTELLECTUAL PROPERTY LAWS, ONLY THAT THIS TENSION DOES NOT EMANATE FROM HAVING DIFFERENT OBJECTIVES.

We are not suggesting that there is no tension between antitrust laws and intellectual property laws, only that this tension does not emanate from having different objectives. Antitrust cases often involve intellectual property and, as with all cases, must take into account the circumstances surrounding that property. There may be ways in which companies can use intellectual property to engage in anticompetitive behavior beyond those that they can use with physical property.

There are also situations in which the courts need to consider relationships between antitrust and intellectual property laws. Requiring consumers who buy a patented product to purchase another product could increase the profits from invention. The intellectual property issue that is raised by such tying concerns

89. Other authors have reached a similar conclusion although from a somewhat different direction. *See* Ward S. Bowman, *Patent and Antitrust Law. A Legal and Economic Appraisal*, 11 J. ECON. LIT. 1403, 1403-1405; Mark Lemley, *A New Balance Between IP and Antitrust* (Stanford Law and Economics Olin Working Paper, Paper No. 340, 2007), *available* at http://ssrn.com/abstract=980045.

90. Hartford-Empire Co. v. United States, 323 U.S. 386 at 452, 65 S.Ct. 373 (1944), Justice Rutledge claims: "Basically these [patent laws and antitrust laws] are opposed in policy, the one granting rights of monopoly, the other forbidding monopolistic activities."

whether there should be limitations on the ways in which an inventor can secure profit from his invention and, ultimately, on his total return. Patent misuse deals with that question. The antitrust issue that is raised by such tying is whether that is the sort of practice the antitrust laws should consider prohibiting and, if so, should it be banned per se or subject to a rule of reason analysis. Given their foundations, an important consideration for both antitrust and intellectual property law is whether local costs outweigh global benefits

This example leads to our next important point.

VI. Consumer and Social Welfare and the Competitive Process

As we have seen, U.S. and, arguably, EC antitrust policy places great value on the dynamic competitive process in which firms can gain significant market power through superior skill, foresight, industry, and even accident but places some limits on how firms play the game. Those limits include cartels and other agreements among competitors, mergers that result in significant increases in market power, and some business practices that are deemed to go too far. That is consistent with policymakers—some combination of the legislators who enacted the laws and the courts and authorities that have interpreted them—believing that the competitive struggle among firms, with many dying and some achieving great success, counterbalanced by light regulation of the excesses, will maximize long-run economy-wide consumer welfare.[91]

A. THE BOUNDARIES OF COMPETITION LAW

These policy objectives are made operational in two related stages.

In the first stage, legislators and the courts, through the development of case law, roughly determine the boundaries of the game and a framework for assessing whether practices cross those boundaries. Sherman Act Section 1 and the Article 81 EC Treaty are reasonably specific that agreements among competitors are highly suspect although the case law has refined that considerably. The Clayton Act and Article 82 EC Treaty are specific that certain kinds of business practices such as tying are suspect. Sherman Act Section 2 and Article 82 EC Treaty provide a flexible mechanism for identifying other business practices that are suspect. Over time several categories have emerged. Some practices move from outside the boundaries to within as a result of legislative, judicial, or prosecutorial choices.

The courts have also devised general approaches for assessing whether firms have crossed the boundaries. The United States has the per se and rule of reason framework. The European Community has also developed a variety of rules-

91. In the long run there is no meaningful distinction between consumer and social welfare.

based approaches although these tend to be closer to per se condemnation for dominant firms.[92] These general approaches involve the assignment of the burden of proof at various stages of the inquiry.[93]

In this first stage the courts (in particular), in determining what sorts of competitive practices should be condemned, have focused on the long-run consequence for economic progress. It is at this stage that the U.S. and EC courts have confirmed that it is not unlawful to have a monopoly or to acquire that monopoly through a myriad of lawful ways.

In the second stage the courts assess whether particular business practices cross those boundaries and should therefore be deemed violations of the antitrust laws. That is usually a fact-intensive inquiry within the framework set out in the first stage. The analysis is usually predicated on a "relevant antitrust market" which is determined as the first step of the inquiry. Many practices never reach court for this second stage, because it has become settled law that they are within the boundaries of the game of competition. Other business practices have come to be avoided because it has become settled law that they are outside the boundaries.

The first and second stages are related. Especially in common-law countries, it is through numerous fact-intensive inquiries at the lower court level that the higher courts fashion competition rules. Nonetheless, there is an important distinction: the development of competition rules and the application of those rules. Figure 2 describes the role of stage 1 and stage 2 in regulating the competitive game.

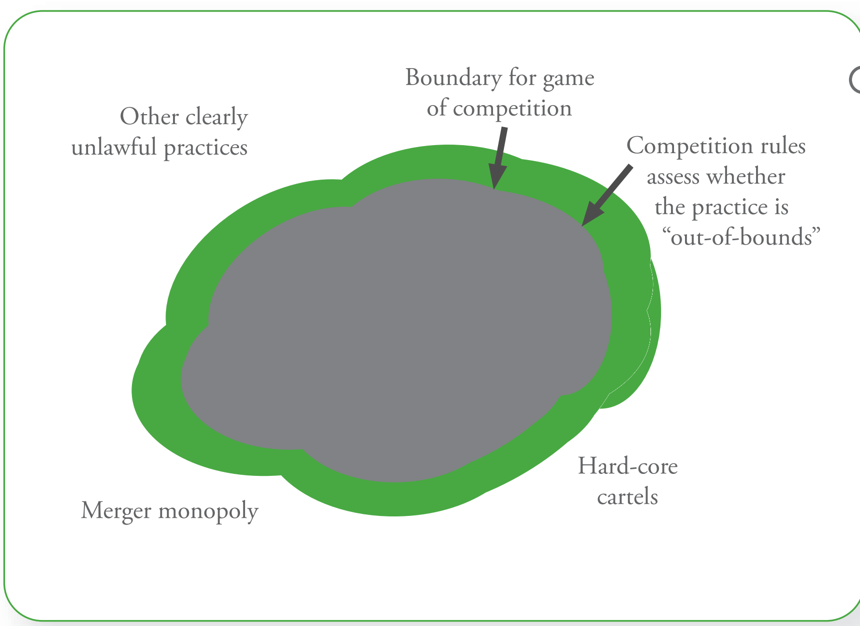


Figure 2

Antitrust rules and the game of competition

92. Evans, *supra* note 25.

93. United States v. Microsoft Corp., 2001 US App. *LEXIS 14324* (DC Cir. June 28, 2001)

B. CONSUMER WELFARE AND THE COMPETITIVE PROCESS

In recent decades the U.S. antitrust community—in particular judges, law professors, economists, and agency officials—have come to accept the premise that the antitrust laws have the singular purpose of protecting (or maximizing) consumer welfare. As a result there has been an attempt in the cases—operating at the second stage—to make this principle operational by assessing whether particular practices reduce consumer welfare. An example is the balancing of anticompetitive and pro-competitive effects that underlies the application of the rule of reason in *U.S. v. Microsoft* in a decision that has become one of the leading explications of the rule of reason analysis.[94] Some treatments of consumer welfare examine whether the practices at issue raise prices or lower output—the drivers of the basic welfare analysis described in Figure 1 and in elementary economics textbooks.[95] Other treatments of consumer welfare focus on whether a business practice "harms the competitive process."[96] Because it is assumed that the competitive process maximizes consumer welfare it is further assumed that harm to the competitive process reduces consumer welfare. Consumer welfare and the impact on the competitive process are usually considered within the context of a relevant antitrust market.

There approaches result in some confusion both in their case applications and in the literature.

First, it is not the case, for the reasons already discussed, that the overarching objective of the antitrust laws is to prohibit business practices that reduce consumer welfare in relevant antitrust markets. It is sensible and often practical to use the impact on consumer welfare in a relevant market in the second stage of the analysis as a basis for assessing whether a business practice crossed the boundaries established in the first stage. But the consumer welfare analysis used in the second stage is obviously different from the consumer welfare analysis used in the first stage since many of the practices allowed in the first stage would fail the competitive effects analysis in the second stage. The market-focused consumer welfare analysis in the second stage is a tactic for achieving the long-run economy-wide economic progress that is the focus of the first stage.

SECOND, THE "COMPETITIVE PROCESS" IS AN EMPTY PHRASE THAT CAN BE USED TO JUSTIFY OR CONDEMN ANY BUSINESS PRACTICE. THE PHRASE HAS NO OBJECTIVE MEANING IN ECONOMICS.

Second, the "competitive process" is an empty phrase that can be used to justify or condemn any business practice. The phrase has no objective meaning in

94. *Id.*

95. Hospital Corporation of America v. Federal Trade Commission 807 F.2d 1381.

96. Data Gen. Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147, 1183 (1st Cir. 1994); Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, (1993).

economics. Economists have developed numerous models of static, and occasionally dynamic, competition and have used those models to assess how perturbations in those models would affect consumer and social welfare. Economists generally recognize that there is a tradeoff between static and dynamic competition. But economists have not reached any professional consensus on the outline of a specific competitive process that would maximize consumer or social welfare, nor is it clear that there is a specific competitive process that would do so. If one could determine that a practice harmed consumer welfare then one could reasonably define that practice as harmful to the competitive process. But there is no scientific basis for inferring harm to consumer welfare from the inchoate notion of harm to the competitive process.

"Competitive process" is a circular concept within the antitrust laws. Antitrust policy assumes, as we have seen, that unfettered competition in the market and for the market is the best approach for achieving economic progress and thus long-run economy-wide social welfare. The antitrust laws recognize that certain kinds of competitive practices may interfere with economic progress and therefore seek boundaries for the competitive game. The competitive process is defined in the first stage of the analysis above as competition that lies in these boundaries and therefore does not violate the rules of the game.

We have seen the assertion of harm to the competitive process used as the core justification of two recent and much discussed Third Circuit opinions on monopolization, *Dentsply* and *LePage's*. In *Dentsply*, an exclusive dealing contract between the defendant, an artificial teeth supplier, and dealers was held to have unlawfully excluded rivals from the market for artificial teeth sales. In *LePage's*, the defendant's policy of offering bundled discounts was held to have excluded rivals from access to key distributors. In its *Dentsply* opinion, the Third Circuit perceptively noted that both cases involved a similar harm to competitive process, and treated both cases as requiring similar outcomes in court. The defendant's practices in both cases were viewed as inherently harmful to the competitive process. Missing in both analyses is an explanation why exclusive dealing contracts and bundled discounts should not be regarded simply as features of "the competitive process." Both are potential tools for seeking the undivided loyalty and promotional efforts of dealers and distributors. These points have been made in the literature, which is developing a sharper scientific basis for examining the welfare effects of exclusive dealing and bundled discounts.[97] Our point, which is in large part independent of the ultimate conclusions from the economic literature, is that the notion of harm to the competitive process, with no rigorous analysis of local or global welfare effects, fails as a theoretical rationale for decisions under the antitrust laws.

97. Benjamin Klein & Andres Lerner, Procompetitive Justifications for Exclusive Dealing: Preventing Free-Riding and Creating Undivided Dealer Loyalty (Nov. 6, 2006, *available at* http://www.ftc.gov/os/sectiontwohearings/docs/kleindealerloyalty111206.pdf); Michael Salinger, *A Graphical Analysis of Bundling* 68 J. Bus. 85 (1995).

Economists are at least in part responsible for sowing this confusion.

VII. The Role of Economics in Competition Policy

Modern economics has played a significant role in the development of antitrust law in the last fifty years. During the 1950s economists and legal scholars associated with the Chicago School demonstrated that a number of anticompetitive theories, especially those involving vertical restraints, were not founded on sound economics.[98] Other economists not associated with the Chicago School also started applying rigorous economic analysis to antitrust law.[99] These contributions have led to considerable refinement in antitrust jurisprudence starting with *Sylvania*[100] in 1977 and leading to *Leegin* in 2007. Economic analysis is regularly cited in decisions by the U.S. Supreme Court as well as lower courts and few antitrust cases proceed without dueling expert economists. Beginning in 1982, the DOJ started incorporating economic reasoning in its merger guidelines. Today, economics has become an almost *lingua franca* for the discussion of competition policy worldwide. Economics is widely and, correctly in our view, credited with making antitrust more rigorous and coherent.

There are, however, two limitations on the role that economics can play in antitrust.

One limitation is purely natural. It results from the fundamental difference between these two disciplines. Antitrust is a policy implemented through a legal process in which learning is built from examining different factual circumstances over time, in which precedents are developed which tend to promote clear and predictable rules of law, and in which making reasoned but ultimately subjective tradeoffs between local costs and global benefits is fundamental. Economics is a science that studies the behavior of consumers and businesses in a world of scarce resources that have alternative uses.[101] Industrial organization, the branch of economics that is most relevant to antitrust, studies the structure of industries and how firms interact in these industries. It largely rests on analyzing theoretical models based on certain assumptions and sometimes testing those models against data. Economic analysis is a valuable input into a judicial process that weighs the

98. For a summary, *see* Herbert Hovenkamp, *The Reckoning of Post-Chicago Antitrust*, *in* Post-Chicago Developments in Antitrust Law 1, 3 (Antonio Cucinotta et al., eds., 2002).

99. Einer R. Elhauge, *Harvard, Not Chicago: Which Antitrust School Drives Recent Supreme Court Decisions?* 3 Competition Pol'y Int'l. 59 (2007)

100. Continental T.V. v. G.T.E. Sylvania, 433 U.S. 36 (1977).

101. Lionel Robbins, An Essay on the Nature and Significance of Economic Science 16 (1945).

value of alternative sources of evidence and considers tradeoffs that go beyond what any particular economic model can handle.

The other limitation—and the one we focus on in this section—results from a mismatch between the necessary focus of antitrust and the chosen focus of the modern industrial organization literature. The dynamic competitive process and its role in promoting economic progress are at the heart of antitrust policy. The big issues in antitrust have to do with whether the global benefits from the competitive struggle, that may well lead to the creation of significant and durable market power, are outweighed by local costs that result from the restriction of output in specific markets. Industrial organization economics has paid little attention to dynamic competition and, therefore, has had little systematic knowledge to contribute to the design of antitrust rules at the first stage of antitrust discussed earlier.[102]

Industrial organization—from the early price theory work by the Chicago School to the most recent game theory work—largely considers static competition in a market.[103] Assumptions are made about certain aspects of the firms' technology, the nature of demand, how the firms interact with each other, and other factors. A model is then developed based on those assumptions and used to examine certain features of the market. Often the model is used to assess how certain business practices affect total welfare in that market. Empirical work may test some of the implications and assumptions of the model (although the ratio of empirics to theory is very low). Such models, and much of the empirical analysis, are based on looking at interactions at a point in time or possibly based on two periods. Longer-run concerns, including effects on incentives, are generally treated as "additional considerations" but are seldom actual features of the model. Moreover, matters that are important to judicial rulemaking such as error cost, ease of administration, predictability, and the indirect consequences on incentives are either ignored or mentioned in passing.

The focus on static competition in the market is not because economists have a bias against dynamic competition. Modern economics is based largely on developing mathematical models. It is hard enough to solve the equations of static models for unique solutions and draw inferences from these equations. Oftentimes the models are very sensitive to assumptions that have been made, for example, about the functional relationships between certain variables. The mathematics of dynamic models is far more challenging and the likelihood that an economist who invests efforts in such models will achieve a publishable result

102. That is not to say that economists, and economic-minded judges and lawyers, have not been influential in expounding on the problems of errors costs and the importance of long-incentives. However, systematic work on these issues is almost nonexistent in the academic literature.

103. *See* DENNIS & PERLOFF, *supra* note 43; JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION (1988). Of the 36 chapters of the Handbook of Industrial Organization, only two cover dynamic issues—one on innovation and the other on entry. HANDBOOK OF INDUSTRIAL ORGANIZATION (R. Schmalensee et al. eds, 2007).

is lower.[104] It is easy to use words to talk about dynamic competition, as Professor Joseph Schumpeter did so eloquently, but it is much more difficult to use mathematics.[105] When realism and relevance butt heads with analytical tractability, tractability almost always wins out in economics.

When realism and relevance butt heads with analytical tractability, tractability almost always wins out in economics.

A. TRACTABILITY BIAS

This "tractability bias" leads to "static competition" bias in antitrust economics. Economists focus on issues that pertain to static competition, not because they are more important than dynamic competition, but because that is what they are able to work out mathematically. This phenomenon is well known in economics and leads to one of the most popular jokes told by economists about themselves: the man who drops his keys at night and looks for them under the streetlamp because the light is better there.

To illustrate the effects of static competition bias we consider the effect of introducing dynamic considerations into several examples of possibly anticompetitive conduct. We do this to illustrate the bias and not to advocate any particular result. Moreover, we are not arguing that the development of more dynamic models would necessarily either provide any basis for changing where the boundaries for the game of competition are currently drawn or the analysis of particular cases.

1. Innovation

Consider the following illustration based in part on Williamson's welfare tradeoff model.[106] Suppose a firm monopolizes a market, as shown in Figure 3, leading to a transfer from consumers of T and a deadweight loss of D. At the same time, the conduct that led to the monopoly also created efficiencies, with the efficiency gain represented by E in the diagram. The diagram could describe the result of an exclusive dealing contract that has the effect of foreclosing market rivals (by blocking access to a key resource, supplier, or distributor) and at the same time reducing supply costs.[107]

---

104. For a flavor of the complexity of dynamic models see Richard Ericson & Ariel Pakes, *Markov-Perfect Industry Dynamics: A Framework for Empirical Work,* 62 Rev. Econ. Stud. 53, 53-82 (1995); Hugo A. Hopenhayn, *Entry, Exit, and Firm Dynamics in Long Run Equilibrium,* 60 Econometrica 1127 (1992); Boyan Jovanovic, *Selection and the Evolution of Industry*, 50 Econometrica 649, 649-670 (1982).

105. Thomas K. McCraw, Prophet of Innovation: Joseph Schumpeter and Creative Destruction (2007).

106. Oliver E. Williamson, *Economies as an Antitrust Defense: The Welfare Tradeoffs*, 58 Am. Econ. Rev. 18 (1968).

107. The conduct that both monopolizes and generates efficiencies could take many different forms, such a merger toward monopoly, as originally analyzed in Williamson's tradeoff analysis, *id*. Alternatively, the conduct could involve technological innovation, *see* Steven C. Salop, *Exclusionary Conduct, Effect on Consumers, and the Flawed Profit-Sacrifice Standard,* 73 Antitrust L. J. 311, 345-46 (2006) (example of change in product design standard with monopolizing effect).

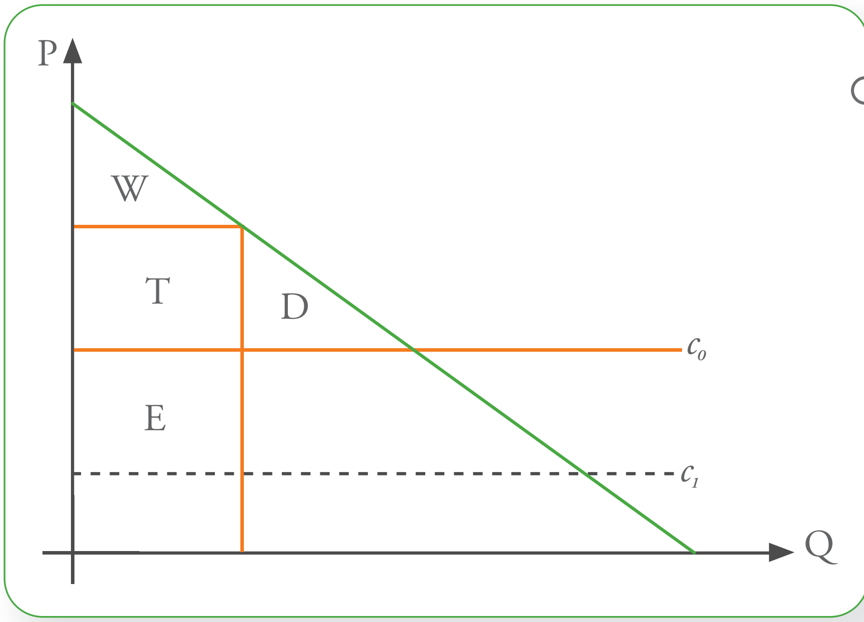


Figure 3

Monopoly, efficiency, and consumer welfare

One central argument of the Chicago School is that firms should not be penalized for efficient conduct. Doing so would discourage efficient business practices, which would reduce total welfare and could reduce consumer welfare as well. In terms of the welfare tradeoff analysis, this argument implies that the optimal penalty imposed for monopolization is the sum of the transfer and deadweight loss components $T + D$. Faced with having to pay the optimal penalty for monopolization, a firm would proceed with its monopolizing conduct whenever the efficiency gain (E) is greater than the deadweight loss imposed on society (D). Thus, if the adoption of a new product standard reduced production costs and also permitted a firm to monopolize its market, the firm would have an incentive to go ahead with the new standard if the profit expected as a result exceeded the total welfare loss imposed on consumers—or, equivalently, if the cost savings exceeded the deadweight loss.

The notion that a monopolist should be penalized an amount that reflects the static welfare costs of monopolization is accepted among analysts today. Even Chicago School critics have referred to it as one of the school's important lessons for antitrust.[108] But, as insightful as this Chicago School lesson on antitrust punishment is, it is still based on a static analysis; the welfare tradeoff model does not incorporate dynamic welfare concerns.

The simplest way to alter the welfare tradeoff model to incorporate the dynamic element is to consider the incentives that lead to the creation of monopolies. Suppose that, in the first period, the firm decides whether to invest in some

108. Areeda & Hovenkamp, *supra* note 1.

activity that could create a new market in the second period. In the second period, the firm decides whether it will adopt some practices that will enable it to monopolize the new market, depending on expected profits and the penalty, if any, for engaging in those practices.

For example, suppose in the first period the firm invests in the design and production of a new artificial tooth that will be ready to market in the second period. Rivals can copy the tooth design easily so the second period market could be highly competitive. However, the firm could reduce competitive pressure by engaging in some exclusionary act at the start of the second period. Ideally, it would like to obtain a legal barrier to entry, such as a patent or a tariff on foreign competitors, but perhaps such options are not available. The new tooth design may not be patentable or there may be too few legislators interested in providing protection from competition to the firm. Suppose the firm's best option for excluding competition, therefore, is entering into an exclusive dealing contract with a key resource supplier. The returns from the creation of the new artificial tooth depend on the firm's later success in excluding competition. It will have an incentive to monopolize if the gains from monopolization exceed the expected antitrust penalties.

THAT IS NOT TO SAY THAT THERE SHOULD NOT BE AN ANTITRUST PENALTY, ONLY THAT THE OPTIMAL PENALTY MUST CONSIDER THE DYNAMIC CONSEQUENCES.

If the firm monopolizes the market, it will impose a welfare loss on consumers equal to the monopoly transfer and deadweight loss ($T + D$), and introduce an efficiency gain (E) in the form of lower supply costs. If the firm is deterred from monopolizing the market in the second period, it will not impose any welfare losses on consumers, because the market will be competitive, and it will not generate the supply-side efficiency gain.

In this alternative "dynamic" description of monopolization, the firm's investment creates the market. The anticipation of an antitrust penalty would diminish its incentive to invest in the activity that creates the market—the new artificial tooth. More generally, the antitrust penalty has dynamic welfare consequences because it could suppress the creation of new products (as in our example) and therefore lead to the loss of the significant social wealth created from new products.[109] That is not to say that there should not be an antitrust penalty, only that the optimal penalty must consider the dynamic consequences.

Consider the private and social returns from investment for the would-be monopolist, on the assumption that it invests and later monopolizes the newly-created market. The private return to the firm would be the monopoly transfer and the efficiency gain ($T + E$). The social return from investment would be the

109. The value of new products equals roughly the area between the demand schedule and the cost schedule while the deadweight cost of monopoly is ordinarily a small fraction of the area between the demand schedule and the cost schedule.

residual consumer surplus after monopolization, the transfer, and the efficiency gain ($W + T + E$). A penalty assessed against the firm for monopolizing imposes a dynamic welfare cost because it could deny society (consumers especially) the residual surplus (W). In view of this, an optimal penalty for monopolization would include, to some degree, a bounty equal to the residual surplus to bring the private and social returns from innovation closer to each other.

An optimal antitrust penalty that includes a bounty equal to the residual surplus could easily be zero or negative. In other words, it may not be optimal to punish the monopolist at all when dynamic incentive effects are taken into account. The static punishment setting would require the optimal penalty to be set equal to the sum of the transfer and deadweight loss ($T + D$). The dynamic punishment setting would require the optimal penalty to internalize the sum of the transfer and deadweight loss *minus* the residual surplus ($T + D – W$).[110] If the residual surplus is greater than the transfer and deadweight loss amounts, the optimal penalty for monopolization may not be positive. It is this sort of reasoning that, at least implicitly, has led the legislatures and the courts to allow many business practices that can lead to monopoly. We are not advocating lower scrutiny for any particular practice. Rather, we are observing that static economic models do not take these considerations into account and therefore provide, at best, incomplete information to those who are designing competition rules.

In this example, we have assumed that the monopolist has created a new market. If, in fact, the monopolist's investment did not create or enhance a market, the standard static analysis—internalize the transfer and deadweight loss—would remain valid. So if the monopolist in this story devoted his entire investment to designing a more efficient way to transfer surplus from consumers, then there would be no case for taking a more lenient approach to punishment.

But if the monopolist creates a new market, which is the core example of the dynamic welfare benefit of innovation, the welfare gain to consumers is substantial even when the firm monopolizes the market it has created. The same can be said of investment that expands a market. In these innovation scenarios, which we think are common in real world markets and go well beyond innovation (the subject of intellectual property laws), the static welfare tradeoff analysis is no longer the best source for an optimal regulatory policy.

Admittedly dynamic models are complicated. The optimal antitrust penalty in our dynamic scenario is a messier rule than the optimal static penalty. But this does not imply that the static model should be applied as the sole source for policy recommendations in settings in which dynamic competitive effects are present.

110. The optimal penalty formula is more complicated because it depends on the probability of monopolization following investment.

2. Entry

For another illustration, consider the economic analysis of the coordinated effects of mergers. Under the coordinated effects theory, mergers can be harmful to consumer welfare because they may facilitate collusion. Modern economic analysis of the coordinated effects builds on the modern analysis of collusion. Jonathan Baker provides an especially clear and straightforward presentation.[111] Let P be the coordinated price and $?_i(P)$ represent the per unit profit of firm i evaluated at the collusive price. The firm's profit in any period at the collusive price is $\pi_i(P)q_i(P)$. If the firm cheats, setting its price just under P, it produces at its capacity k_i. The firm will avoid detection for T periods, after which the industry price falls to the zero-profit level as punishment. The firm will prefer to remain in the collusive network rather than cheat if the discounted value of the stream of profits from collusion is greater than the discounted value of the stream of profits from cheating. Thus, if the firm's discount rate is ?, it will prefer to collude rather than cheat if $[\pi_i(P)q_i(P)]/(1 - \delta) > [\pi_i(P)k_iT(1 - \delta^T)]/(1 - \delta)$.

Under the modern analysis of coordinated effects, coordination may be hampered by the existence of a firm for which the discounted value of profits from collusion is equal to the discounted value of profits from cheating. These firms have been referred to as mavericks.[112]

Suppose a firm within the collusive network chooses to acquire a maverick. Such a merger can reduce consumer welfare by eliminating the pricing constraint imposed by the maverick's existence. This analysis has led to the suggestion that if the market is conducive to coordination, the acquisition of maverick firms should establish a presumption of harm to competition.[113]

As is well known, entry constrains prices, as does the existence of maverick firms. Any policy that eliminates mavericks and permits the collusive price P to be maintained also enhances the incentive to enter and undercut the collusive price. Of course, the coordinated-effects analysis assumes that entry is not attractive at the collusive price, otherwise it would have occurred. Thus, no entry occurs at the collusive price because the expected profits from undercutting the collusive price are less than the cost of entry.

This analysis of coordinated effects suggests that entry incentives are greater than under a model that ignores the effects of mergers. Presumably the firms within the collusive network would prefer a merger over charging the competi-

111. Jonathan B. Baker, *Market Concentration in the Antitrust Analysis of Horizontal Mergers, in* ANTITRUST LAW & ECONOMICS (Keith Hylton, ed. forthcoming 2009)

112. Jonathan B. Baker & Carl Shapiro, Reinvigorating Horizontal Merger Enforcement 35 (Oct. 2007, *available* at http://ssrn.com/abstract=991588).

113. *Id.*

tive price in all future periods. But doing so would be a bad policy for them because it would undermine the threat of punishment. Each prospective entrant therefore knows that it should enter not as a "cheater" (which would not be profitable anyway) but as a maverick firm. Entering as a maverick is potentially attractive because it allows the new firm to gain the same profits as from cheating (which are insufficient to cover the entry cost) plus the option value of the merger. And given the consistent finding that acquiring firms pay a substantial premium over the market, the share of the merged entity's profits going to the entrant should be assumed to exceed the entrant's contribution to the merged entity's profits. As the merger option's value to the entrant increases, the cost of entry loses its relevance as a constraint on entry incentives in this analysis.

The prospect of a merger, in this analysis, is like a golden parachute for the entering firm.

The prospect of a merger, in this analysis, is like a golden parachute for the entering firm. A policy of acquiring troublesome mavericks, in order to maintain the collusive equilibrium, calls forth more prospective mavericks.[114] Mergers with potential coordinated effects induce entry.

We are not proposing that either of these dynamic extensions is complete or should be used to modify current competition rules. Rather, the point is that dynamic considerations are important, courts and legislatures consider them implicitly, and modern economic models often do not.

B. THE STATIC-IZATION OF ANTITRUST

Economists are playing an increasing role in antitrust. Many of the antitrust scholars writing on antitrust are economists, economic analysis is playing an increasingly important role in antitrust authorities, and it is not uncommon in countries around the world for economists to head the antitrust authority.[115] By and large economists have helped improve antitrust analysis considerably. A downside to the increased role of economists is the possible infection of antitrust with "tractability bias"— an excessive focus on static competition simply because that is what economists are most at ease in analyzing, as the parable of the keys emphasizes.

114. The policy of acquiring mavericks encourages entry. There is also the more obvious argument that the threat of entry is a function of the coordinated price. If the acquisition policy is implemented, and the pricing constraint of the maverick removed, the firms might move to a higher coordinated price. The decision to move to a higher coordinated price level could induce entry. Entry was not desirable at the initial coordinated price, which was constrained by the maverick. But after the acquisition policy is put into effect, this changes and the threat of entry may become sufficient to prevent coordination at a higher price.

115. For example, the former European Commissioner in charge of antirust was an economics professor and the current one has her undergraduate degree in economics and no law degree. The current and past heads of the U.K.'s Office of Fair Trade were economics professors. Economists are or recently served in top positions at authorities in Brazil, South Korea, and Mexico.

The excellent survey of the economic principles of antitrust by Kaplow and Shapiro illustrates the bias.[116] They examine the economic underpinnings of market power, collusion, merger, and monopolization. Every model they present is based on static competition within a relevant antitrust market. There is no formal analysis of, and but a few afterthoughts on, dynamic considerations. The local versus global tradeoff that underlies modern antitrust is largely neglected. This same statement is true for every major treatment of antitrust by economists that we know of.[117] These models therefore provide some utility for the application of competition rules adopted by the courts and some information that is relevant for the development of competition rules. But if a judge wanted to know whether any particular business practice should fall on one side or the other of the boundaries for the game of competition she would not find the answer—or even much of what she would need to know to make an informed judgment—in the modern industrial organization literature.

If the economic approach to antitrust were only of academic interest the tractability bias would be of no concern. However, the static economic approach is becoming infused in the practice of antitrust. This has become most apparent in the analysis of unilateral effects for Section 2 and Article 82. Most of the economic analysis related to determining the scope of antitrust rules concerning unilateral practices concerns competing models of largely static competition. The global benefits of unfettered competition largely get introduced through discussions of error costs.[118]

For antitrust enforcement in the United States, there is some irony here. Posner lamented years ago that lawyers dominated enforcement decisions within the antitrust enforcement agencies, allowing economists to serve largely as handmaidens.[119] The critique of enforcement as excessively lawyer-driven led to the belief that better enforcement decisions would be made if economists played a greater role in reviewing antitrust enforcement decisions.[120] Circumstances have changed and economists now play important roles in the enforcement

116. Louis Kaplow & Carl Shapiro, *Antitrust* (Harvard Law and Econ. Discussion Paper No. 575, 2007), *available* at http://ssrn.com/abstract=961264.

117. Other superb expositions of modern antitrust economics have the same bias. *See e.g.,* WHINSTON, *supra* note 67; MOTTA, *supra* note 3; HANDBOOK OF ANTITRUST ECONOMICS (Paolo Buccirossi ed.) (2008).

118. Evans and Padilla argue that firms are likely to be reluctant to implement alternative businesses practices that replicate the one found anticompetitive (such as price competition and tying), as such practices are likely to also be found anticompetitive. *See* David S. Evans & A. Jorge Padilla, *Designing Antitrust Rules for Assessing Unilateral Practices: A Neo-Chicago Approach,* 72 U. CHI. L. REV. 73 (2005); for a basic static model that considers unilateral effects see Joseph Farrell & Carl Shapiro, *Horizontal Mergers: An Equilibrium Analysis*, 80 AM. ECON. REV. 107 (1990).

119. Richard A. Posner, *A Program for the Antitrust Division,* 38 U. CHI L. REV. 500, 533 (1971).

120. *See* FRANKLIN FISHER ET. AL., FOLDED, SPINDLED, AND MUTILATED: ECONOMIC ANALYSIS AND U.S. V. IBM 348 (1983).

agencies, and some improvements have resulted. Perhaps the most important is a shift away from reliance on subjective intent evidence and toward the use of objective and empirical evidence of consumer harm.[121] However, because of the tendency to focus on static welfare models at the expense of dynamic competition, the enhanced stature of economists in the federal enforcement agencies may be not be sufficient to lead to a substantial improvement in the quality of enforcement decisions.

Outside of the United States, antitrust law is largely enforced by competition authorities with limited judicial oversight. In the European Community, for example, the Commission acts as investigator, prosecutor, and judge. Its decisions can be appealed but the higher level courts defer to its findings of facts, especially those involving complex economic assessments. The static-ization of antitrust is particularly problematic in these jurisdictions. Static economic analysis forms the basis for guidelines that provide the framework for assessing whether particular business practices violate the rules.[122]

The static focus of modern industrial organization is a problem both for itself as a branch of economic science and as a body of knowledge that is relevant to the big issues within antitrust. The academic literature needs to move from the static to the dynamic within markets and from the effects of policies within markets to the effects of policies for long-term economic progress. That will require a change in the reward systems in academic economics. The economic profession will need to provide a premium to researchers who work on dynamic competition and one that either compensates them for the especially hard mathematical work necessary for robust dynamic models or provides bonus points that skew incentives towards less mathematical dynamic analysis and away from highly technical, clever, and irrelevant static analysis.

VIII. Concluding Thoughts

The recognition of the importance of monopoly in promoting economic progress has been a key part of antitrust policy since its inception and is implicitly recognized in U.S. and EC law, which are the foundations for most global competition policy. However, there seems to have been great confusion on this point in the literature, perhaps most readily seen in the debate over the tension between IP and antitrust law and the role of antitrust and the new economy. This confusion seems to have resulted, ironically, from the increased role of modern economic analysis in the law which has imparted a bias towards static analysis. The United

121. Fisher complained that the case against IBM seemed to be based largely on evidence of subjective intent found in company memoranda. *Id*. at 347. Today, internal memoranda and emails are still used to suggest anticompetitive intent, but they are seldom the focus of a case.

122. U.S. Dept. of Justice & Fed. Trade Comm'n, Horizontal Merger Guidelines, (1992, Revised 1997).

States, European Community, and other jurisdictions around the world should avoid attempts to turn antitrust into a branch of static consumer welfare maximization. At the same time economists should spend more effort understanding how the pursuit of monopoly power affects long-run economic progress and the role of antitrust policy in this competition for the market.

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

"Dynamic Competition" Does Not Excuse Monopolization

Jonathan Baker

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

"Dynamic Competition" Does Not Excuse Monopolization

*Jonathan B. Baker**

I. Introduction

In the 2004 *Trinko* decision,[1] Justice Antonin Scalia, writing for the Supreme Court, depicted "monopoly power, and the concomitant charging of monopoly prices" as "an important element of the free-market system."[2] Scalia argued that "[t]he opportunity to charge monopoly prices—at least for a short period . . . induces risk taking that produces innovation and economic growth."[3] According to Scalia, this benefit of monopoly explains a long-standing element of the antitrust prohibition against monopolization: "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*."[4]

1. Verizon Commc'ns Inc. v. Law Office of Curtis V. Trinko, 540 U.S. 398 (2004). *Trinko* held that a regulated telephone company's alleged refusal to share its network with rivals, as required by the regulatory scheme for the telecommunications industry, did not state an antitrust claim for monopolization where the regulatory framework provided for a non-antitrust means of deterring and remedying harm to competition. In a more recent decision, the Court expanded the antitrust immunity implied by the presence of a parallel regulatory scheme. Credit Suisse Sec. (USA) LLC v. Billing, 127 S.Ct. 2383 (2007).

2. *Trinko*, 540 U.S. at 407.

3. *Id.*

4. *Id.* This observation was unnecessary to reach the decision in the case.

*Professor of Law, American University Washington College of Law. The author is indebted to Andy Gavil, Jim May, Dave Reitman, Dick Schmalensee, and Steve Salop for valuable comments and discussions, and to Michael Turner for research assistance.

In that brief passage, Justice Scalia made two controversial claims, one about economics and the other about antitrust law. He argued first that the prospect of achieving monopoly fosters innovation, and, second, that this economic proposition explains one important aspect of antitrust doctrine. The provocative new article by David S. Evans and Keith N. Hylton offers a detailed justification for Scalia's claims (though, surprisingly, without reference to Scalia's views).[5]

Neither Justice Scalia nor Professors Evans and Hylton draw out the implication of these claims for antitrust policy.[6] Indeed, it is difficult for Evans and Hylton to say more about how they would change antitrust law while simultaneously relying on the "revealed preferences" of policy-makers to infer the goals of antitrust, as that method subtly equates "is" with "ought."[7]

But it is evident that the argument will in practice be deployed to justify, on innovation-promoting grounds, the exercise of market power, and, consequently, to call for a relaxation in antitrust enforcement, particularly against monopolization.[8]

This implication was drawn by Assistant Attorney General Thomas Barnett, the current head of the Justice Department's Antitrust Division ("DOJ"). In a recent article on antitrust and innovation, Barnett endorsed Scalia's economic argument from *Trinko*, stating that "the ability to charge monopoly prices, at least for a short while, can be what induces firms to take the risks that produce inno-

5. David S. Evans & Keith N. Hylton, *The Lawful Acquisition and Exercise of Monopoly Power and Its Implications for the Objectives of Antitrust*, 4(2) Competition Pol'y Int'l (2008) [hereinafter Evans & Hylton]. Professors Evans and Hylton do not limit their antitrust law discussion to the Sherman Act §2 rules prohibiting anticompetitive single firm conduct. But the rules regarding monopolization are the focus of much of their article and are emphasized here.

6. *See* Evans & Hylton, at 236 ("We are not advocating lower scrutiny for any particular practice.")

7. The revealed preference approach is predicated either on the dubious assumption that the existing body of law—the product of the past choices of Congress, the enforcement agencies and the courts—successfully implements throughout the economic principles currently accepted by those policy-makers, or on the related and suspect claim that legal and political institutions evolve to capture efficiencies. For criticism of the efficiency view of political institutions, *see, e.g.,* Douglass C. North, Institutions, Institutional Change and Economic Performance 7 (1990) (explaining that North abandoned the efficiency view of institutions when he recognized that rulers devised property rights in their own interests and that transactions costs typically resulted in typically inefficient property rights prevailing); Daron Acemoglu, *Why Not a Political Coase Theorem? Social Conflict, Commitment and Politics*, 31 J. Comp. Econ. 620 (2003); *cf.* Richard E. Wagner, *Common Law, Statute Law and Economic Efficiency*, *in* 1 The New Palgrave Dictionary of Economics and the Law 313 (Peter Newman ed. 1998) (reviewing arguments for and against the efficiency of the common law and statutes); Jürgen G. Backhaus, *Efficient Statute Law*, *in* 2 The New Palgrave Dictionary of Economics and the Law 24 (Peter Newman ed. 1998) (same).

8. *See* Andrew I. Gavil, *Exclusionary Distribution Strategies by Dominant Firms: Striking a Better Balance*, 72 Antitrust L.J. 3, 44 (2004) (arguing that Justice Scalia's "deliberate goal" in *Trinko* was "to build the case for a more tolerant monopolization standard"). It is hard to reconcile the recent concern about the impact of monopolization standards on innovation with the lack of evidence of successful Sherman Act §2 challenges directed at innovative dominant firm conduct.

vation and other efficiencies, which ultimately benefits consumers."[9] Barnett saw that argument as a reason to call for "appropriate caution in enforcement of the antitrust laws against single firm conduct."[10] Consistent with his views, the DOJ has brought no monopolization cases during the George W. Bush administration.[11]

To the contrary: nothing is more important to economic welfare than innovation and growth, and competition and antitrust enforcement are essential for fostering them.

This comment critically evaluates Evans and Hylton's defense of Justice Scalia's legal and economic claims, and the policy implication drawn by Assistant Attorney General Barnett. It shows, first, that the legal claim is at best only partially correct, as the conduct requirement for the monopolization offense was importantly prompted by concerns other than for innovation. Second, it shows that the economic claim misleads unless qualified by the observation that the push of competition generally spurs innovation more than the pull of monopoly. Third, it explains why greater attention to fostering innovation does not call for relaxing antitrust enforcement, contrary to the policy implication.

As Evans and Hylton emphasize, innovation is important, and an appropriate concern of antitrust policy. But considerations of "dynamic competition" do not argue against antitrust enforcement. To the contrary: nothing is more important to economic welfare than innovation and growth, and competition and antitrust enforcement are essential for fostering them.

9. Thomas O. Barnett, *Maximizing Welfare Through Technological Innovation*, 15 Geo. Mason L. Rev. 1191, 1201 (2008). Barnett prefaced this aspect of his article with an explanation of how competition encourages innovation.

10. *Id*. Others favoring relaxation of antitrust's concern with market power and monopoly argue for less intervention on the ground that markets are self-correcting. *See* Frank Easterbrook, *The Limits of Antitrust*, 63 Texas L. Rev. 1, 15 (1984) (making self-correction argument). For criticism of this argument, see Ariel Ezrachi & David Gilo, *Are Excessive Prices Really Self-Correcting*?, J. Comp. L. & Econ (forthcoming 2008) *available at* http://ssrn.com/abstract=1237802) (working paper).

11. Although the Justice Department's workload statistics list one monopolization case brought in 2002, that figure appears to be an error. *See* Antitrust Division Workload Statistics, FY 1998-2007 *available at* http://www.usdoj.gov/atr/public/workstats.htm) (last consulted Oct. 13, 2008). In 2007, one West Virginia newspaper's acquisition of its rival and joint venture partner was challenged under the statutes prohibiting anticompetitive mergers and agreements (Clayton Act §7 and Sherman Act §1), along with a Sherman Act §2 count. *U.S. v. Daily Gazette Co.*, No. 2:070329 (S.D. W.Va. filed May 22, 2007). But this is only a technical monopolization case: the monopolization claim is not the gravamen of the violation; the Justice Department's press release emphasizes the acquisition frame for the case, and the Antitrust Division's workload statistics for 2007 do not record it (or any other case) as a monopolization filing. By contrast, the Justice Department brought at least seven monopolization cases during the Clinton administration. William E. Kovacic, *The Modern Evolution of U.S. Competition Policy Enforcement Norms*, 71 Antitrust L.J. 377, 449 (2003), reports seven cases and the Antitrust Division's workload statistics as of mid-2001 record eleven Sherman §2 cases filed in district court between 1994 and 2000. *See* Antitrust Division Workload Statistics, FY 1991-2000 *available at* http://web.archive.org/web/20010101000000-20011231235959/http://www.usdoj.gov/atr/public/7344.htm.

II. Why Monopolization Is Not a Status Offense

Professors Evans and Hylton correctly observe that antitrust law does not make mere monopoly pricing illegal. Monopolization is a conduct offense, not a status offense: the monopolization prohibition applies only if the monopolist has also inappropriately obtained or maintained its monopoly power. This doctrinal point was made clear during antitrust's structural era.[12] It was suggested in the seminal monopolization decision, *Alcoa*,[13] and confirmed by the Supreme Court in the still-cited elaboration of monopolization doctrine in *Grinnell*.[14]

Evans and Hylton follow Justice Scalia's *Trinko* dictum in explaining why mere monopoly power is insufficient to prove a Sherman Act §2 violation: they interpret this aspect of the longstanding doctrinal rule as proof that antitrust accepts monopoly when doing so provides incentives for innovation.[15] This interpretation of the mid-twentieth century case law is incomplete. While *Alcoa* recognized the potential for adverse incentive effects of a rule condemning monopoly pricing, it did not articulate clearly what those incentive concerns would be. It is hard to say whether the *Alcoa* court was more concerned that a sleepy monopolist would fail to minimize costs or that the monopolist would fail to pursue the development of new products and processes.[16] Moreover, the no-fault deconcentration proposals of antitrust's structural era—a mainstream idea during the 1970s (though ultimately not adopted by Congress or the courts)—suggest more of a concern with production efficiency than innovation incentives, as those proposals generally

12. Antitrust's "structural era" lasted from the 1940s through the late 1970s. *See generally* Jonathan B. Baker, *A Preface to Post-Chicago Antitrust, in* Post Chicago Developments in Antitrust Analysis 60, 63-64 (Roger van den Bergh, Roberto Pardolesi & Antonio Cucinotta, eds., 2002). Monopolization had previously been recognized as a conduct offense rather than a standard offense in *United States v. Standard Oil Co.*, 221 U.S. 1, 62 (1911) (noting "the omission of any direct prohibition against monopoly in the concrete" from the Sherman Act).

13. United States v. Aluminum Co. Of America, 148 F.2d 416, 430 (2d Cir. 1945) (*Alcoa*) (monopoly power not objectionable when acquired through "superior skill, foresight, and industry"). Ironically, *Alcoa* may have been the structural era monopolization decision that came the closest toward making monopolization a status offense, through an expansive definition of exclusionary conduct. *Cf. In re* E.I. DuPont de Nemours & Co. (*TiO_2*), 96 F.T.C. 653 (1980) (declining to find monopolization with conduct similar to the basis for a violation in *Alcoa*).

14. United States v. Grinnell Corp., 384 U.S. 563, 570-71 (1966) (distinguishing unlawful conduct from "growth or development as a consequence of a superior product, business acumen, or historic accident").

15. *See* Evans & Hylton, at 220.

16. *Alcoa*, 148 F.2d at 427 ("Many people believe that possession of unchallenged economic power deadens initiative, discourages thrift and depresses energy; that immunity from competition is a narcotic . . . that the spur of constant stress is necessary to counteract an inevitable disposition to leave well enough alone."); *see* Berkey Photo, Inc. v. Eastman Kodak Co., 603 F.2d 263, 273 (2d Cir. 1979) (recognizing that if monopolies were deemed unlawful *per se*, the antitrust laws would "compel the very sloth they were intended to prevent").

exempted large firms benefiting from substantial economies of scale without explicitly exempting firms in industries experiencing rapid innovation.[17]

Evans and Hylton neglect another reason for the acceptance of a conduct predicate for the monopolization offense during the structural era that has nothing to do with incentives to innovate: if mere monopoly pricing were deemed a violation of the antitrust laws, the possible judicial remedies—divestiture and price regulation—would be unattractive, particularly in a private case.[18] Price regulation is particularly troublesome, as courts are ill-suited for determining a reasonable price in the first instance, and, of equal importance, poorly-equipped to adjust the price over time as costs and other market conditions change.[19]

Evans and Hylton's explanation for why monopolization law historically insisted on anticompetitive conduct along with monopoly power—their claim that antitrust law values monopolies for their role in promoting innovation—is far from

---

17. *See* White House Task Force Report on Antitrust Policy, *reprinted in* 2 Antitrust L. & Econ. Rev. 11 (1968-69) (Neal Task Force Report) (proposed "Concentrated Industries Act"); Industrial Reorganization Act, S. 1167, 93rd Cong. (1973) (Hart bill), *reprinted in* Industrial Concentration: The New Learning 444 (Harvey J. Goldschmid et. al, eds., 1974); Oliver E. Williamson, *Dominant Firms and the Monopoly Problem: Market Failure Considerations*, 85 Harv. L. Rev. 1512, 1525 (1972) (persistent dominance should be presumptively unlawful under Sherman Act §2, rebuttable only by a showing of scale economies, an unexpired patent, or absolute managerial superiority). *But see* Carl Kaysen & Donald T. Turner, Antitrust Policy: An Economic and Legal Analysis 111-13 (1959) (recommending deconcentration legislation, but allowing dominant firms to rebut a presumption of unreasonableness by showing that their market power flowed from scale economies or the introduction of new products or processes). *See generally* Harlan M. Blake, *Legislative Proposals for Industrial Concentration*, in Industrial Concentration: The New Learning 340 (Harvey J. Goldschmid et. al, eds., 1974); William E. Kovacic, *Failed Expectations: The Troubled Past and Uncertain Future of the Sherman Act as a Tool for Deconcentration*, 74 Iowa L. Rev. 1105, 1137 (1989).

18. This concern was highlighted by Donald Turner, one of the most influential antitrust commentators during that period. *See* Donald F. Turner, *The Scope of Antitrust and Other Economic Regulatory Policies*, 82 Harv. L. Rev. 1207, 1223-24 (1969). Turner saw no bar to a government monopolization case "based solely on the fact that the monopoly has been retained for a substantial period of time," *id. at 1223*, but emphasized that "there is no public interest" in such a government case "unless an effective remedy is available," *id.* at 1223. He saw restructuring through divestiture or dissolution as the best remedy, *see id.* at 1213-17, and preferred public to private actions against monopolists in part because private plaintiffs, which can seek damages, *id.* at 1223, "may well be biased toward relief" that impaired the efficiency of the surviving firms," *id.* at 1224. Turner was skeptical about the utility of direct regulation of prices and entry, even when conducted by an expert administrative agency rather than a court. *Id.* at 1231. Moreover, Turner had previously rejected the idea that the Sherman Act could go farther, and simply make unlawful "the charging of a monopoly price by a monopolist" on the primary ground that Congress could not possibly have "intended the courts, under the Sherman Act, to act as price regulators," Donald F. Turner, *The Definition of Agreement Under the Sherman Act: Conscious Parallelism and Refusals to Deal*, 75 Harv. L. Rev. 655, 669-70 (1962) (stating that "the practical problems imposed on a court would of course be immense").

19. These problems may well be particularly acute in rapidly changing markets where innovation is important, as the reasonable price will turn in part on an assessment that may frustrate judicial determination: identifying the economic cost of investments in research and development, including the competitive return on such investments in the industry at issue after accounting for their risk. But that is a different point from Evans and Hylton's incentive claim.

the only serious candidate. In consequence, Evans and Hylton must argue for their view of appropriate antitrust policy on economic rather than legal grounds.

III. Monopolization Discourages Innovation

Evans and Hylton view antitrust prohibitions as chosen by courts to balance the harms from the exercise of market power against the benefits to innovation that they expect market power to confer.[20] In their view, the exercise of market power creates both a social benefit, in the form of enhanced incentives to innovate, and a social harm, in the form of the cost to consumers resulting from the reduction of output and increase in price within the market.[21] After making that tradeoff, they say, monopolization that may seem harmful when looking only to its effects on price and output within a relevant market might turn out on balance to be beneficial.[22]

The idea that monopoly could be beneficial on innovation-promoting grounds has limited policy relevance for two reasons. First, in practice, even the most aggressive antitrust enforcement regime would not remove entirely the ability of firms, whether dominant or not, to profit from their new ideas, and thus would not completely destroy incentives to innovate. There are in general many important sources of appropriability for innovating firms—including first-mover advantages, intellectual property rights, brand reputation, and the sale of complementary products and services—and it is unlikely that enforcement against monopolization would subvert them all. Even when appropriability is weak, innovation incentives may be strong.[23] With other important sources of appropriability, moreover, the monopolist's incremental incentive to innovate arising from the challenged conduct may be small or even non-existent; one cannot simply assume it is substantial relative to the other welfare losses the same conduct creates.

Second, the economic analysis proffered by Evans and Hylton ignores the possibility—indeed, the likelihood—that the exercise of market power harms aggregate innovation incentives rather than enhancing them. In the particular case of monopolization, if a dominant firm finds a way to raise its expected reward from successful innovation, that conduct may increase the dominant firm's incentive to

20. Evans & Hylton, at 220. They contend that the antitrust laws, like the intellectual property laws, are based on a "fundamental recognition that profits from securing significant market power serve as a reward for expending effort on things that will ultimately benefit society and that securing this effort is worth the price of deviations from the static competitive outcome." Evans & Hylton, at 226.

21. Evans & Hylton, at 220.

22. *See* Evans & Hylton, at 236 (arguing that the optimal penalty for monopolization could turn out to require no penalty at all, or even a subsidy, if the monopolist creates a new product or invests to expand a market).

23. Jonathan B. Baker, *Beyond Schumpeter Vs. Arrow: How Antitrust Fosters Innovation*, 74 ANTITRUST L.J. 575, 580-81 & 581 n.14 (2007).

invest in research and development ("R&D"). But as a guide to antitrust policy, this proposition is incomplete. Whether total industry R&D and the aggregate likelihood of innovation success rise depends on the magnitude of the effect and on the extent to which the dominant firm's conduct simultaneously reduces the incentive of rival firms to invest in R&D. The available empirical evidence resolves the question in favor of competition by showing that as a general rule, greater product market competition strongly encourages innovation and productivity, its close cousin.[24] Hence, even if antitrust is concerned solely with innovation—even if antitrust enforcement is undertaken without regard for the static welfare losses that Evans and Hylton point to as antitrust's justification—antitrust law should still be concerned with monopolization and other exercises of market power.

Antitrust enforcement against monopolization most obviously benefits innovation when it targets "cheap exclusion"—exclusionary practices by a dominant firm that are inexpensive for the dominant firm to implement and have no efficiency justification.[25] When such conduct impedes rival innovation, as by limiting the rival's access to key inputs or the post-innovation market, it reduces the aggregate industry probability of innovation success. The government cases against Microsoft[26] and Rambus,[27] for example, can be understood as challenging cheap exclusion.[28]

ACCORDINGLY, ANTITRUST ENFORCEMENT ATTACKING CHEAP EXCLUSION INCREASES THE AGGREGATE PROBABILITY OF INDUSTRY INNOVATION.

Cheap exclusion benefits an innovative dominant firm by increasing the reward to that firm from its own success in developing new products or processes. But that greater reward makes no difference to the probability of successful dominant firm innovation; it is simply the by-product of conduct that impedes rival innovation with no countervailing efficiency benefit. Accordingly, antitrust enforcement attacking cheap exclusion increases the aggregate probability of industry innovation.[29]

24. *See generally id.* at 583-86 (2007) (surveying literature). Additional empirical work on this topic would be useful. *Cf.* Evans & Hylton, at 240 (encouraging academic economists working on antitrust-related issues to pay more attention to dynamic competition).

25. Susan A. Creighton, D. Bruce Hoffman, Thomas G. Krattenmaker & Ernest A. Nagata, *Cheap Exclusion*, 72 ANTITRUST L.J. 975 (2005).

26. *See* United States v. Microsoft Corp., 253 F. 3d 34 (D.C. Cir. 2004).

27. Rambus Inc., 2007-1 Trade Cas. (CCH) ¶75,585 (2006), *rev'd Rambus Inc. v. FTC*, 522 F.3d 456 (D.C. Cir. 2008).

28. *See* Baker, *supra* note 23, at 592-93.

29. For a technical statement of this argument, see Jonathan B. Baker, , *"Dynamic Competition" Does Not Excuse Monopolization* 13-15 (Oct. 15, 2008), *available at* http://ssrn.com/abstract=1285223.

Suppose instead that the greater reward to the dominant firm from its successful innovation raises the incentive of the dominant firm to invest in research and development, consistent with the dynamic Justice Scalia and Professors Evans and Hylton emphasize. Antitrust enforcement can still lead to greater industry innovation, notwithstanding some reduction in the dominant firm's incentive to invest in R&D, because enforcement may simultaneously increase the R&D investment incentives of the dominant firm's rivals.[30]

Even if enforcement reduces a dominant firm's reward from innovation substantially, moreover, the marginal benefit of that firm's R&D investments need not decline markedly, so enforcement may not greatly lessen the dominant firm's likelihood of innovation success.[31] This idea may explain why antitrust enforcers have paid attention to monopolization allegations in "winner take all" (or "winner take most") markets, such as operating system software or microprocessors.[32] In those markets, the "prize" for successful innovation by the dominant firm is likely to remain large even after a monopolization case, so antitrust enforcement is likely to make little difference to the dominant firm's incentive to innovate.[33]

At the same time, the increased product market competition that results from antitrust enforcement may provide strong encouragement to R&D by the dominant firm's rivals, and consequently generate a substantial increase in rival prospects for innovation success. If so, the greater competition resulting from antitrust enforcement against monopolization would increase the aggregate odds of innovation success in the market as a whole.[34] This outcome would be contrary to what Evans and Hylton suppose, but it is consistent with the empirical evidence that competition spurs innovation.

30. Increased product market competition, as may result from antitrust enforcement, affects every firm's incentives to innovate in two ways: greater pre-innovation competition encourages innovation by feeding each firm's desire to escape product market competition, but it also discourages innovation by increasing firm fears that post-innovation competition will limit the profits from investment in R&D. The latter force is emphasized by Justice Scalia and Professors Evans and Hylton, but the desire to escape competition is often more important. *See generally* Baker, *supra* note 23. Antitrust enforcement may also encourage innovation by protecting competition in innovation markets (that is, by fostering competition in innovation itself). *Id.*

31. *Cf.* Gavil, *supra* note 8, at 43 (most innovation is encouraged by the prospect of profits rather than the prospect of monopoly profits). Similarly, the granting of intellectual property rights does not equate to the award of monopoly profits. It is now well established in antitrust, for example, that patents do not necessarily confer monopoly power.

32. *See, e.g.* United States v. Microsoft Corp., 253 F. 3d 34 (D.C. Cir. 2004) (exclusionary conduct by dominant firm in operating system software); In the Matter of Intel Corp. 128 F.T.C. 213 (1999) (exclusionary conduct by dominant firm in microprocessors).

33. *See* Baker, *supra* note 23, at 593-94.

34. For a technical statement of this argument, see Jonathan B. Baker, *"Dynamic Competition" Does Not Excuse Monopolization* 15-21 (Oct. 15, 2008), *available at* http://ssrn.com/abstract=1285223.

IV. Conclusion

Justice Scalia, supported by Professors Evans and Hylton, essentially argues that monopolization cases are brought in spite of their deleterious effects on incentives to innovate. That argument reflects an incomplete view of antitrust history, economic theory and the empirical literature. It takes one side of an old debate between Schumpeter and Arrow that today's antitrust policy can and should go beyond.[35]

GREATER ATTENTION TO "DYNAMIC COMPETITION," AS PROFESSORS EVANS AND HYLTON RECOMMEND, PROVIDES NO JUSTIFICATION FOR RELAXING ANTITRUST'S LONGSTANDING CONCERN WITH MONOPOLIZATION.

As a general matter, current antitrust rules target conduct and industries where antitrust intervention will tend to encourage innovation—as by attacking cheap exclusion, for example, or monopolization in winner-take-all markets.[36] Greater attention to "dynamic competition," as Professors Evans and Hylton recommend, provides no justification for relaxing antitrust's longstanding concern with monopolization.

35. *See generally* Baker, *supra* note 23.

36. *See generally id.* at 588-600. *Cf.* Dennis W. Carlton & Ken Heyer, *Appropriate Antitrust Policy Towards Single-Firm Conduct* (U.S. Department of Justice, Economic Analysis Group Discussion Paper EAG 08-2 March 2008) (arguing that current U.S. antitrust policy toward monopolization properly allows dominant firms to extract monopoly rents so long as those firms do not impair the competitive constraints imposed by rivals) *available at* http://www.usdoj.gov/atr/public/eag/231610.pdf.

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Competition and Innovation: Dangerous Myopia of Economists in Antitrust?

Christian Ewald

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853),
Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Competition and Innovation: Dangerous "Myopia" of Economists in Antitrust?

*Christian Ewald**

It seems fairly unlikely that the seminal papers of Professor Joseph A. Schumpeter would today have a good chance to be published in one of the leading journals specialized in industrial organization. This judgment, however, is a remarkable contrast with his still profound relevance in the world of antitrust. His warning that putting too much emphasis on static efficiency may risk killing endogenous technological change and growth has already inspired numerous policy debates in the past. A new paper by David S. Evans and Keith N. Hylton[1] ("Evans & Hylton") provides telling evidence that this is still true exactly 100 years after Schumpeter, for the first time, outlined the basis of what is known today as "Schumpeterian tradeoff".[2]

1. David Evans & Keith Hylton, *The Lawful Acquisition and Exercise of Monopoly Power and its Implications for the Objectives of Antitrust*, (4)2 Competition Pol'y Int'l (Autumn 2008) [hereinafter Evans & Hylton].

2. Schumpeter's famous concept of "creative destruction" was first presented explicitly in 1942; *see* Joseph Schumpeter, Capitalism, Socialism and Democracy, 83 (1942), Schumpeter's major steps were however already performed and anticipated in Joseph Schumpeter, Wesen und Hauptinhalt der theoretischen Nationalökonomie (1908). This never translated German-language book which might have the English title *"Essence and Limits of Equilibrium Economics"* was already published in 1908; for further details *see*: Esben Andersen, The essence of Schumpeter's evolutionary economics: A centennial appraisal of his first book, (Paper for the International Schumpeter Society Conference, Rio de Janeiro, July 2008).

*Head of Section "Economic Issues," Bundeskartellamt (Federal Cartel Office), Bonn, Germany; the views expressed in this article are those of the author alone, and do not represent the views of the Bundeskartellamt.

Despite this long period of discussion, the views on the implications which should be drawn from Schumpeter's notion that some degree of monopoly power[3] is a necessity to keep the process of innovation going are still far from unanimous. A very pronounced position can be found in the recent U.S. Department of Justice Report on the assessment of unilateral conduct under Section 2 of the Sherman Act ("Report")[4]. In particular, the Report's assessment of the risks of over- and underdeterrence shows that the positive dynamic effects of monopoly power highlighted by Schumpeter are obviously considered to be the most relevant concern of antitrust enforcers. Accordingly, the focus of the Report is much more on the negative consequences of overdeterrence and the risk of creating dynamic inefficiencies by undermining innovation.[5] The (static and dynamic) inefficiencies emanating from underdeterrence are, on the contrary, only mentioned in passing.[6]

At least from a transatlantic perspective, many other public statements currently seem to indicate that in the United States—at least in the Justice Department's Antitrust Division[7]—the Schumpeterian tradeoff provides the major intellectual underpinning for an extremely cautious "hands-off" approach in antitrust. Using the words of the current Assistant Attorney General for Antitrust:

> "Dynamic efficiency is a particular focus, and helps explain why U.S. antitrust enforcers have devoted so much time to issues surrounding innovation. Their work has a clear policy implication: antitrust enforcers must be careful not to pursue immediate, static efficiency gains at the expense of long-term, dynamic efficiency improvements, since the latter are likely to create more consumer welfare than the former. Accordingly, U.S. enforcers

3. In the following, I use the term "monopoly power" in its strict economic sense, i.e. a company's ability to raise price above marginal costs. Therefore, in particular, the term should not be equated with the legal concept of "dominance" or "significant market power".

4. U.S. Department of Justice, Competition and Monopoly: Single Firm Conduct under Section 2 of the Sherman Act, 2008 [hereinafter Report]; *available at* www.usdoj.gov/atr/public/reports/236681.htm

5. *Id*. at 14.

6. The Report explicitly mentions dynamic inefficiencies resulting from persistent monopoly power only in one very short paragraph which summarizes quite generally the impact of monopoly power on consumer welfare: *"Firms with ill-gotten monopoly power can inflict on consumers higher prices, reduced output and poorer quality goods or services. In addition, in certain circumstances, the existence of a monopoly can stymie innovation. Section 2 enforcement saves consumers from these harms by deterring or eliminating exclusionary conduct that produces or preserves monopoly"*; *Id.* at 10.

7. For quite skeptical statements by the FTC on the Justice Department's Section 2 Report, *see* P. Harbor, J. Liebowitz, & J. T. Rosch, *Statement of Commissioners Harbor, Liebowitz, and Rosch on the Issuance of the Section 2 Report by the Department of Justice* 5 (Sept. 8, 2008) [hereinafter Harbor, Liebowitz, and Rosch] *available at* http://www.ftc.gov/os/2008/09/080908section2stmt.pdf.

approach practices that bear on innovation incentives with something close to the medical principle of "first, do no harm."'"[8]

Evans & Hylton provide some very interesting but—at least from an economist's point of view—quite provocative arguments in favor of such an approach. They argue that a very cautious enforcement approach might be inevitably necessary to compensate for the increased involvement of potentially "myopic" economists in antitrust enforcement. Due to a severe deficiency of antitrust economics, economists emphasize the risk of overestimating short-term static inefficiencies at the expense of the tremendous long-term blessings stemming from all the innovations fostered by the prospect of monopoly power. Because Evans & Hylton found such apt words to describe the deficiency, it is appropriate to depict their core argument as a quote:

They argue that a very cautious enforcement approach might be inevitably necessary to compensate for the increased involvement of potentially "myopic" economists in antitrust enforcement.

"The dynamic competitive process and its role in promoting economic progress are at the heart of antitrust policy. The big issues in antitrust have to do with whether the global benefits from the competitive struggle, that may well lead to the creation of significant and durable market power, are outweighed by local costs that result from the restriction of output in specific markets. Industrial organization economics has paid little attention to dynamic competition and, therefore, has had little systematic knowledge to contribute to the design of antitrust rules [...]. Industrial organization—from the early price theory work by the Chicago School to the most recent game theory work—largely considers static competition in a market. [...] Modern economics is based largely on developing mathematical models. [...] It is easy to use words to talk about dynamic competition [...], but it is much more difficult to use mathematics. When realism and relevance butt heads with analytical tractability, tractability almost always wins out in economics. [...] This "tractability bias" leads to "static competition" bias in antitrust economics."[9]

8. Thomas Barnett, Maximizing Welfare Through Technological Innovation, (Presentation to the George Mason University Law Review 11th Annual Symposium on Antitrust, 15, October 31, 2007) *available at*: http://www.usdoj.gov/atr/public/speeches/227291.pdf

9. Evans & Hylton, *supra* note 1, 232–233.

My comments on Evans & Hylton's arguments are twofold: First, I consider it necessary to put at least two question marks behind their diagnosis that there is a severe risk of a "myopic" application of state-of-play antitrust economics. Second, in my view at least two further qualifications have to be made regarding Evans & Hylton's perception of the scope and limitations of antitrust enforcement which—explicitly or implicitly—drives their argument. Both pillars together carry my view that—to stay within the picture—prescribing antitrust enforcers strong glasses which are in the risk of leading to a severe hyperopia or even blindness seems not to be a suitable therapy for an alleged myopia in antitrust; the Schumpeterian tradeoff should not provide the justification for an overly cautious "hands-off" approach.

I. On the "Myopia" of Economists in Antitrust

To avoid any misunderstanding as regards the first pillar of my argument: I do not argue against Evans & Hylton's highly welcome appeal that more effort in academic economics should be directed toward a better understanding of the dynamic dimension of competition. It seems beyond doubt that the marginal benefits of increased research efforts are most likely to be higher than by producing further refinements of highly sophisticated models to add a small increment to an already huge bulk of literature.

My critical assessment is, rather, based on the following points: First, in my view Evans & Hylton exaggerate an indisputably existing asymmetry in theoretical economics. Second, I want to stress that Evans & Hylton's fear of a "static competition bias" only materializes if antitrust enforcement is based on a wrong idea of the role of economics and economists in antitrust enforcement.

> FIRST, IN MY VIEW EVANS & HYLTON EXAGGERATE AN INDISPUTABLY EXISTING ASYMMETRY IN THEORETICAL ECONOMICS.

On the first point: Evans & Hylton judge that there is a severe risk that antitrust enforcement systematically underestimates the merits of monopoly power for dynamic efficiency mainly from the fact that the level of mathematical formalization in dynamic theory is significantly lower than in static analysis. But—as Evans & Hylton correctly spotted—the reason for what they call "tractability bias"[10] is by no means intentional but the consequence of the complex issues concerned. But at least we have some basic models on dynamic efficiencies[11] and economics already has moved far beyond the times when the seminal papers of Schumpeter were published. Ironically, the fear of a "static

10. *Id*. at 233.

11. *See* e.g. the nice presentation in MASSIMO MOTTA, COMPETITION POLICY: THEORY AND PRACTICE 60 et. seq (2004), with some further references.

competition bias" in antitrust would be most convincing if Schumpeter hadn't entered the stage to butt the then prevailing paradigm of "perfect competition" from the throne of antitrust and the very productive (admittedly mostly non-technical) following discussion had never taken place.

Because the gap between high performance formal modeling and the focal point of what Schumpeter famously called "process of creative destruction" is still so large, it is currently only a pious hope that this gap may be closed a little bit by further research. The argument, however, that until the gap has vanished sufficiently only a cautious "hands-off" approach in antitrust can avoid a very harmful "static competition bias" deserves no support since its advocates pretend to be able to anticipate what is impossible to know: the outcome of the future academic work Evans & Hylton so forcefully ask for. Should the relevant research results finally confirm that—as Jonathan Baker puts it—"the push of competition spurs more innovation than the pull of monopoly,"[12] any caution would not only be useless but simply wrong.

To stress the core of my argument, it might be useful to refer to another Austria-born professor[13] with (at least) the same worldwide impact as Professor Schumpeter: Sir Karl R. Popper.[14] His seminal work on the theory of science[15] can be summarized as follows: the truth of our economic theories, even the best of them, cannot be verified by scientific testing but can only be falsified. He also held that theory, and human knowledge generally, is irreducibly conjectural or hypothetical, and is generated by the creative imagination needed to solve problems. Accordingly, Popper's view of scientific progress is essentially driven by the same "process of creative destruction" Schumpeter has highlighted.

If Popper's premise is accepted, there is only one reasonable approach of integrating economics and economists in antitrust: to use the relevant state-of-play of antitrust economics (to be stressed: all of it!) to make sure that the outcome of an antitrust investigation is economically sound and the best possible decision at that point in time. I am quite sure that Professor Popper would strongly support such an approach; and he would also be very reluctant to accept a policy approach in antitrust which just bets on some possible future results of an inherently open and never-ending research process.

This brings me directly to my second point: When first reading Evans & Hylton's claim that "the static-ization of antitrust is particularly problematic" in

---

12. Jonathan Baker, *Dynamic Competition' Does Not Excuse Monopolization*, 4 (2) Competition Pol'y Int'l at p. 245 (Autumn 2008).

13. Joseph A. Schumpeter was born in 1883 in Triesch (Austria-Hungary).

14. Karl R. Popper was born in Vienna in 1902.

15. Karl Popper, The Logic of Scientific Discovery (1959, original published in Vienna in 1935).

jurisdictions outside the United States where—like on the EU-level or in Germany—the competition authority acts "as investigator, prosecutor, and judge,"[16] two questions immediately jumped into my mind. First, do Evans & Hylton really assume that the recruitment policy of competition authorities as regards economists is so poor? And second, is it really true that in other jurisdictions the risk of an inappropriate definition of the role of economists in antitrust enforcement is higher than in the United States? Since no chief economist of any competition authority in the world would consider the first question to be a relevant one, I decided to think more deeply only about the second one. And the answer I arrived at is "No."

THE QUESTION WHETHER THE CHALLENGE OF INTEGRATING ECONOMIC ANALYSIS PROPERLY IN ANTITRUST ENFORCEMENT IS MANAGED SUCCESSFULLY IS NOT LINKED TO A SPECIFIC INSTITUTIONAL FRAMEWORK OF LAW ENFORCEMENT.

The question whether the challenge of integrating economic analysis properly in antitrust enforcement is managed successfully is not linked to a specific institutional framework of law enforcement. Adversarial enforcement systems like in the United States and administrative systems like in the European Community or Germany may develop different views of how a successful integration should look. The underlying principles are, nevertheless, the same.[17] My critical assessment of Evans & Hylton's hypothesis of a static bias and the "static-ization" of antitrust stems from my conjecture that they overestimate the risk that these principles are disregarded.

One of the most important lessons economists have to take to heart is the quite obvious fact that an antitrust case cannot be translated into a list of elegant formulas and equations and then solved mechanically with something like the quantitative impact on consumer or total welfare being the output. The scenario in which the Evans & Hylton's fears would really have some relevance is, however, just the one in which this lesson is totally disregarded: some number-crunching economists who are caught in the world of "highly technical, clever, and useless static analysis"[18] feed their computers with data of dubious quality and then present one single figure as the relevant evidence which should give the lead.

16. Evans & Hylton, *supra* note 1, 240.

17. For a very good description of these principles and the EU-Commission's approach of integrating economists in competition law enforcement *see*: Lars-Hendrik Röller, *Economic Analysis and Competition Policy Enforcement in Europe*, in MODELING EUROPEAN MERGERS: THEORY, COMPETITION POLICY AND CASE STUDIES 13 – 26 (Peter Bergejk & Erik Kloosterhuis ed., 2005) for some very apt considerations based on the specific institutional setup in the United States, see: HERBERT HOVENKAMP, THE ANTITRUST ENTERPRISE: PRINCIPLES AND EXECUTION (2005).

18. Evans & Hylton, *supra* note 1, 240.

Accordingly, I also think that Evans & Hylton's statement that there is some irony in the fact that "because of the tendency to focus on static welfare models at the expense of dynamic competition, the enhanced stature of economists in [. . .] enforcement agencies may not be sufficient to lead to a substantial improvement in the quality of enforcement decisions"[19] misses the point. What really should be seen by economists with some irony is the fact that it is necessary to refer to a worst-case scenario of an unsuccessful integration of economists in antitrust enforcement to underpin the fear of its systematic myopia.

Or to put it in another way: The recommendation of a cautious "hands-off" approach in antitrust should not be based on the general assumption that economists are not able to reasonably apply the state-of-play in antitrust economics. If there is some empirical evidence that they did so in the past, it seems much more appropriate to think about a more suitable integration of economic analysis in antitrust enforcement than to stop enforcement.

II. On the Scope of Antitrust Enforcement—and Its Limits

In addition to these more general thoughts on the role of economics and economists in antitrust enforcement, I believe that Evans & Hylton's argument deserves at least two further qualifiers:

1. Evans' and Hylton's paper does not properly reflect the scope of antitrust and its concept of (abuse of) monopoly power or monopolization.
2. What Evans & Hylton consider to be the objective of antitrust is too far away from being operational to form the guard rail for practical antitrust enforcement.

To develop the first point, I intend to use a method economists are very familiar with: to think in terms of an ideal world. In an ideal Schumpeterian world of competition, which in particular leaves aside the risk of failing innovation efforts, each company would, at any point in time, get exactly the reward it can reasonably expect for its innovation efforts. This is the core of the perpetual motion machine of economic progress highlighted by Schumpeter. All companies (including those not even existing today) have the same question driving their incentive to innovate: Will I get—due to some monopoly power—what I can reasonably expect as a reward for my innovation efforts?

So far, in this ideal world, no antitrust enforcer is present. But if he or she enters the stage, the relevant scope of his/her task could be described as follows:

19. *Id*. at 240.

to make sure that companies with some significant monopoly power (or let us say a "dominant market position") cannot successfully implement a strategy to get more reward for innovation than they can reasonably expect also taking into account the profits and incentives to innovate of all other companies inside or close to the relevant market concerned. The antitrust enforcers in our ideal world are therefore just focused on the monopoly profits a company with significant monopoly power intends to get at the expense of the profits and hence the incentive to innovate of other companies.

To use a quite prominent antitrust case to illustrate my point: the relevant (and much disputed) antitrust issue of the Microsoft case was not whether Microsoft has some monopoly power. The relevant question was whether Microsoft's conduct had to be assessed as an abuse of it[20]—an abuse of monopoly power in the sense of an attempt to effectively reap more profits from monopoly power than needed as a reward for its innovation efforts in the past.

Accordingly, in the context of a Schumpeterian world, the tradeoff antitrust enforcers are mostly interested in is not one of static versus dynamic efficiency; the core issue of antitrust and innovation is exclusively a dynamic one. In their analysis, Evans & Hylton lost sight of this important point which results from the quite common knowledge that static monopoly power by no means only creates static inefficiencies but also severely damages dynamic efficiency.[21] The reason for this is quite simple: the innovation effort fueled by the prospect of monopoly power significantly cools down should the prospect become reality. When monopoly power becomes reality and even goes along with significant market power ("dominance"), a company's interest in maintaining this comfortable situation as long as possible is stronger than its sense that only the pressure from other companies which follow the same "pursuit of happiness of monopoly power" has brought it into the position it currently enjoys.

Based on the (static) illustration of the famous Williamson tradeoff[22] Evans & Hylton develop the intriguing concept of a "bounty equal to the residual surplus to bring the private and social returns from innovation closer to each other."[23]

20. For this reason I consider the term "monopolization" used in U.S. antitrust law a little bit misleading. The terms "abuse of monopoly power" or "abuse of a dominant position" seem to better fit this issue.

21. *See*, for a nice and quite simple formal representation presented under the heading "Monopoly gives fewer incentives to innovate: An Example": Massimo Motta, Competition Policy: Theory and Practice, 58 (2004). One may also refer to the quite famous hypothesis of an inverted-U shaped relationship between the degree of static market power and dynamic efficiency and innovation; see Philippe Aghion, et al., *Competition and Innovation: an inverted-U relationship,* 120 Q. J. Econ. 701-728 (2005).

22. Oliver Wlliamson, *Economies as an Antitrust Defense: The Welfare Tradeoffs*, 58 Am. Econ. Rev. 18-36 (1968).

23. Evans & Hylton, *supra* note 1, 236.

For all those companies, however, which are under scrutiny of antitrust enforcers, the likelihood of a positive bounty is close to zero. Or can we really assume that a company which finally reached the paradise of significant monopoly power suddenly strives to nothing else than being driven out of it? Quite the opposite seems to be much more likely, i.e. a very strong incentive for dominant companies to barricade the doors of the paradise against all other companies wanting to come in. The objective of antitrust is to keep the door open to the paradise of monopoly power; to expect that this is deliberately done by the most powerful of all its occupants is an illusion.

THE OBJECTIVE OF ANTITRUST IS TO KEEP THE DOOR OPEN TO THE PARADISE OF MONOPOLY POWER; TO EXPECT THAT THIS IS DELIBERATELY DONE BY THE MOST POWERFUL OF ALL ITS OCCUPANTS IS AN ILLUSION.

With my last qualifier, I leave the field of the theoretical discussion laid out by Schumpeter and look into the practical limitations of antitrust enforcement. In this regard, a further expansion of the already established virtual panel of Austria-born Professors may help a bit. The new participant I would like to welcome is Professor Friedrich A. von Hayek.[24] The support I expect from Professor Hayek would probably look like this:

While reading or hearing Evans & Hylton's statement that the objective of antitrust law "is economic progress broadly defined or, in the language of economics, long-run economy-wide consumer welfare,"[25] Professor Hayek may first show a frown. Afterwards he would probably say something like: "Gentlemen, I would strongly recommend you be a little bit less ambitious and more humble." And then he would highlight some of the main elements of his work. Because his good old friend Professor Popper also recently joined the panel, he would most likely refer to his own philosophy of science which is also highly critical of what he terms scientism, i.e. pretending to know what in fact cannot be known.[26] But in any case he would make the point that all the professors of industrial organization and antitrust enforcers taken together still would know much less than what a benevolent social planner would need to know to maximize welfare in the long run.

To avoid again any misunderstanding: as an economist, I am far from disputing that the welfare standard currently provides the only suitable point of reference for sound theoretical analysis. But considering only some of the issues connected with this concept, strong doubts arise whether it is also a good practical point of reference for antitrust enforcers. How can the consumer benefits of future innovation be measured? And even if we would know how, what should

24. Hayek was born in Vienna in 1899.

25. Evans & Hylton, *supra* note 1, 220.

26. See e.g. Friedrich A. Hayek, *The Use of Knowledge in Society*, 34 AM. ECON. REV 519-530 (1945).

be the discount factor to take them properly into account? How to take on board all the indirect dynamic effects across markets? And so on and so forth.

At this point, one may see in our virtual panel probably Professor Schumpeter himself asking for the floor. And he would mention that all this reminds him very much of a rather fierce dispute he has had with a British professor—admittedly in the field of macroeconomics and not in antitrust. In the course of this discussion, Professor John M. Keynes had stated:

> "Now 'in the long run' this is probably true. [. . .] But this long run is a misleading guide to current affairs. In the long run we are all dead. Economists set themselves too easy, too useless a task if in tempestuous seasons they can only tell us that when the storm is long past the ocean is flat again."[27]

Schumpeter would surely stress that he is—like me—far away from asking for a "Keynesian Revolution" in antitrust. But if applied to the world of antitrust, Keynes' famous words may nevertheless provide the basis for a strong warning: to argue that antitrust intervention is horribly dangerous because one cannot exclude that an intervention today may probably hinder or postpone innovations in the future, is not far away from asking consumers today to pay the bill for hoped-for innovations of already very powerful companies which will probably never materialize.

III. Conclusion

The most severe issue connected with the Schumpeterian tradeoff is—that it is a tradeoff. I intended to show that adding the thoughts of some other Austria-born professors to the seminal work of Professor Schumpeter must lead to the conclusion that it's wrong to ask antitrust enforcers to be mainly concerned about monopoly power as the carrot and less concerned about competitive pressure as the stick.

The most suitable policy approach to cope with a tradeoff should be to be neither myopic nor hyperopic but to have the clearest view possible. To get this view, however, it is useful and even indispensable to have a very intensive and controversial discussion. Therefore, the current debate on the right view of antitrust on innovation can be interpreted as a "process of creative destruction of antitrust enforcement approaches." I am quite confident that Professor Schumpeter would appreciate this outcome of his work very much.

27. John M. Keynes, A Tract on Monetary Reform ch. 3 (1924).

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Injecting Innovation into The Rule of Reason: A Comment on Evans and Hylton

Richard Gilbert

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Injecting Innovation into The Rule of Reason: A Comment on Evans and Hylton

*Richard Gilbert**

The Evans and Hylton paper on *The Lawful Acquisition and Exercise of Monopoly Power and its Implications for the Objectives of Antitrust*[1] arrived in my in-box at about the same time as the U.S. Department of Justice's report on *Competition And Monopoly: Single-Firm Conduct Under Section 2 Of The Sherman Act* ("DOJ Report").[2] The two documents have much in common. Both place the historical development of the legal treatment of monopoly in an historical context and consider appropriate tests to evaluate when single-firm conduct should run afoul of the Sherman Act.

The DOJ Report generated considerable controversy. The Federal Trade Commission co-organized hearings on Section 2 enforcement with the Department of Justice, but did not endorse the final report.[3] Among other criti-

1. David Evans & Keith Kylton, *The Lawful Acquisition and Exercise of Monopoly Power and Its Implications for the Objectives of Antitrust*, 4(2) COMPETITION POL'Y INT'L (Autumn, 2008) [hereinafter Evans & Hylton].

2. U.S. DEP'T OF JUSTICE, COMPETITION AND MONOPOLY: SINGLE-FIRM CONDUCT UNDER SECTION 2 OF THE SHERMAN ACT (2008), [hereinafter DOJ Report], *available at* http://www.usdoj.gov/atr/public/reports/236681.htm.

3. *See* P. Harbor, J. Liebowitz, & J.T. Rosch, *Statement of Commissioners Harbor, Liebowitz, and Rosh on the Issuance of the Section 2 Report by the Department of Justice*, Federal Trade Commission (September 8, 2008), available at http://www.ftc.gov/opa/2008/09/section2.shtm, last accessed October 7, 2008.

*Emeritus Professor of Economics and Professor of the Graduate School, University of California at Berkeley.

cisms, Commissioners Harbour, Leibowitz, and Rosch faulted the DOJ Report for relying too heavily on economic theory in the consideration of applying antitrust law.[4] Evans and Hylton would appear to agree with this critique if economic theory is interpreted to be a static analysis of competitive effects. The authors fault economists for a "… focus on issues that pertain to static competition, not because they are more important than dynamic competition, but because that is what they are able to work out mathematically."[5] This leads to a "tractability bias" that emphasizes static competition concerns at the expense of potentially more important dynamic effects.

I AM SYMPATHETIC WITH THE CONCERN THAT DYNAMIC CONSIDERATIONS ARE OFTEN NEGLECTED IN COMPETITION ANALYSIS.

I am sympathetic with the concern that dynamic considerations are often neglected in competition analysis. Dynamic competitive effects, while complex to analyze, are too important to ignore and I have emphasized dynamic competition in my own evaluations of the state of competition policy.[6] Dynamic considerations influence competition policy in two general ways. The first is the role of dynamic competition in identifying the *types of conduct* that should raise antitrust concerns under the antitrust laws. The second is the role of dynamic competition in evaluating the *effects of conduct* that is challenged under the antitrust laws.

Evans and Hylton recognize that competition analysis is a two-stage evaluation in which the law seeks boundaries for the competitive game in the first stage (the types of conduct that raise antitrust concerns) and analyzes the effects of the conduct in the second stage. As an illustration, they note that the antitrust laws in both the United States and the European Community treat harshly the acquisition of market power through collusion by competitors, yet neither legal system challenges market power attained by a single firm through industry, foresight, or sheer luck, even though the market power that is attained can be similar in both cases. Collusion, they observe, adversely distorts the dynamic process of the competitive market, while competition to win a market and acquire market power is part and parcel of a well-functioning economy.[7]

4. *Id.*

5. Evans & Hylton, *supra* note 1, at 233.

6. See, e.g., Richard Gilbert & Steven Sunshine, *Incorporating Dynamic Efficiency Concerns in Merger Analysis: The Use of Innovation Markets*, 63 (2) Antitrust L.J. 569-602 (1995) and Richard Gilbert, *Competition and Innovation*, Issues in Competition Law and Policy (Wayne D. Collins, ed.) (2008), [hereinafter Gilbert].

7. Antitrust law distinguishes market power from monopoly power, although my comments in this paragraph relate to both. According to the U.S. Supreme Court, monopoly power is "the ability to control prices or exclude competition". Market power is the ability to price profitably above marginal cost. See, e.g., Thomas G. Krattenmaker, Robert H. Lande, & Steven C. Salop, *Monopoly Power And Market Power In Antitrust Law*, 76 Geo. L.J. 241, December, 1987.

The collection of conduct that is suspect under the antitrust laws has evolved largely from legislation and legal precedent. While economics has helped to sharpen our understanding of why certain types of conduct may or may not raise antitrust concerns, economic theory has not articulated a scientific epistemology to explain why conduct should be put in the suspect category in the first place. The primary focus of economic analysis regarding the acquisition and exercise of monopoly power has been to analyze the effects of conduct that is exposed to antitrust review. Evans and Hylton note this limited role of economists, but they too devote most of their article to the evaluation of the effects of conduct that is challenged under the antitrust law rather than evaluating the types of conduct that should raise antitrust concerns.

Evans and Hylton advocate a rule of reason approach that balances likely competitive effects against likely efficiencies from the challenged conduct. They promote a rule of reason standard that measures the effects of conduct on total economic welfare, measured by the sum of consumer benefits and producer profits. Debate over the appropriate welfare standard has long raged in antitrust circles, with some arguing that antitrust policy should focus solely on consumer welfare, while others have argued that antitrust policy should recognize total economic welfare or, at a minimum, place a positive weight on producer profits.[8] This is not the place to settle this debate, but only to note that it remains an open issue.

A central argument in the Evans and Hylton paper is that the rule of reason balancing should not be limited to a static analysis of the effects of conduct on economic welfare, but should also include a dynamic analysis of the effects of the conduct on product development and investment in productive efficiencies. They use the example of conduct associated with a monopolized new product to illustrate their argument. The monopoly price imposes a consumer cost T and a deadweight loss D from restriction of output. Under a purely static analysis, with all costs and benefits localized to the market in which the firm operates, a penalty levied on the monopolist equal to $T+D$ would provide incentives for the monopolist to choose conduct that maximizes total economic welfare. The firm would engage in the conduct only when the deadweight loss exceeds the value of any firm-specific cost savings that the conduct may achieve. If E is the profit derived from the efficiencies, the monopolist would engage in the conduct only if $T+E > T+D$, or if $E > D$. This is the correct static test under a total economic welfare standard.

8. Examples of positions in this debate are Kenneth Heyer, *Welfare Standards and Merger Analysis: Why Not the Best*, 2(2) COMPETITION POL'Y INT'L 29 (2006) (arguing for a total economic welfare standard) and Steven C. Salop, *Exclusionary conduct, effect on consumers, and the flawed profit-sacrifice standard*, 73(2) ANTITRUST L.J. 311, 336 (2006) (asserting that antitrust law focuses on consumer welfare). Joseph Farrell and Michael Katz express a more ambivalent position. Joseph Farrell & Michael Katz, *Welfare Standards in Competition Policy*, 2(2) COMPETITION POL'Y INT'L 3, 28 (2006) (economic justification for total welfare standard, but a consumer welfare standard can lead to more efficient enforcement in some instances). A further complication is that consumer welfare may include some or all of producer surplus to the extent that consumers benefit from firm-specific profits or efficiencies.

Evans and Hylton astutely point out that this is not the correct calculation if the firm would not have developed the product in the absence of the challenged conduct. The new product generates a residual consumer welfare W at the monopoly price, measured by the area between the demand curve and the monopoly price. If the conduct is pivotal to the creation of the product, the correct penalty under a total economic welfare standard is $T+D-W$. The conduct is socially desirable if and only if $E+W > D$. Under the optimal penalty, the firm would engage in the conduct if and only if $T+E > T+D-W$, or if $E+W > D$. Note that the optimal penalty can be negative (meaning that no liability is incurred) even though the conduct may incur a static welfare loss.

The utility of the Evans and Hylton rule depends on whether the conduct at issue is pivotal to the creation of the new product. If the product would have been created with or without the conduct, then society would suffer the loss of consumer surplus from the conduct with no offsetting dynamic benefits. Furthermore, the monopoly that the conduct helps to create may have other potentially adverse effects on innovation. Monopoly profits can be a disincentive for a firm to invest in new and improved products that might make its existing monopoly obsolete. And monopolizing conduct may erect artificial barriers to competition from rival firms that are potential sources of innovative products and production techniques.

THE UTILITY OF THE EVANS AND HYLTON RULE DEPENDS ON WHETHER THE CONDUCT AT ISSUE IS PIVOTAL TO THE CREATION OF THE NEW PRODUCT.

Evans and Hylton argue that static evaluations of competitive effects have dominated the economic analysis of conduct that is suspect under the antitrust laws, because that is what economists do best, despite the fact that "static economic models . . . provide, at best, incomplete information to those who are designing competition rules."[9] Certainly, conduct can have dynamic effects that swamp the consequences for static economic efficiency. But Evans and Hylton underestimate the challenge of subjecting firm conduct to a thorough rule of reason analysis, even one that is limited to static competitive effects. The DOJ Report considers a rule of reason test that inquires whether challenged conduct "reduces competition without creating a sufficient improvement in performance to fully offset these potential adverse effect[s] on prices and thereby prevent consumer harm."[10] The DOJ Report notes that "The effects-balancing (rule of reason) test confronts a court with the administrative challenge of conducting an open-ended measuring of effects that includes comparing the existing world with a hypothetical world that is subject to debate. These administrability problems include limitations on both the ability of economists accurately to measure the net consumer-welfare effects

9. Evans & Hylton, *supra* note 1, at 236.

10. DOJ Report, *supra* note 2, at 37. (footnote omitted)

of particular conduct and the ability of judges and juries to evaluate this evidence."[11] The DOJ Report concludes that "The Department does not believe that the effects-balancing test should be the general test for analyzing conduct under section 2"[12] because, in plain language, it is too hard to do.

While I am not as pessimistic as the DOJ Report about the ability of economists to balance competitive effects and efficiencies, Evans and Hylton are perhaps too confident about the practicality of such balancing, particularly when the exercise includes dynamic competitive effects. Indeed, some of the participants in the hearings that led to the DOJ Report testified that courts have never engaged in an actual quantitative balancing of competitive harms and efficiencies in a Section 2 case, even when the evaluation has been limited to static impacts.[13]

A prominent example of the application of the rule of reason to alleged monopolizing conduct is the antitrust case brought by the DOJ and several states against Microsoft.[14] Among other inquiries, the appellate court considered whether three elements of the Windows operating system and Internet Explorer browser harmed competition. The Court evaluated the product design conduct by applying the following steps:[15]

- The plaintiff must demonstrate that the conduct harmed consumers (an anticompetitive effect);
- if a plaintiff successfully demonstrates anticompetitive effect, then the monopolist may proffer a pro-competitive justification for its conduct; and
- the plaintiff can rebut the proffered pro-competitive justification or, if the justification stands unrebutted, then the plaintiff must demonstrate that the anticompetitive harm of the conduct outweighs the pro-competitive benefit.

The third step implies a rule of reason type of balancing of benefit and harm. But the *Microsoft* Court did not balance benefits and harms because it never got to the third step. For two of the three design elements, the Court concluded that Microsoft's conduct harmed competition and Microsoft had not demonstrated

11. *Id.* (footnote omitted).

12. *Id.* at 37.

13. *Id.* at 38, footnote 38.

14. U.S. v. Microsoft, U.S. Court of Appeals for the DC Circuit, 253 F.3d 34 (2001).

15. The Court described five principles. I have condensed the first two principles into one principle dealing with competitive effects.

any pro-competitive benefits. Therefore its conduct failed the test without the need for a quantitative balancing. For the third element the Court concluded that Microsoft offered a pro-competitive justification, which the plaintiff neither rebutted nor demonstrated was outweighed by the harm to competition.

I have argued elsewhere that a rule of reason standard for product innovation would be difficult to implement with an acceptable degree of accuracy.[16] New products have spillover effects that can advantage or disadvantage other firms. Conduct that enhances market power can increase or decrease incentives to invest in new or improved products or production methods. Dynamic innovation incentives depend on technological opportunities, the nature of the new product or method, the ability of the firm to appropriate the benefits of the new product or method, and possibly many other market, technological, and human factors. Moreover, it is possible that a dynamic analysis would lead to systematic errors because some effects, such as spillovers that benefit firms or consumers in other industries or at future points in time, are inherently difficult to measure.

Evans and Hylton suggest that an explicit consideration of dynamic effects would lead antitrust enforcers to excuse conduct that they might otherwise challenge if they limit their analysis to static impacts. They illustrate their argument in Figure 3, which shows that a positive dynamic effect of conduct on costs can outweigh negative effects on static total welfare. They describe a stylized version of the Dentsply case, in which the DOJ successfully argued on appeal that Dentsply had monopolized the market for artificial teeth by requiring dental supply dealers to refrain from distributing competing teeth as a condition to accepting Dentsply's premium teeth products.[17] Evans and Hylton observe that Dentsply's exclusive dealing arrangements can adversely affect static welfare by raising prices, but also can promote welfare by enabling Dentsply to profit from lower supply costs from its exclusive dealer network. The Court considered both of these effects in its verdict that Dentsply's conduct was anticompetitive.

Evans and Hylton introduce a new wrinkle, which is an assumption that Dentsply had invested to develop a new and improved type of artificial tooth. They note that Dentsply's incentive to make this investment is its expected profit, which depends on whether Dentsply is permitted to engage in exclusive dealing and particularly on any penalties assessed for its conduct.[18] At a minimum, they argue that any antitrust penalty should take into account potential adverse effects on incentives to create new products, such as the new artificial tooth.

16. Richard Gilbert, *Holding Innovation to an Antitrust Standard*, 3(1) COMPETITION POL'Y INT'L 3 (2007).

17. United States v. Dentsply International, Inc., 277 F.Supp. 2d 387 (D.Del. 2003), 399 F.3d 181 (3d Cir. 2005), cert. denied, 126 S. Ct. 1023 (2006).

18. "The anticipation of an antitrust penalty would diminish its incentive to invest in the activity that creates the market—the new artificial tooth."

In this hypothetical, the firm's investment creates the market. The anticipation of an antitrust penalty would diminish the firm's incentive to invest in the activity that creates the market—the new artificial tooth. More generally, an antitrust penalty has dynamic welfare consequences because it can reduce the incentives to create new products that may incur antitrust liability. Evans and Hylton stop far short of a conclusion that there should be no antitrust penalties for monopolization. However, they argue that "the optimal penalty must consider the dynamic consequences."[19]

I do not question the importance of including dynamic incentives for innovation in an analysis of the competitive effects of monopolizing behavior, even if one might question the scope for innovation in artificial teeth. But the quantification of dynamic incentives is a formidable task. Furthermore, in many market situations, dynamic competitive effects are likely to reinforce static concerns about monopolizing conduct.

Furthermore, in many market situations, dynamic competitive effects are likely to reinforce static concerns about monopolizing conduct.

The authors focus on the example in which firms compete for a durable monopoly. In a winner-take-all or winner-take-most competitive environment, increasing the reward to the winner is likely to strengthen incentives for investments such as research and development that make victory more likely. The canonical example is a patent race.[20] Because a larger reward may generate more innovation, the authors suggest that allowing a firm such as Dentsply to engage in monopolizing behavior could increase welfare by encouraging Dentsply to invest in better artificial teeth to improve the odds that it will enjoy the benefits of a profitable monopoly.

In addition to the difficulties of quantifying these effects, there are two basic reasons to question this logic. First, it goes too far. If increasing rewards generates more innovation in artificial teeth, then why limit the rewards to the monopoly profit that a firm can earn from artificial teeth? Why not permit the firm to engage in conduct that monopolizes other markets as an inducement to invest in more R&D?

A more basic objection to an innovation defense for monopolization is that there is no reason to believe that monopoly encourages innovation in many market environments. At the most general level, there are two basic forces that affect incentives for innovation: the reward to an innovator and the reward to incumbency. Joseph Schumpeter emphasized the innovator's reward in his theory of creative destruction and in his argument that competition is not necessari-

19. Evans & Hylton, *supra* note 1 at 235.

20. *See*, e.g., Jennifer F. Reinganum, *Dynamic games of innovation*, 25 J. ECON. THEORY 21 (1981).

ly the most efficient institution to promote technical progress.[21] Schumpeter emphasized the value of monopoly and large scale as a means to promote investment in research and development and to reap its benefits. The arguments presented in Evans and Hylton reflect a Schumpeterian view of market incentives.

Kenneth Arrow articulated the argument that Schumpeter overlooked the benefits from incumbency.[22] The incentive to innovate is the difference between a firm's profits if it is a successful inventor and its profits if it does not invest in R&D. A monopolist's flow of profits from existing businesses reduces the increase in profit that the firm can earn by innovating. This incumbency or obsolescence effect is a potential drag on the incentive to innovate.

There are many variations on the central themes in Schumpeter and Arrow. These variations can produce incentives for innovation that differ from the pure Schumpeterian or Arrow constructs or that combine elements of the two. Which theory best describes the effects of monopoly power on the incentives to innovate depends on many factors, such as the ability of inventors to appropriate the values of their innovations, whether innovations create new products or lower the costs of producing existing products, and whether innovations increment or drastically change the competitive landscape.

Both economic theory and empirical studies reinforce a conclusion that one cannot presume that monopoly promotes innovation. Many innovations are valuable, but incremental improvements for which Arrow's theory of obsolescence likely would apply.[23] Most empirical studies find little or no support for the proposition that highly concentrated markets invest more in research in development or produce more innovations.[24] One cannot be confident of these conclusions without a comprehensive assessment of market conditions and technological opportunities that affect the opportunities and incentives to invest in innovative effort. But there is little basis in economic theory or empirical research to justify a presumption that monopoly necessarily is good for innovation.

---

21. Joseph A. Schumpeter, The Theory of Economic Development. (Oxford University Press 1961) (1912) and Joseph A. Schumpeter, Capitalism, Socialism and Democracy (Harper Colophon 1976) (1942).

22. Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources to Invention*, *in* The Rate and Direction of Economic Activity, (R. R. Nelson ed., 1962).

23. *See*, e.g., V. Kadiyali, N. J. Vilcassim, & P. K. Chintagunta, *Product line extensions and competitive market interactions: an empirical analysis*, 89(1-2) J. Econometrics 339-63 (1998), Ernst R. Berndt, Iain M. Cockburn, & Karen A. Grépin, *The Impact of Incremental Innovation in Biopharmaceuticals*, 24(2) Pharmacoeconomics 69 (2006), and Morris A. Cohen, Jehoshua Eliashberg, & Teck H. Ho, *An Anatomy of a Decision-Support System for Developing and Launching Line Extensions*, 34(1) J. Marketing Res. 117-29 (1997).

24. For a review of the economic theory of research and development and empirical studies of the relationship between market structure and innovation, see Gilbert, *supra* note 6.

Evans and Hylton fault the economics profession for not rewarding those stalwart researchers who study dynamic competition and attempt to incorporate dynamic effects in competition policy. Instead, they argue that the profession rewards those who search under the lamppost, because that is where the tools of static competition theory shed the most light. The complexities of dynamic competition are sufficiently daunting to limit most econometric studies of market competition to static models. But I disagree that professional rewards pose a barrier to innovation for the analysis of dynamic competition.

But I disagree that professional rewards pose a barrier to innovation for the analysis of dynamic competition.

Most economists would agree with the basic premise in the Evans and Hylton paper that dynamic incentives for innovation are critical to market performance and, where feasible, evaluation of the antitrust consequences of monopolizing conduct should account for these incentives. My impression is that the economics profession looks favorably on research in this area. To test my view of professional incentives, I conducted a simple JSTOR search of recent publications in major economics journals. From 1995 to 2005, the *American Economic Review*, *The Journal of Political Economy*, *Econometrica*, and *The RAND Journal of Economics* published 1,775 articles that mentioned dynamic competition, innovation, or research and development in the abstract. Over the same period, these journals published 641 articles that mentioned merger, monopoly, or monopolization in the abstract. While hardly a definitive study, the evidence suggests that there are significant rewards to those who can unlock the secrets of dynamic competition.

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Schumpeterian Competition and Antitrust

Herbert Hovenkamp

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Schumpeterian Competition and Antitrust

*Herbert Hovenkamp**

Joseph Schumpeter's vision of competition saw it as a destructive process in which effort, assets, and fortunes were continuously destroyed by innovation. This endless process displaced older technologies in order to make way for new ones, but led to economic growth far greater than more stable, conservative alternatives.[1] Schumpeter's vision was striking—in sharp contrast with the conventional neoclassical model of competitive markets, where the focus was on changes in output and price, relatively leisurely shifts in consumer tastes, and exceptional strategic behavior that occasionally dislodged one technology and displaced it by another. Neoclassical competition is a little like watching the ocean when it is calm, while Schumpeterian competition is like watching a raging storm or perhaps even a tidal wave.

As Evans and Hylton so powerfully observe, neoclassical economics is much more comfortable modeling the relatively stable situation than the Schumpeterian one.[2] Economists since Alfred Marshall have observed that the static, partial equilibrium analysis that dominates industrial economics is readily susceptible to mathematics, and many of its rather specific propositions are testable.[3] The Schumpeter model may be testable at a very general level, but

1. Most famously in Joseph A. Schumpeter, Capitalism, Socialism and Democracy (1942), particularly ch. 7, on *The Process of Creative Destruction*. Some of his argument was anticipated in Joseph A. Schumpeter, The Theory of Economic Development (1912).

2. David S. Evans & Keith N. Hylton, *The Lawful Acquisition and Exercise of Monopoly Power and its Implications for the Objectives of Antitrust*, 4(2) Competition Pol'y Int'l (Autumn 2008) [hereinafter Evans & Hylton].

3. On this point, see Joseph A. Schumpeter, History of Economic Analysis 835-838 (1954; revised ed, 1984). See also Herbert Hovenkamp, The Neoclassical Crisis in U.S. Competition Policy, 1890-1955 (SSRN working paper, July 2008), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1156927).

*Ben V. & Dorothy Willie Professor of Law, University of Iowa.

probably not in any sense that antitrust policy finds useful. Schumpeter's analysis is much too concerned with the mostly unmanageable realities of the economy as a whole and with largely unanticipated developments that cannot readily be modeled within the equilibrium-searching forces of neoclassical economics.

To be sure, at a fairly general level the contributions that innovation makes to economic growth can be modeled, and to an extent the models can be empirically tested. For example, the neoclassical growth model developed by Robert W. Solow in the 1950s assumed that innovation is an exogenous factor in the economy, and one can test for its presence and magnitude by assessing the impact of endogenous factors and then assuming that the "residual," or the amount by which growth exceeds these expectations, must be the result of innovation.[4] By contrast, endogenous growth models tend to see innovation as growing out of variables that are within the model of the economy.[5]

Today Schumpeter's conclusion that innovation results much more from convulsive, unexpected changes than from the gradual movement of a market toward competitive equilibrium is fairly well established. What we cannot do, however, is *ex ante* measurement of the long-run effects of specific innovation efforts. Nor can we predict the long-run impact of some observed practice on innovation, certainly not in marginal cases. While innovation overall creates an enormous payoff to society, predicting successful innovations on a case-by-case basis is a fool's errand. Testing like that done of Solow's neoclassical growth model is entirely *ex post*, looking back at the impact of previous innovation in a defined place and time period. Further, it measures aggregate productivity only.

INNOVATION OVERALL CREATES AN ENORMOUS PAYOFF TO SOCIETY, PREDICTING SUCCESSFUL INNOVATIONS ON A CASE-BY-CASE BASIS IS A FOOL'S ERRAND.

A very high percentage of innovation programs fail, but the ones that succeed frequently provide enormous payoffs. And of course the problem is that *ex ante* separation is impossible. If we could predict successful innovations accurately then we could avoid launching the unsuccessful ventures and save enormous resources. These problems have proven to be significant obstacles for economic analysis of specific antitrust claims where the question is likely effects on innovation in the future.

Another problem with measuring innovation or its impact from an *ex ante* perspective is that innovation is so badly behaved in comparison with the ordinary

4. Robert M. Solow, *A Contribution to the Theory of Economic Growth*, 70 Q. J. Econ. 65 (1956); Robert M. Solow, *Technical Change and the Aggregate Production Function*, 3 Rev. Econ. Statistics 312 (1957).

5. See Paul M. Romer, *The Origins of Endogenous Growth*, 8 J. Econ. Persp. 3 (1994).

price and output functions of neoclassical economics. Most changes in price and output are continuous and related to one another. We know enough about many types of practices (price fixing, predatory pricing, mergers, etc.) to predict price and output effects. But the consequences of innovation are often radically indeterminate—sometimes rewarding a large investment by producing nothing at all, or sometimes by producing results that were far different than anyone anticipated.[6] The classic example is Viagra, which was the result of a project seeking treatments for angina. Protracted male erections were initially regarded by the researchers as an undesirable side effect of what would later become one of the most successful pharmaceuticals ever.[7]

As Evans and Hylton observe, in antitrust economic analysis we tend to look at the price and output effects of practices. We evaluate them by asking whether they tend toward increased or decreased output, higher or lower prices, or whether they injure consumers over a testable time period, which is typically quite short. We do not try to show more, because for the most part we cannot answer second-order questions about long-run welfare implications. In the short run a practice may destroy a rival, produce monopoly, and may even appear to impair consumer welfare. But in the longer run it may be part of the very process of creative destruction that Schumpeter believed to be the bedrock of economic progress. Or to say it differently, it may be quite easy for an antitrust economist to predict that a particular exclusionary practice will tend to produce lower market-wide output and higher prices. But it is very likely impossible to predict whether some inchoate innovation that is part of the monopolist's scheme might produce long-term gains that greatly outweigh these short-term losses.

That argument is difficult to dispute, but it is subject to several limitations that serve to dilute its importance. Indeed, the observation may do little more than act as a warning that antitrust economics, and more importantly federal judges, must keep one wandering eye on the long run. Here are the qualifiers I would add:

1. We should not confuse the prospect of innovation with the scope of the intellectual property laws.

2. For many practices positive innovation effects are difficult to foresee even on Schumpeter's own expansive and nonmathematical terms.

3. Many antitrust violations restrain rather than promote innovation.

6. Herbert Hovenkamp, *Restraints on Innovation*, 29 Cardozo L. Rev. 247 (2007) [hereinafter Hovenkamp].

7. *See id*. at 256-257; and Ian H. Osterloh, *The Discovery and Development of Viagra (sildenafil citrate) in* Sildenafil 1, 3 (U. Duzendorfer ed., 2004).

On these points.

First, one must never confuse the prospect of innovation with the scope of the intellectual property ("IP") laws. While Evans and Hylton are speaking generally about competition and innovation as complementary rather than competing products, they refer to this principle by suggesting that there is "no fundamental tension between the policies of antitrust law and intellectual property law; both balance the benefits and costs of static and dynamic competition for the economy as a whole."[8] While that might be true of an economy with ideal competition law and intellectual property law systems, it is hardly true of the world that we actually live in. In fact, both the Patent Act and the Copyright Act have produced bloated regulatory regimes that probably serve to undermine innovation as often as they promote it, and almost certainly do more damage to the innovation process than the antitrust laws themselves. Indeed, there is reason to believe that the patent system fails to carry its freight in any market except perhaps chemicals and pharmaceuticals;[9] and the copyright system has become a playground for special interest groups.[10]

FIRST, ONE MUST NEVER CONFUSE THE PROSPECT OF INNOVATION WITH THE SCOPE OF THE INTELLECTUAL PROPERTY ("IP") LAWS.

Of course, federal judges are not at liberty to rewrite the detailed patent and copyright codes simply because they believe them to be badly designed. But the fact is that one cannot infer that if a conflict appears between competition policy and IP, siding with the latter is more conducive to further innovation. Further, as noted below, the antitrust laws may do a better job of furthering innovation than IP does, provided that it is sufficiently sensitive to the problem of innovation restraints.

The fact is that in the legal situation we currently face, we can probably do far more to promote innovation by reformulating IP policy than by reformulating antitrust policy. For example, a more serious proof of harm requirement could go a long way, perhaps more in copyright than in patent. If an infringement benefits the infringer and its customers and causes no harm to the IP holder then it is a Pareto improvement. Injury should be measured in terms of the *ex ante* incentive to create the protected work in the first place.[11]

---

8. Evans & Hylton, *supra* note 2 *at* 204.

9. James Bessen & Michael J. Meurer, Patent Failure: How Judges, Bureaucrats, and Lawyers put Innovators at Risk (2008); and Robert P. Merges, *One Hundred Years of Solicitude: Intellectual Property Law, 1900-2000*, 88 Cal. L. Rev. 2187, 2336 (2000). Even relative conservatives such as Landes and Posner find overprotection. *See* William M. Landes & Richard A. Posner, The Economic Structure of Intellectual Property Law, Ch. 1 (2003).

10. Christina Bohannan, *Reclaiming Copyright*, 23 Cardozo Arts & Ent. L. J. 567, 568 (2006).

11. *See* Christina Bohannan, *Copyright Harm, Foreseeability and Fair Use*, 85 Washington Univ. L. Rev. 969 (2007).

My second point is that, for many practices challenged by the antitrust laws, innovation effects are difficult to assess or even foresee on Schumpeter's own non-mathematical terms. Not every antitrust violation has significant implications for innovation. Pricing practices are a good example. When properly defined, both price fixing and predatory pricing involve changing the price of a good in anti-competitive ways. Neither one has obvious implications for innovation subject to one exception: one can always argue that a firm will use monopoly profits to innovate more, and that the gains from the resulting innovation might possibly far exceed the losses from short-run consumer injuries. But this argument proves too much and justifies monopoly no matter how created or maintained.

On the other hand, if a practice challenged under the antitrust laws actually furthers innovation, one would expect that the defendant could provide an explanation and some evidence. For example, if exclusive dealing really is being used to protect the market for an incipient product then the defendant should be able to tell us and this information should be incorporated into our rule of reason analysis.

Finally, the third point is that many antitrust violations restrain rather than promote innovation. Indeed, there are good reasons for believing that market-dominating firms or joint ventures with a significant investment in their technology are more likely to use exclusionary practices to restrain the innovations of rivals or potential rivals than to develop or promote their own innovations. For the most part, the technology and markets of dominant firms are well established and they tend to profit from stable growth. By contrast, the small firm seeking entry must shake up the pot.

On the other hand, if a practice challenged under the antitrust laws actually furthers innovation, one would expect that the defendant could provide an explanation and some evidence.

Evans and Hylton give the very interesting example borrowed from the *Dentsply* case[12] of a firm that develops a new and innovative but unpatentable tooth. It must then use exclusive dealing in order to capitalize on its investment by excluding rivals via a restraint on market access rather than the IP laws.[13] Whether that story is plausible or not, there is an alternative story that is at least as plausible. Suppose that a smaller rival has developed an innovative artificial tooth that may very well be patentable, but success depends on market access. Further, this artificial tooth threatens to take a significant share of the market once it is successfully deployed. Dentsply's exclusive dealing serves to deny it market access.

12. United States v. Dentsply Int'l, Inc., 399 F.3d 181 (3d Cir. 2005), cert. denied, 546 U.S. 1089 (2006) (condemning exclusive dealing as unlawful monopolization under §2 of the Sherman Act). See 3B Phillip E. Areeda & Herbert Hovenkamp, Antitrust Law ¶768 (3d Ed. 2008).

13. Evans & Hylton, *supra* note 2 at 235. [TAN 108].

In this case the antitrust violation has served to restrain rather than promote innovation. The story is more plausible than the Evans/Hylton story for two reasons. First, in this setting market-shifting innovations are more likely to come from smaller firms. Once it has attained dominance, a firm's interest in creative destruction becomes greatly diminished because it is as likely to be the victim as the enabler. Indeed, often a firm's investment in its own technology creates a form of path dependence. Its vested interests lie much more in preserving what it has rather than producing a huge market-shifting innovation. As a result, it may be inclined to innovate in ways that take advantage of technology and property rights in which it has already made an investment. By contrast, the smaller rival succeeds by differentiating its product from that of the dominant firm.[14] Second, exclusive dealing by a dominant firm is very likely a more effective means of excluding a smaller rival's innovation than it is of promoting the dominant firm's own innovation. Indeed, Evans and Hylton have to assume that the IP laws provide no protection in order to make their story work.

Finally, modeling the incentives to restrain innovation is at least potentially more tractable than modeling innovation itself, although measuring long-run effects is often just as difficult. Restraints on innovation typically show up in creation or perpetuation of monopoly prices, reduced output, and the like. That is, a dominant firm typically restrains innovation in order to prevent its market position from eroding. Such gains to the monopolist are subject to the ordinary measurement tools of forensic economics.

The boycott situation is similar to the vertical exclusion story. Consider the *Allied Tube* case, which involved a boycott by the manufacturers of steel electrical conduit intended to exclude a market shifting innovation—conduit made from PVC (polyvinyl chloride).[15] PVC conduit was cheaper, easier to work with, and did not short out when it came into contact with an electrical wire. Allied, whose manufacturing commitment was entirely to steel, plainly foresaw what later became a market reality: plastic conduit would swamp the field. It therefore organized a boycott designed to exclude PVC conduit from the market by writing its use out of municipal building codes.

The *Allied Tube* story is a particularly easy and obvious one, because PVC conduit was an innovation in its final stages of market preparation. Its market success was reasonably foreseeable by the time the antitrust violation occurred.[16]

14. *See*, e.g., S. J. Liebowitz & S. E. Margolis, *Path Dependence, Lock-in and History*, 11 J. L., ECON., AND ORG. 205 (1995); J. Farrell & G. Saloner, *Installed Base and Compatability: Innovation, Product Preannouncements, and Predation*, 76 AM. ECON. REV. 940 (1986).

15. Allied Tube & Conduit Corp. v. Indian Head, Inc., 486 U.S. 492, 496-497 (1988).

16. For a similar story, see Am. Soc'y of Mech. Eng'rs v. Hydrolevel Corp., 456 U.S. 556 (1982), which involved an agreement among the members of an accreditation association to suppress a superior valve technology, with the result that the plaintiff's valve could not be marketed. See 13 HERBERT HOVENKAMP, ANTITRUST LAW ¶2115 (2d ed. 2005).

More incipient innovations are easier for dominant firms to exclude. Further, the violations are more difficult to detect, and it is certainly more difficult to prove injury. Consider the pressure that Microsoft placed on Intel to stop its Java-enabled chip R&D program lest Microsoft stop cooperating with Intel on future projects for chips that ran on the Windows platform.[17] Java is a multi-platform processing language. At the time the Java-enabled chip threatened to make alternative operating systems "compatible" with Microsoft Windows by enabling software developers to write software that would operate on multiple platforms and communicate seamlessly with one another. By excluding Java, Microsoft stood to gain the higher market share and prices that resulted from suppressing the innovative competition that threatened to make Windows one of many alternative platforms. Consumers lost uncertain value, depending on the likelihood that the chip would have succeeded and its market impact. Or consider the many, many cases involving *Walker Process* style patent infringement lawsuits based on improperly obtained patents or on irrationally broad patent claims.[18] Many of these are lawsuits brought by large firms with a heavy investment in their existing technology, designed to oust the innovative technology of a less well financed rival.

The obvious question raised is: When is an antitrust violation more likely to be innovation enhancing rather than innovation restraining?

In sum, one place the antitrust laws could be more aggressive than they are today is when the stars are in alignment. An important corollary of the premise that innovation contributes much more to economic growth than does price competition and short run efficiency is that a *restraint* on innovation can do much more harm. Restraints such as the ones at issue in *Allied Tube* and *Microsoft* simultaneously produce higher prices in the dominant firm's market and innovation in incipient markets is delayed or not permitted to materialize.

The obvious question raised is: When is an antitrust violation more likely to be innovation enhancing rather than innovation restraining? While that question may be very difficult to answer in some cases, in others it appears not to be. For example, where a dominant firm is using an exclusionary practice to protect its established investment from an incipient technology, harm to innovation seems to be the most likely outcome.

17. *See* Hovenkamp, *supra* note 6 *at* 249-250 (discussing United States v. Microsoft Corp., 84 F. Supp. 2d 9, 107 (D.D.C. 1999) (fact finding #396), aff'd in part, rev'd in part, but affirmed on this issue, United States v. Microsoft Corp., 253 F.3d 34, 77 (D.C. Cir. 2001) (noting Microsoft's statements to Intel that "cooperation with Sun and Netscape to develop a Java runtime environment . . . was one of the issues threatening to undermine cooperation between Intel and Microsoft"), cert. denied, 534 U.S. 952 (2001)).

18. Walker Process Equip., Inc. v. Food Mach. & Chem. Corp., 382 U.S. 172 (1965). See 3 Phillip E. Areeda & Herbert Hovenkamp, Antitrust Law ¶706 (3d ed. 2008); Herbert Hovenkamp, The Walker Process Doctrine: Infringement Lawsuits as Antitrust Violations (SSRN working paper, Sep. 2008) *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1259877).

The most difficult set of cases is likely to involve joint ventures and at least some mergers, where the dangers of collusion must be set against the very real possibilities that the union will promote significant innovation. Standard setting is another area. The potential cost savings from reliable standards can be enormous, but the process can be used to exclude novel technologies. For example, the *Hydrolevel case* involved a situation where a standard setting committee within the American Society of Mechanical Engineers was manipulated into denying approval, and thus market access, to the plaintiff's innovative valve when the dominant firm perceived a market threat.[19]

It is also worth noting that restraints on innovation can be addressed under both antitrust policy and a properly formulated IP policy. For example, the doctrine of patent or copyright "misuse" can be a device for combating contractual devices or overly broad claims by IP holders that tend to restrain rival innovations. But misuse claims apply only against IP holders, and typically only in defenses against infringement lawsuits.[20] The restraints at issue in cases like *Allied Tube* and *Microsoft* did not involve firms acting as IP holders but rather as market participants with considerable leverage over others and existing technologies they wished to protect.

An increased antitrust concern with restraints on innovation places a premium on government enforcement for the very reason that Evans and Hylton suggest: economic proof of the effects of restraints on innovation is so difficult to obtain, thus making proof of private injury and damages very difficult. A case in point is the tagalong litigation in *Kloth v. Microsoft*, where the Fourth Circuit ultimately held that private plaintiffs could not obtain damages for Microsoft's suppression of Intel's Java chip program because they were too speculative. As the court observed, "It would be entirely speculative and beyond the competence of a judicial proceeding to create in hindsight a technological universe that never came into existence. . . .[21] While private plaintiffs must show causation and actual injury for damages or threatened injury for an injunction, the United States or Federal Trade Commission acting as enforcer need show only that the antitrust laws have been violated.[22]

19. Am. Soc'y of Mech. Eng'rs v. Hydrolevel Corp., 456 U.S. 556 (1982). See also Allied Tube & Conduit Corp. v. Indian Head, Inc., 486 U.S. 492, 496-497 (1988).

20. See 10 Phillip E. Areeda, Einer Elhauge, & Herbert Hovenkamp, Antitrust Law ¶1781 (2d ed. 2004); Christina Bohannan, Intellectual Property Misuse and Foreclosure (2008).

21. Kloth v. Microsoft Corp., 444 F.3d 312, 323 (4th Cir. 2006); see Hovenkamp, *supra* note 6, *at* 259.

22. See 2 & 2A Phillip E. Areeda & Herbert Hovenkamp, Antitrust Law ¶¶303, 326 (3d ed. 2007).

Conclusion

Schumpeter was correct that over the long run the gains from innovation dwarf the gains from government intervention to make the economy more competitive under the traditional criteria of price and output. It follows that the losses resulting from restraints on innovation could be very large as well. The problem of ex ante measurement of the social losses that result from a restraint on an undeveloped innovation is equivalent to the problem of ex ante measurement of the gains that the innovation would have produced had the innovation process been permitted to run its course. In both cases an ex ante assessment could be virtually impossible and in any case would be highly speculative.

But that does not necessarily mean that antitrust cannot do anything about the problem. In some cases, all that is necessary is to consider short-run consequences for competition and ignore innovation possibilities that are too remote to see. In other cases, one should consider whether an innovation or a restraint on innovation is the more likely outcome. The likelihood that a practice furthers innovation should serve to weaken or perhaps even undermine the antitrust concern. By contrast, the likelihood that a practice restrains innovation should deserve a much closer look.

Volume 4 | Number 2 | Autumn 2008

Competition Policy International

Joseph Schumpeter on Competition

Thomas K. McCraw

Published in Competition Policy International (print ISSN 1554-0189, online ISSN 1554-6853), Autumn 2008, Vol. 4, No. 2. For articles and more information, visit www.globalcompetitionpolicy.org.

Joseph Schumpeter on Competition

*Thomas K. McCraw**

The following documents illustrate the relevance of Schumpeter's thought to competition policy. Part I is an introduction to Schumpeter's ideas; Part II a series of excerpts from his book, *Capitalism, Socialism and Democracy*;[1] Part III a 1951 critique of his stance toward antitrust by the economist Edward S. Mason; and Part IV an evaluation of the current use of Schumpeter's theories in discussions of competition policy.

1. JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY, (1st edition 1942) [hereinafter CAPITALISM]

*The author is the Isidor Straus Professor of History Emeritus at Harvard Business School. A Pulitzer Prize-winning historian, he is the author of *Prophet of Innovation: Joseph Schumpeter and Creative Destruction* (Cambridge, Mass.: Belknap Press of Harvard University Press, 2007).

I. Introduction

During the 1980s, there began a spirited revival of interest in the writings of Joseph Schumpeter (1883-1950), spurred by renewed attention to his seminal works on entrepreneurship and innovation. The movement gathered so much strength that citations to Schumpeter by scholars and journalists began to exceed those to Keynes, a phenomenon that would have seemed unthinkable only a few years earlier. In 2000, *Business Week* ran a two-page spread titled "America's Hottest Economist Died 50 Years Ago."[2]

This upsurge of interest has migrated to numerous areas of inquiry, including competition policy. Three pertinent articles among many that might be cited are Schmalensee's "Antitrust in Schumpeterian Industries",[3] Katz and Shelanski's "'Schumpeterian'" Competition and Antitrust Policy in High-Tech Markets",[4] and Baker's "Beyond Schumpeter vs Arrow: How Antitrust Fosters Innovation".[5]

This sudden attention to Schumpeter's work by antitrust scholars is a bit surprising, because very little of his vast body of writing even purports to address antitrust directly. In all, he published eleven books and scores of articles and reviews—a staggering total of about 3.5 million words. But the word "antitrust" appears almost nowhere.

A reading of Schumpeter's work in its entirety makes it clear that he disavowed advocacy of any kind. He opposed the development of "schools" of economic thought, even though he had tremendous respect for the achievements of great scholars across the ideological spectrum, from his fellow Austrians Mises and Hayek on the far Right to Marx on the far Left. Schumpeter believed economics to be a science, and he conceived the task of scientists as the quest for truth, not the service of policy. He was convinced that direct pursuit of the second goal would inevitably corrupt the first. So, if one searches his work for explicit guides to antitrust policy, one may find, as Gertrude Stein said of the city of Oakland, that "There is no there there."

This is probably why Schumpeter's writings were neglected for so long by antitrust scholars and policymakers. He makes no appearance, for example, in Bork's *The Antitrust Paradox*[6]. Nor is his name prominent in most other antitrust

2. Charles J. Whalen, *America's Hottest Economist Died 50 Years Ago*, Bus. Wk. (December 11, 2000)

3. Richard Schmalensee, *Antitrust in Schumpeterian Industries*, 90 Am. Econ. Rev. 192-196 (2000)

4. Michael L. Katz & Howard A. Shelanski, "*Schumpeterian" Competition and Antitrust Policy in High-Tech Markets*, 14 Competition 47 (2005)

5. Jonathan B. Baker, *Beyond Schumpeter vs. Arrow: How Antitrust Fosters Innovation*, 74 Antitrust L. J. 575-602 (2007)

6. Robert Bork, The Antitrust Paradox: A Policy at War with Itself (1978)

treaties and texts that appeared either before or after Bork's book. There is no explicit there there.

Taken as a whole, Schumpeter's writings fall into the tradition of grand social theory exemplified by European thinkers such as August Comte, Karl Marx, John Stuart Mill, and Max Weber. Although he spent his academic career as a professor of economics—teaching at two universities in Austria and one in Germany before moving permanently to Harvard in 1932—his work freely crosses disciplinary lines. In addition to economics, it encompasses sociology, psychology, law, business, and some mathematics.[7]

TAKEN AS A WHOLE, SCHUMPETER'S WRITINGS FALL INTO THE TRADITION OF GRAND SOCIAL THEORY EXEMPLIFIED BY EUROPEAN THINKERS SUCH AS AUGUST COMTE, KARL MARX, JOHN STUART MILL, AND MAX WEBER.

Among all the sister disciplines of economics, Schumpeter most prized history. Concerning what he regarded as the three basic building blocks of economics—theory, statistics, and history—he wrote that the last "is by far the most important." In his final book, he issued this remarkable credo:

> "I wish to state right now that if, starting my work in economics afresh, I were told that I could study only one of the three but could have my choice, it would be economic history that I should choose. And this on three grounds. First, the subject matter of economics is essentially a unique process in historic time. Nobody can hope to understand the economic phenomena of any, including the present, epoch who has not an adequate command of historical *facts* and an adequate amount of historical *sense* or of what may be described as *historical experience*. Second, the historical report cannot be purely economic but must inevitably reflect also "institutional" facts that are not purely economic: therefore it affords the best method for understanding how economic and non-economic facts *are* related to one another and how the various social sciences *should* be related to one another. Third, it is, I believe, the fact that most of the fundamental errors currently committed in economic analysis are due to lack of historical experience more often than to any other shortcoming of the economist's equipment."[8]

7. He had little talent for advanced math, but he thought it vitally important: along with Irving Fisher and Ragnar Frisch, he founded the Econometric Society, and he wrote the lead article for the first issue of *Econometrica*.

8. JOSEPH SCHUMPETER, HISTORY OF ECONOMIC ANALYSIS, 12-13 (1954)

Schumpeter came to this conclusion reluctantly. For almost his entire life he regarded himself primarily as a theorist, and he achieved some unique successes. His book *The Theory of Economic Development*[9] is one of the classic economic texts of the twentieth century. It remains to this day the best argument for the addition of entrepreneurship as a fourth factor of production along with land, labor, and capital. He was a leader not only in the study of entrepreneurship, but also in his emphasis on credit creation, business strategy, and—above all—innovation.

It was during the 1930s, some 25 years after his first important publications, that Schumpeter began fully to appreciate the importance of history. His 1,095 page *Business Cycles*[10] is as much a work of history as of theory; and his history (which highlights innovation and covers the entire capitalist epoch in Britain, Germany, and the United States), coheres far better than his theory. The latter is spoiled by a heroic but futile attempt to fit patterns of booms and busts into determinate periods defined by other theorists: Joseph Kitchin (40 month cycles), Clément Juglar (8-10 year) and Nikolai Kondratieff (50-60 year).

HE WAS A LEADER NOT ONLY IN THE STUDY OF ENTREPRENEURSHIP, BUT ALSO IN HIS EMPHASIS ON CREDIT CREATION, BUSINESS STRATEGY, AND—ABOVE ALL—INNOVATION.

But even in this book, written during the Great Depression, Schumpeter explicitly disavows advocacy and offers no solution to the economic crisis. "I recommend no policy and propose no plan," he writes in the preface; his book can "be used to derive practical conclusions of the most conservative or the most radical complexion." *Business Cycles* was an exercise in value-neutral science, and in this respect it typified nearly all of Schumpeter's writings. Some 33 years earlier, in the preface to his very first book, he had written something quite similar: "I hold aloof from practical politics and recognize no purpose other than knowledge."[11]

The subtitle of *Business Cycles*—"A Theoretical, Historical, and Statistical Analysis of the Capitalist Process"—well expresses the extraordinary reach of what Schumpeter was trying to do in 1939. Although the book failed as the *magnum opus* he was hoping for, the immense amount of empirical research on specific firms and industries that went into it prepared him, as nothing else could have done, to write his most famous work, *Capitalism, Socialism and Democracy*. The book appeared only three years after *Business Cycles* and is one of the seminal non-fiction works of the last hundred years, in any field. For competition policy, it is the most relevant of all his works, but, again, it offers no explicit formulas.

9. Joseph Schumpeter, The Theory of Economic Development (1911, Eng. Translation 1934)

10. Joseph Schumpeter, Business Cycles: A Theoretical, historical, and Statistical Analysis of the Capitalist Process (1939) [hereinafter Cycles]

11. *Id.* at vi. *Also* Joseph Schumpeter, Das Wesen und der Hauptinhalt von theoretichen Nationalökonomie (The Nature and Content of Theoretical Economics) (1908) i-vi (Erich Schneider trans.)

Although very much a book of its time—Schumpeter wrote it in 30 months during 1939-1942, against the uniquely atypical backdrop of the Great Depression and World War II—it is also a book for the ages. Among its many virtues, *Capitalism, Socialism and Democracy* contains one of the best explications of capitalism ever written. The book's most quoted phrase, "creative destruction," is perhaps second only to Adam Smith's "invisible hand" as the best-known metaphor in all of economics, a discipline rich in metaphors.

AMONG ITS MANY VIRTUES, *CAPITALISM, SOCIALISM AND DEMOCRACY* CONTAINS ONE OF THE BEST EXPLICATIONS OF CAPITALISM EVER WRITTEN.

II. Passages from *Capitalism, Socialism and Democracy*

The analysis in *Capitalism, Socialism and Democracy* is profound, and it provides at least some implicit guides to competition policy—though one must be extremely careful in applying anything Schumpeter says to a particular case or controversy. In parts of the book, he may appear to prefer large firms to small ones, but this is not what he believed, as his many other writings clearly show. His litmus test for competition policy, and almost any other policy, has little to do directly with firm size or industry structure, and everything to do with innovation. This is clear in *The Theory of Economic Development* and equally so in *Business Cycles*, where he writes repeatedly of "New Men" founding "New Firms" and thereby forcing "Innovation" (he capitalizes all three terms). The problem in specific cases involving public policy—as the mixed record of antitrust shows so clearly—is in making such a judgment *about the future* with much accuracy.

In *Capitalism, Socialism and Democracy*, Schumpeter felt it necessary to explain the workings of big business because when he began writing the book during the late 1930s, large firms were under very severe attack. They stood in lower popular repute than at any other time in American history. Hence the candid tone and very strong language in the following excerpts from the book, which contain the heart of Schumpeter's analysis of capitalism (footnotes are quoted as cited):

> "If we look more closely at the conditions (. . .) that must be fulfilled in order to produce perfect competition, we realize immediately that outside of agricultural mass production there cannot be many instances of it. (. . .) every grocer, every filling station, every manufacturer of gloves or shaving cream or handsaws has a small and precarious market of his own, which he tries—must try—to build up and to keep by price strategy, quality strategy—"product differentiation"—and advertising. Thus we get a completely different pattern which there seems to be no reason

to expect to yield the results of perfect competition and which fits much better into the monopolistic schema. In these cases we speak of Monopolistic Competition. Their theory has been one of the major contributions to postwar economics.[12] (. . .)

As soon as the prevalence of monopolistic competition or of oligopoly or of combinations of the two is recognized, many of the propositions which the Marshall-Wicksell generation of economists used to teach with the utmost confidence become either inapplicable or much more difficult to prove. This holds true, in the first place, of the propositions turning on the fundamental concept of equilibrium, i.e. a determinate state of the economic organism, toward which any given state of it is always gravitating and which displays certain simple properties. In the general case of oligopoly there is in fact no determinate equilibrium at all and the possibility presents itself that there may be an endless sequence of moves and countermoves, an indefinite state of warfare between firms. (. . .)

The theories of monopolistic and oligopolistic competition and their popular variants may in two ways be made to serve the view that capitalist reality is unfavorable to maximum performance in production. One may hold that it always has been so and that all along output has been expanding in spite of the secular sabotage perpetrated by the managing bourgeoisie. Advocates of this proposition would have to produce evidence to the effect that the observed rate of increase can be accounted for by a sequence of favorable circumstances unconnected with the mechanism of private enterprise and strong enough to overcome the latter's resistance However, those who espouse this variant at least avoid the trouble about historical fact that the advocates of the alternative proposition have to face. This avers that capitalist reality once tended to favor maximum productive performance, or at all events productive performance so considerable as to constitute a major element in any serious appraisal of the system; but that the later spread of monopolist structures, killing competition, has by now reversed that tendency.

First, this involves the creation of an entirely imaginary golden age of perfect competition that at some time somehow metamorphosed itself into the monopolistic age, whereas it is quite clear that perfect competition has at no time been more of a reality than it is at present. Secondly, it is necessary to point out that the rate of increase in output did not decrease from the nineties from which, I suppose, the prevalence of the largest-size concerns, at least in manufacturing industry, would have to be dated; that there is nothing in the behavior of the time series of total output to suggest a "break in trend"; and, most important of all, that the modern standard of life of the masses evolved during the period

12. See, in particular, E.S. Chamberlin, Theory of Monopolistic Competition (1933), and Joan Robinson, The Economics of Imperfect Competition (1933).

of relatively unfettered "big business." If we list the items that enter the modern workman's budget and from 1899 on observe the course of their prices not in terms of money but in terms of the hours of labor that will buy them—i.e., each year's money prices divided by each year's hourly wage rates—we cannot fail to be struck by the rate of the advance which, considering the spectacular improvement in qualities, seems to have been greater and not smaller than it ever was before. If we economists were given less to wishful thinking and more to the observation of facts, doubts would immediately arise as to the realistic virtues of a theory that would have led us to expect a very different result. Nor is this all. As soon as we go into details and inquire into the individual items in which progress was most conspicuous, the trail leads not to the doors of those firms that work under conditions of comparatively free competition but precisely to the doors of the large concerns—which, as in the case of agricultural machinery, also account for much of the progress in the competitive sector—and a shocking suspicion dawns upon us that big business may have had more to do with creating that standard of life than with keeping it down.

The conclusions alluded to at the end of the preceding chapter are in fact almost completely false. Yet they follow from observations and theorems that are almost completely[13] true. Both economists and popular writers have once more run away with some fragments of reality they happened to grasp. These fragments themselves were mostly seen correctly. Their formal properties were mostly developed correctly. But no conclusions about capitalist reality as a whole follow from such fragmentary analyses. If we draw them nevertheless, we can be right only by accident. That has been done. And the lucky accident did not happen.

THE ESSENTIAL POINT TO GRASP IS THAT IN DEALING WITH CAPITALISM WE ARE DEALING WITH AN EVOLUTIONARY PROCESS.

The essential point to grasp is that in dealing with capitalism we are dealing with an evolutionary process. It may seem strange that anyone can fail to see so obvious a fact which moreover was long ago emphasized by Karl Marx. Yet that fragmentary analysis which yields the bulk of our propositions about the functioning of modern capitalism persistently neglects it. Let us restate the point and see how it bears upon our problem.

---

13. As a matter of fact, these observations and theorems are not completely satisfactory. The usual expositions of the doctrine of imperfect competition fail in particular to give due attention to the many and important cases in which, even as a matter of static theory, imperfect competition approximates the results of perfect competition. There are other cases in which it does not do this, but offers compensations which, while not entering any output index, yet contribute to what the output index is in the last resort intended to measure—the cases in which a firm defends its market by establishing a name for quality and service for instance. However, in order to simplify matters, we will not take issue with that doctrine on its own ground.

Capitalism, then, is by nature a form or method of economic change and not only never is but never can be stationary. And this evolutionary character of the capitalist process is not merely due to the fact that economic life goes on in a social and natural environment which changes and by its change alters the data of economic action; this fact is important and these changes (wars, revolutions and so on) often condition industrial change, but they are not its prime movers. Nor is this evolutionary character due to a quasi-automatic increase in population and capital or to the vagaries of monetary systems of which exactly the same thing holds true. The fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers' goods, the new methods of production or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates.

As we have seen in the preceding chapter, the contents of the laborer's budget, say from 1760 to 1940, did not simply grow on unchanging lines but they underwent a process of qualitative change. Similarly, the history of the productive apparatus of a typical farm, from the beginnings of the rationalization of crop rotation, plowing and fattening to the mechanized thing of today—linking up with elevators and railroads—is a history of revolutions. (...) of industrial mutation—if I may use that biological term—that incessantly revolutionizes[14] the economic structure *from within*, incessantly destroying the old one, incessantly creating a new one. This process of Creative Destruction is the essential fact about capitalism. It is what capitalism consists in and what every capitalist concern has got to live in. This fact bears upon our problem in two ways.

THIS PROCESS OF CREATIVE DESTRUCTION IS THE ESSENTIAL FACT ABOUT CAPITALISM. IT IS WHAT CAPITALISM CONSISTS IN AND WHAT EVERY CAPITALIST CONCERN HAS GOT TO LIVE IN.

First, since we are dealing with a process whose every element takes considerable time in revealing its true features and ultimate effects, there is no point in appraising the performance of that process *ex visu* of a given point of time; we must judge its performance over time, as it unfolds through decades or centuries. A system—any system, economic or other—that at *every* given point of time fully utilizes its possibilities to the best advantage may yet in the long run be inferior to a system that does so at *no* given point of time, because the latter's failure to do so may be a condition for the level or speed of long-run performance.

Second, since we are dealing with an organic process, analysis of what happens in any particular part of it—say, in an individual concern or

14. Those revolutions are not strictly incessant; they occur in discrete rushes which are separated from each other by spans of comparative quiet. The process as a whole works incessantly however, in the sense that there always is either revolution or absorption of the results of revolution, both together forming what are know as business cycles.

industry—may indeed clarify details of mechanism but is inconclusive beyond that. Every piece of business strategy acquires its true significance only against the background of that process and within the situation created by it. It must be seen in its role in the perennial gale of creative destruction; it cannot be understood irrespective of it or, in fact, on the hypothesis that there is a perennial lull.

But economists who, *ex visu* of a point of time, look for example at the behavior of an oligopolist industry—an industry which consists of a few big firms—and observe the well-known moves and countermoves within it that seem to aim at nothing but high prices and restrictions of output are making precisely that hypothesis. They accept the data of the momentary situation as if there were no past or future to it and think that they have understood what there is to understand if they interpret the behavior of those firms by means of the principle of maximizing profits with reference to those data. The usual theorist's paper and the usual government commission's report practically never try to see that behavior, on the one hand, as a result of a piece of past history and, on the other hand, as an attempt to deal with a situation that is sure to change presently—as an attempt by those firms to keep on their feet, on ground that is slipping away from under them. In other words, the problem that is usually being visualized is how capitalism administers existing structures, whereas the relevant problem is how it creates and destroys them. As long as this is not recognized, the investigator does a meaningless job. As soon as it is recognized, his outlook on capitalist practice and its social results changes considerably.[15] (. . .)

It is hardly necessary to point out that competition of the kind we now have in mind acts not only when in being but also when it is merely an ever-present threat. It disciplines before it attacks. The businessman feels himself to be in a competitive situation even if he is alone in his field or if, though not alone, he holds a position such that investigating government experts fail to see any effective competition between him and any other firms in the same or a neighboring field and in consequence conclude that his talk, under examination, about his competitive sorrows is all make-believe. In many cases, though not in all, this will in the long run enforce behavior very similar to the perfectly competitive pattern. (. . .)

In the case of retail trade the competition that matters arises not from additional shops of the same type, but from the department store, the chain store, the mail-order house and the supermarket which are bound

15. It should be understood that it is only our appraisal of economic performance and not our moral judgment that can be so changed. Owing to its autonomy, moral approval or disapproval is entirely independent of our appraisal of social (or any other) results, unless we happen to adopt a moral system such as utilitarianism which makes moral approval and disapproval turn on them *ex definitione*.

to destroy those pyramids sooner or later.[16] Now a theoretical construction which neglects this essential element of the case neglects all that is most typically capitalist about it; even if correct in logic as well as in fact, it is like *Hamlet* without the Danish prince. (. . .)

Practically any investment entails, as a necessary complement of entrepreneurial action, certain safeguarding activities such as insuring or hedging. Long-range investing under rapidly changing conditions, especially under conditions that change or may change at any moment under the impact of new commodities and technologies, is like shooting at a target that is not only indistinct but moving—and moving jerkily at that. Hence it becomes necessary to resort to such protecting devices as patents or temporary secrecy of processes or, in some cases, long-period contracts secured in advance. But these protecting devices which most economists accept as normal elements of rational management[17] are only special cases of a larger class comprising many others which most economists condemn although they do not differ fundamentally from the recognized ones.

If for instance a war risk is insurable, nobody objects to a firm's collecting the cost of this insurance from the buyers of its products. But that risk is no less an element in long-run costs, if there are no facilities for insuring against it, in which case a price strategy aiming at the same end will seem to involve unnecessary restriction and to be productive of excess profits. Similarly, if a patent cannot be secured or would not, if secured, effectively protect, other means may have to be used in order to justify the investment. Among them are a price policy that will make it possible to write off more quickly than would otherwise be rational, or additional investment in order to provide excess capacity to be used only for aggression or defense. Again, if long-period contracts cannot be entered into in advance, other means may have to be devised in order to tie prospective customers to the investing firm.

In analyzing such business strategy *ex visu* of a given point of time, the investigating economist or government agent sees price policies that seem to him predatory and restrictions of output that seem to him synonymous with loss of opportunities to produce. He does not see that restrictions of this type are, in the conditions of the perennial gale, incidents, often unavoidable incidents, of a long-run process of expansion

16. The mere threat of their attack cannot, in the particular conditions, environmental and personal, or small-scale retail trade, have its usual disciplining influence, for the small man is too much hampered by his cost structure and, however well he may manage within his inescapable limitations, he can never adapt himself to the methods of competitors who can afford to sell at the price at which he buys.

17. Some economists, however, consider that even those devices are obstructions to progress which, though perhaps necessary in capitalist society, would be absent in a socialist one. There is some truth in this. But that does not affect the proposition that the protection afford by patents and so on is, in the conditions of a profit economy, on balance a propelling and not an inhibiting factor.

which they protect rather than impede. There is no more of paradox in this than there is in saying that motorcars are traveling faster they otherwise would *because* they are provided with brakes.

2. This stands out most clearly in the case of those sectors of the economy which at any time happen to embody the impact of new things and methods on the existing industrial structure. The best way of getting a vivid and realistic idea of industrial strategy is indeed to visualize the behavior of new concerns or industries that introduce new commodities or processes (such as the aluminum industry) or else reorganize a part or the whole of an industry (such as, for instance, the old Standard Oil Company).

As we have seen, such concerns are aggressors by nature and wield the really effective weapon of competition. Their intrusion can only in the rarest of cases fail to improve total output in quantity or quality, both through the new method itself—even if at no time used to full advantage—and through the pressure it exerts on the preexisting firms. But these aggressors are so circumstanced as to require, for purposes of attack and defense, also pieces of armor other than price and quality of their product which, moreover, must be strategically manipulated all along so that at any point of time they seem to be doing nothing but restricting their output and keeping prices high.

On the one hand, largest-scale plans could in many cases not materialize at all if it were not known from the outset that competition will be discouraged by heavy capital requirements or lack of experience, or that means are available to discourage or checkmate it so as to gain the time and space for further developments. (. . .)

Again this requires strategy that in the short run is often restrictive. In the majority of successful cases this strategy just manages to serve its purpose. In some cases, however, it is so successful as to yield profits far above what is necessary in order to induce the corresponding investment. These cases then provide the baits that lure capital on to untried trails. Their presence explains in part how it is possible for so large a section of the capitalist world to work for nothing: in the midst of the prosperous twenties just about half of the business corporations in the United States were run at a loss, at zero profits, or at profits which, if they had been foreseen, would have been inadequate to call forth the effort and expenditure involved.

Old concerns and established industries, whether or not directly attacked, still live in the perennial gale.

Our argument however extends beyond the cases of new concerns, methods and industries. Old concerns and established industries, whether or not directly attacked, still live in the perennial gale. Situations emerge in the process of creative destruction in which many firms may have to perish that nevertheless would be able to live on vigorously and usefully if they could weather a particular storm. (. . .)

All this is of course nothing but the tritest common sense. But it is being overlooked with a persistence so stubborn as sometimes to raise the question of sincerity. And it follows that, within the process of creative destruction, all the realities of which theorists are in the habit of relegating to books and courses on business cycles, there is another side to industrial self-organization than that which these theorists are contemplating. (. . .)

It is certainly as conceivable that an all-pervading cartel system might sabotage all progress as it is that it might realize, with smaller social and private costs, all that perfect competition is supposed to realize. This is why our argument does not amount to a case against state regulation. It does show that there is no general case for indiscriminate "trust-busting" or for the prosecution of everything that qualifies as a restraint of trade. Rational as distinguished from vindictive regulation by public authority turns out to be an extremely delicate problem which not every government agency, particularly when in full cry against big business, can be trusted to solve.[18] (. . .)

Of course, plenty of cases of genuine price rigidity remain—of prices which are being kept constant as a matter of business policy or which remain unchanged because it is difficult to change, say, a price set by a cartel after laborious negotiations. In order to appraise the influence of this fact on the long-run development of output, it is first of all necessary to realize that this rigidity is essentially a short-run phenomenon. There are no major instances of long-run rigidity of prices. Whichever manufacturing industry or group of manufactured articles of any importance we choose to investigate over a period of time, we practically always find that in the long run prices do not fail to adapt themselves to technological progress—frequently they fall spectacularly in response to it[19]—unless prevented from doing so by monetary events and policies or, in some cases, by autonomous changes in wage rates which of course should be taken into account by appropriate corrections exactly as should changes in quality of products. And our previous analysis shows sufficiently why in the process of capitalist evolution this must be so.

What the business strategy in question really aims at—all, in any case, that it can achieve—is to avoid seasonal, random and cyclical fluctua-

18. Unfortunately, this statement is almost as effective a bar to agreement on policy as the most thoroughgoing denial of any case for government regulation could be. In fact it may embitter discussion. Politicians, public officers, and economists can stand what I may politely term the whole-hog opposition of "economic royalists." Doubts about their competence, such as crowd upon us particularly when we see the legal mind at work, are much more difficult for them to stand.

19. They do not as a rule fall as they would under conditions of perfect competition. But this is true only *ceteris paribus*, and this proviso robs the proposition of all practical importance. I have adverted to this point before and shall return to it below.

tions in prices and to move only in response to the more fundamental changes in the conditions that underlie those fluctuations. Since these more fundamental changes take time in declaring themselves, this involves moving slowly by discrete steps—keeping to a price until new relatively durable contours have emerged into view. In technical language, this strategy aims at moving along a step function that will approximate trends. (. . .)

Perhaps the reader feels some surprise that so little remains of a doctrine of which so much has been made in the last few years. The rigidity of prices has become, with some people, the outstanding defect of the capitalist engine and—almost—the fundamental factor in the explanation of depressions. But there is nothing to wonder at in this. Individuals and groups snatch at anything that will qualify as a discovery lending support to the political tendencies of the hour. The doctrine of price rigidity, with a modicum of truth to its credit, is not the worst case of this kind by a long way.

Another doctrine has crystallized into a slogan, viz., that in the era of big business the maintenance of the value of existing investment—conservation of capital—becomes the chief aim of entrepreneurial activity and bids fair to put a stop to all cost-reducing improvement. Hence the capitalist order becomes incompatible with progress.

Progress entails, as we have seen, destruction of capital values in the strata with which the new commodity or method of production competes.

Progress entails, as we have seen, destruction of capital values in the strata with which the new commodity or method of production competes. In perfect competition the old investments must be adapted at a sacrifice or abandoned; but when there is no perfect competition and when each industrial field is controlled by a few big concerns, these can in various ways fight the threatening attack on their capital structure and try to avoid losses on their capital accounts; that is to say, they can and will fight progress itself.

So far as this doctrine merely formulates a particular aspect of restrictive business strategy, there is no need to add anything to the argument already sketched in this chapter. Both as to the limits of that strategy and as to its functions in the process of creative destruction, we should only be repeating what has been said before. This becomes still more obvious if we observe that conserving capital values is the same thing as conserving profits. Modern theory tends in fact to use the concept Present Net Value of Assets (= capital values) in place of the concept of Profits. Both asset values and profits are of course not being simply conserved but maximized. (. . .)

I have entitled this chapter as I did [Monopolistic Practices] because most of it deals with the facts and problems that common parlance associates with monopoly or monopolistic practice. So far I have as much as possible refrained from using those terms in order to reserve for a separate section some comments on a few topics specifically connected with them. Nothing will be said however that we have not already met in one form or another.

(a) To begin with, there is the term itself. Monopolist means Single Seller. Literally therefore anyone is a monopolist who sells anything that is not in every respect, wrapping and location and service included, exactly like what other people sell: every grocer, or every haberdasher, or every seller of "Good Humors" on a road that is not simply lined with sellers of the same brand of ice cream. This however is not what we mean when talking about monopolists. We mean only those single sellers whose markets are not open to the intrusion of would-be producers of the same commodity and of actual producers of similar ones or, speaking slightly more technically, only those single sellers who face a given demand schedule that is severely independent of their own action as well as of any reactions to their action by other concerns. The traditional Cournot-Marshall theory of monopoly as extended and amended by later authors holds only if we define it in this way and there is, so it seems, no point in calling anything a monopoly to which that theory does not apply.

But if accordingly we do define it like this, then it becomes evident immediately that pure cases of long-run monopoly must be of the rarest occurrence and that even tolerable approximations to the requirements of the concept must be still rarer than are cases of perfect competition. The power to exploit at pleasure a given pattern of demand—or one that changes independently of the monopolist's action and of the reactions it provokes—can under the conditions of intact capitalism hardly persist for a period long enough to matter for the analysis of total output, unless buttressed by public authority, for instance, in the case of fiscal monopolies. A modern business concern not *so* protected—i.e., even if protected by import duties or import prohibitions—and yet wielding that power (except temporarily) is not easy to find or even to imagine. Even railroads and power and light concerns had first to create the demand for their services and, when they had done so, to defend their market against competition. Outside the field of public utilities, the position of a single seller can in general be conquered—and retained for decades—only on the condition that he does not behave like a monopolist. Short-run monopoly will be touched upon presently.

Why then all this talk about monopoly? The answer is not without interest for the student of the psychology of political discussion. Of course, the concept of monopoly is being loosely used just like any other. (. . .) But this is not all. Economists, government agents, journalists and politicians in this country obviously love the word because it has come to be a term of opprobrium which is sure to rouse the public's hostility against any interest so labeled. In the Anglo-American world monopoly has been cursed and associated with functionless exploitation ever since, in the sixteenth and seventeenth centuries, it was English administrative practice to create monopoly positions in large numbers which, on the one hand, answered fairly well to the theoretical pattern of monopolist behavior and, on the other hand, fully justified the wave of indignation that impressed even the great Elizabeth.

Nothing is so retentive as a nation's memory. Our time offers other and more important instances of a nation's reaction to what happened centuries ago. That practice made the English-speaking public so monopoly-conscious that it acquired a habit of attributing to that sinister power practically everything it disliked about business. To the typical liberal bourgeois in particular, monopoly became the father of almost all abuses—in fact, it became his pet bogey. Adam Smith,[20] thinking primarily of monopolies of the Tudor and Stuart type, frowned on them in awful dignity. (. . .) And in this country monopoly is being made practically synonymous with any large-scale business.

The theory of simple and discriminating monopoly teaches that, excepting a limiting case, monopoly price is higher and monopoly output smaller than competitive price and competitive output. This is true provided that the method and organization of production—and everything else—are exactly the same in both cases. Actually however there are superior methods available to the monopolist which either are not available at all to a crowd of competitors or are not available to them so readily: for there are advantages which, though not strictly unattainable on the competitive level of enterprise, are as a matter of fact secured only on the monopoly level, for instance, because monopolization may increase the sphere of influence of the better, and decrease the sphere of influence of the inferior, brains,[21] or because the monopoly enjoys a disproportionately higher financial standing. Whenever this is so, then

20. There was more excuse for that uncritical attitude in the case of Adam Smith and the classics in general than there is in the case of their successors because big business in our sense had not then emerged. But even so they went too far. In part this was due to the fact that they had no satisfactory theory of monopoly which induced them not only to apply the term rather promiscuously (Adam Smith and even Senior interpreted for instance the rent of land as a monopoly gain) but also to look upon the monopolists' power of exploitation as practically unlimited which is of course wrong even for the most extreme cases.

21. The reader should observe that while, as a broad rule, that particular type of superiority is simply indisputable, the inferior brains, especially if their owners are entirely eliminated, are not likely to admit it and that the public's and the recording economist's hearts go out to them and not to the others. This may have something to do with a tendency to discount the cost or quality advantages of quasi-monopolist combination that is at present as pronounced as was the exaggeration of them in the typical prospectus or announcement of sponsors of such combinations.

that proposition is no longer true. In other words, this element of the case for competition may fail completely because monopoly prices are not necessarily higher or monopoly outputs smaller than competitive prices and outputs would be at the levels of productive and organization efficiency that are within the reach of the type of firm compatible with the competitive hypothesis.

There cannot be any reasonable doubt that under the conditions of our epoch such superiority is as a matter of fact the outstanding feature of the typical large-scale unit of control, though mere size is neither necessary nor sufficient for it. These units not only arise in the process of creative destruction and function in a way entirely different from the static schema, but in many cases of decisive importance they provide the necessary form for the achievement. They largely create what they exploit. Hence the usual conclusion about their influence on long-run output would be invalid even if they were genuine monopolies in the technical sense of the term. (. . .)

In the short run, genuine monopoly positions or positions approximating monopoly are much more frequent. The grocer in a village on the Ohio may be a true monopolist for hours or even days during an inundation. Every successful corner may spell monopoly for the moment. A firm specializing in paper labels for beer bottles may be so circumstanced—potential competitors realizing that what seem to be good profits would be immediately destroyed by their entering the field—that it can move at pleasure on a moderate but still finite stretch of the demand curve, at least until the metal label smashes that demand curve to pieces.

New methods of production or new commodities, especially the latter, do not *per se* confer monopoly, even if used or produced by a single firm. The product of the new method has to compete with the products of the old ones and the new commodity has to be introduced, i.e. its demand schedule has to be built up. As a rule neither patents nor monopolistic practices avail against that. But they may in cases of spectacular superiority of the new device, particularly if it can be leased like shoe machinery; or in cases of new commodities, the permanent demand schedule for which has been established before the patent has expired.

Thus it is true that there is or may be an element of genuine monopoly gain in those entrepreneurial profits which are the prizes offered by capitalist society to the successful innovator. But the quantitative importance of that element, its volatile nature and its function in the process in which it emerges put it in a class by itself. The main value to a concern of a single seller position that is secured by patent or monopolistic strategy does not consist so much in the opportunity to behave temporarily according to the monopolist schema, as in the protection it affords against temporary disorganization of the market and the space it secures for long-range planning. Here however the argument merges into the analysis submitted before.

Glancing back we realize that most of the facts and arguments touched upon in this chapter tend to dim the halo that once surrounded perfect competition as much as they suggest a more favorable view of its alternative. (. . .)

If we try to visualize how perfect competition works or would work in the process of creative destruction, we arrive at a still more discouraging result. This will not surprise us, considering that all the essential facts of that process are absent from the general schema of economic life that yields the traditional propositions about perfect competition. At the risk of repetition I will illustrate the point once more.

Perfect competition implies free entry into every industry. It is quite true, within that general theory, that free entry into all industries is a condition for optimal allocation of resources and hence for maximizing output. If our economic world consisted of a number of established industries producing familiar commodities by established and substantially invariant methods and if nothing happened except that additional men and additional savings combine in order to set up new firms of the existing type, then impediments to their entry into any industry they wish to enter would spell loss to the community. But perfectly free entry into a *new* field may make it impossible to enter it at all. The introduction of new methods of production and new commodities is hardly conceivable with perfect—and perfectly prompt—competition from the start. And this means that the bulk of what we call economic progress is incompatible with it. As a matter of fact, perfect competition is and always has been temporarily suspended whenever anything new is being introduced. (. . .)

The firm of the type that is compatible with perfect competition is in many cases inferior in internal, especially technological, efficiency. If it is, then it wastes opportunities. It may also in its endeavors to improve its methods of production waste capital because it is in a less favorable position to evolve and to judge new possibilities. And, as we have seen before, a perfectly competitive industry is much more apt to be routed—and to scatter the bacilli of depression—under the impact of progress or of external disturbance than is big business. In the last resort, American agriculture, English coal mining, the English textile industry are costing consumers much more and are affecting *total* output much more injuriously than they would if controlled, each of them, by a dozen good brains. (. . .)

In this respect, perfect competition is not only impossible but inferior, and has no title to being set up as a model of ideal efficiency.

In this respect, perfect competition is not only impossible but inferior, and has no title to being set up as a model of ideal efficiency. It is hence a mistake to base the theory of government regulation of industry on the principle that big business should be made to work as the respective industry would work in perfect competition."[22]

III. Passages from Edward S. Mason's Critique of 1951

Mason (1899-1992) was a good friend and a member of what Schumpeter called his "inner circle" of younger colleagues. Along with two other Harvard economists (E.H. Chamberlin and Joe S. Bain), Mason was one of the pioneers of industrial organization theory. He and Bain led the development of the structure-conduct-performance paradigm that dominated the sub-field of industrial organization from about the late 1940s to the 1980s, when it began to yield to game theory and other approaches.

The following excerpts comprise about 20 percent of Mason's article in *The Review of Economics and Statistics*.[23] This issue was dedicated to Schumpeter and its contents were devoted entirely to his work (he had died in 1950). In addition to Mason, the 14 contributors comprised something of an all-star lineup of the profession at that time: Paul Samuelson, Alvin Hansen, Jan Tinbergen, Gottfried Haberler, Fritz Machlup, Seymour Harris, E.H. Chamberlin, Erich Schneider, Arthur Marget, David McCord Wright, Wolfgang Stolper, Arthur Smithies, and A.P. Usher. Six of these economists had been Ph.D. students of Schumpeter's.

Overall, the authors were appropriately generous, but most pulled no punches in their evaluations. Tinbergen, for example, argued (correctly) that Schumpeter was not really a mathematical economist. Chamberlin argued (incorrectly) that Schumpeter had misunderstood his work on monopolistic competition. Mason, as is evident in the following passages, credits Schumpeter with real insight but contends that he provides no practical guide to antitrust policy. This is a fair assessment within the limits of the structure-conduct-performance framework in which Mason was writing, although part of the last sentence in his first paragraph (on the necessity of market power for innovation) is a gross distortion of Schumpeter's thinking. On the whole, Mason's comments go far in explaining why, for so long, Schumpeter's analysis had so little impact on competition theory (footnotes are quoted as cited):

22. CAPITALISM, *supra* note 1, *at* 78-106.

23. Edward Mason, *Schumpeter on Monopoly and the Large Firm*, 33 REV. ECON. & STAT. 139-144 (1951)

"These chapters [VII and VIII of *Capitalism, Socialism and Democracy*] which bring together and sharpen earlier views on the role of the large firm in the competitive process, represent one of the most effective as well as most drastic critiques extant concerning traditional patterns of anti-trust thought. The critique is drastic and effective because it plausibly undermines the two main pillars of the traditional ideology: first that market power is the proper object of attack since power means the ability to exploit; and, second, that the preservation of competition, meaning the exclusion of position of market power, will assure the efficient use of resources. The essence of Schumpeter's position is that market power is necessary to innovation and that innovation is the core of effective competition. (. . .)

Schumpeter maintains that his argument is not a case against all anti-monopoly policy but only a particular variety of policy. There may be "cases of restrictive or regulating strategy" that have "that injurious effect on the long-run development of output which is uncritically attributed to all of them."[24] He does not, however, give us much help in determining what business practices or strategies might be expected to produce expansive rather than restrictive results. What he has to say in criticism of existing policy constitutes a challenge that every serious student of the "monopoly problem" must take to heart. But whether his view of competition as the process of "creative destruction" could be made to yield principles applicable by government agencies and the courts in pursuit of a "rational" as opposed to a "vindictive" anti-monopoly policy is a different matter.

Schumpeter maintains that his argument is not a case against all anti-monopoly policy but only a particular variety of policy.

American anti-trust policy, as distinguished from the anti-monopoly policy of most other countries, purports to be—and to some extent is—an attack upon positions of market power. Whereas legislation and administrative practice elsewhere has emphasized *abuse* of power, including the charging of unreasonable prices, as the proper object of attack, and has recognized the possibility of "good" monopolies, American practice, within certain areas at least, has attacked market power as such. "The reasonable prices fixed today—may become the unreasonable prices of tomorrow" runs the language of a famous anti-trust decision.[25] And with respect to certain kinds of agreements in restraint of trade, i.e., certain attempts to secure a position of market power, the judicial position has been that they are unreasonable and illegal *per se*.

24. Capitalism, *supra* note 1, *at* 91.

25. U.S. vs. Trenton Potteries Co. et al. 273 U.S. 392

Needless to say, however, U.S. anti-trust policy has not been entirely consistent. Large firms enjoying a position of market power have remained immune, while associations with much less power have been broken up. (. . .)

Schumpeter is on surer—and also more important—ground in his evaluation of the results of innovation, that is to say, the relation of innovation to effective competition. Here he denies completely the significance for public policy purposes of any standard of evaluation derived from pure competition, marginal cost-price relationships, or other formulations of static economic analysis. His general position is best stated in a proposition quoted with approval by Pigou.

> A system—any system, economic or other—that at every point of time fully utilizes its possibilities to the best advantage may yet in the long run be inferior to a system that does so at *no* given time, because the latter's failure to do so may be a condition for the level or speed of long-run performance.[26] (. . .)

During the nineteenth century innovation, according to Schumpeter, was typically the product of new firms. "The new processes do not, and generally cannot, evolve out of the old firms, but place themselves side by side with them and attack them." [27] In the twentieth century epoch of "trustified" capitalism, however, innovations issue from existing firms and, as indicated above, usually from large ones. Furthermore, although the creation of giant firms represents a high form of innovating ability that could not be expected to be brought to fruition except in a capitalism that gives full scope to exceptional talent, the process of concentration ends up by making innovations quasi-automatic.

> It meets with much less friction, as failure in any particular case loses its dangers, and tends to be carried out as a matter of course on the advice of specialists. Progress becomes "automized," increasingly impersonal and decreasingly a matter of leadership and individual initiative.[28] (. . .)

Particularly serious difficulties are presented when the attempt is made to apply Schumpeter's analysis in the field of public policy. Here the problems presented are what to do about a specific agreement in

26. The quotation is from CAPITALISM, *supra* note 1, *at* 83. It is cited in A.C. Pigou, LAPSES FROM FULL EMPLOYMENT 71 (1945).

27. "*The Instability of Capitalism*," 28 ECON. J. 384 (1928).

28. *Ibid*, Cf. also *Der Unternehmer in der Volkswirtschaft von heute*, in STRUKTUR WANDLUNGEN DER DEUTSCHEN VOLKSWIRTSCHAFT, I (1928), p. 303, where these ideas are worked out in greater detail.

restraint of trade, a particular combination of hitherto independent firms, or a concrete set of business practices. If one took at face value his admonition that, since we are dealing with an organic process that takes time, a judgment on the consequences of any particular part of it—say a combination of hitherto independent firms—can only be an historical judgment, as these consequences "unfold over decades," and a partial judgment, since the repercussions reverberate throughout an economy which is in process of "organic development," informed public action would clearly be impossible. However, Schumpeter assures us that what he is opposed to is not every anti-monopoly policy but only certain kinds of monopoly policy.[29]

PARTICULARLY SERIOUS DIFFICULTIES ARE PRESENTED WHEN THE ATTEMPT IS MADE TO APPLY SCHUMPETER'S ANALYSIS IN THE FIELD OF PUBLIC POLICY.

What a "sensible" as opposed to a "vindictive" anti-monopoly policy would presumably emphasize are mainly the possibility that various restrictive activities may be a necessary concomitant to innovation with its accompanying investment decisions, and that a firm producing new products and processes may be a more effective stimulant to efficient behavior on the part of others than a large number of routine competitors. What this appears to boil down to in terms of practical application is a useful admonition that the existence of a large firm or a few large firms in a market is not necessarily incompatible with effective competition. (. . .)

Schumpeter most certainly exaggerated the extent of the influence exerted on American business organization and business practices by anti-trust policy. Furthermore, he painted a picture of anti-trust objectives and of the ideological justification of these objectives that is in many respects distorted and out of focus. Nevertheless, his powerful attack on the limitations of static economic analysis as an intellectual foundation for a public anti-monopoly policy is highly salutary and profoundly correct. And his discussion of the political environment in which public policy toward business organization and business practices actually gets shaped is a useful corrective to the thinking of those colleagues who conceive that policy can be divorced from politics. Finally, although it is difficult to the point of impossibility to derive from Schumpeter's "process of creative destruction" an analytical framework on which applicable and effective anti-trust standards might be built, his analysis suggests lines of research and invokes considerations that must play a role in formulating an acceptable public policy in this area."

29. CAPITALISM, *supra* note 1, *at* 134,

IV. Current Use of Schumpeter's Theories

For three decades after the appearance of Mason's article in 1951, relatively few economists read or cited Schumpeter. But then the Schumpeterian revival began in earnest. In a retrospective analysis of *Capitalism, Socialism and Democracy*, published in 1981, another all-star lineup once again paid tribute to his work. His great student Paul Samuelson wrote that "a century after Schumpeter's birth, we take his writings seriously and treat them as living contributions to contemporary debate."[30] In 1983, the centennial of the birth of both Keynes and Schumpeter, *Forbes*[31] ran a cover story, written by Peter Drucker, arguing that it was Schumpeter, not Keynes, who would provide the better guide to the economic changes that were beginning to engulf the world. In 1984, the German economist Herbert Giersh suggested in the *American Economic Review*[32] that the Age of Keynes was about to yield to the Age of Schumpeter. In the 1991 edition of his best-selling *The Worldly Philosophers*,[33] Robert Heilbroner devoted an entire chapter to Schumpeter, and concluded that more than any other great economist depicted in his book, "Schumpeter speaks to us with a voice that is unmistakably contemporary." Time has proved all of these judgments correct.

Before we get too enthusiastic about Schumpeter's work as a beacon of public policy, however, we should keep in mind three caveats:

1. Grand social theorists are not always reliable guides in specific cases. Their ideas can easily be distorted, either deliberately or inadvertently, in service to some immediate goal that the theorists themselves would not have supported. Karl Marx, for example, who urged that "workers of the world unite," would never have endorsed the "socialism in one country" doctrine set forth by Nikolai Bukharin and adopted by Joseph Stalin in 1925, let alone the Stalinist terrors that became institutionalized in 1927. The same point holds true of great economists who did not aspire to grand social theory. Many ostensibly "Keynesian" public policies—especially in the U.K. and the U.S. between about 1950 and 1980—would not likely have been approved by Keynes had he been alive to evaluate them.

2. In the case of Joseph Schumpeter, he addressed so many topics over so long a period (his first work appeared in 1905, his last posthumously, in 1954), that he frequently adjusted his thinking. He wrote so volu-

30. Paul Samuelson, *Schumpeter's Capitalism, Socialism and Democracy in* SCHUMPETER'S VISION: CAPITALISM, SOCIALISM, AND DEMOCRACY AFTER 40 YEARS 21 (Arnold Heertje, ed., 1981)

31. Peter Drucker, *Schumpeter and Keynes*, FORBES 124 (May 23, 1983)

32. Herbert Giersh, *The Age of Schumpeter,* 74 AM. ECON. REV. 103-109 (1984)

33. ROBERT HEILBRONER, THE WORLDLY PHILOSOPHER: THE LIVES, TIMES, AND IDEAS OF THE GREAT ECONOMIC THINKERS, 6th ed, 291 (1991)

minously during this half-century that it is not hard to find apparently contradictory statements in his work, most of which reflect altered external conditions. This characteristic is so pronounced in Schumpeter's writings that it calls to mind the famous lines from the poet Walt Whitman's *Song of Myself* (1851):

> "Do I contradict myself?
> Very well, I contradict myself.
> (I am large, I contain multitudes)."

3. Beyond arguing against mindless trust busting and the conflation of big business with monopoly, Schumpeter very seldom addressed antitrust concerns directly. His central interests had much less to do with industrial organization per se than with entrepreneurship, innovation, business cycles, and the history of economic analysis. The courses he taught at Harvard were mostly on economic theory and on the history of economics as a discipline.

During the 1960s, before Schumpeter's work was taken up for purposes of antitrust analysis, a substantial related literature began to develop around what, unfortunately, became known as "the Schumpeter hypothesis." This alleged hypothesis held that large firms were better at innovation than small firms. Numerous articles appeared—many from prominent scholars—either supporting or attacking the hypothesis. But, as Anne Mayhew correctly pointed out in 1980, Schumpeter had never even formulated such a hypothesis.[34]

It is true, as is evident in the quoted excerpts from *Capitalism, Socialism and Democracy*, that Schumpeter thought that certain kinds of innovation required

34. About 20 useful articles have appeared on the misnamed "Schumpeter hypothesis." Some of the most useful are Franklin Fisher and Peter Temin, *Returns to Scale in Research and Development: What Does the Schumpeterian Hypothesis Imply?* 81 J. POL. ECON.56-70 (1973); F.M Scherer, *Schumpeter and Plausible Capitalism*, 30 J. ECON. LITERATURE 1416-1433 (1992); and Tom Nicholas, *Why Schumpeter was Right: Innovation, Market Power, and Creative Destruction in 1920s America,* 63 J. ECON. HIST. 1023-1058 (2003). Part of this debate is analyzed in DAVID REISMAN, SCHUMPETER'S MARKET: ENTERPRISE AND EVOLUTION Ch. 5 (2004). A particularly good example of the frequent misreadings of the "Schumpeter hypothesis" is J.B. Rosenberg, *Research and Market Share: A Reappraisal of the Schumpeter Hypothesis*, 25 J. IND. ECON. 101-112 (1976): "Schumpeter believed that technological innovations are more likely to be initiated by large rather than small firms" *at* 101. This statement, and many like it from other scholars, is incorrect, but plausible from a selective reading of Schumpeter's sometimes contradictory and ambiguous language. See the useful corrective by Anne Mayhew, *Schumpeterian Capitalism versus the "Schumpeterian Thesis"*, 14 J. ECON. ISSUES 583-592 (1980). Mayhew points out that most of the support for the existence of the "Schumpeterian thesis" derives from a single sentence on p. 106 of Capitalism, Socialism and Democracy—a sentence which is often taken out of context and which does not begin to express the complexity of Schumpeter's thinking. That sentence is: "What we have got to accept is that . . . [the large-scale establishment] has come to be the most powerful engine of . . . progress and in particular of the long run expansion of total output."

teams of researchers. But it is equally clear from his writings that he believed innovation could emerge from almost any source: the lone entrepreneur (the New Man founding a New Firm); the medium-sized company; or the giant corporation with its institutionalized R&D labs.

Throughout his career, Schumpeter admired entrepreneurial startups, and he almost surely would have been delighted by phenomena such as the evolution of Silicon Valley, a center of creative destruction if there ever was one. As for whether he would have taken the side of a company such as Microsoft in its major antitrust suits, it's impossible to say. From the totality of his writings, and allowing for certain self-contradictions, it seems likely that he'd have admired Microsoft greatly in its early years, but would then have turned his preferences to some (not all) of its many scrappy challengers.

Throughout his career, Schumpeter admired entrepreneurial startups, and he almost surely would have been delighted by phenomena such as the evolution of Silicon Valley, a center of creative destruction if there ever was one.

It is here that Schumpeter's enthusiasm for history becomes most relevant to his stance toward competition policy. One of the many lessons of history, as the Cambridge historian F.W. Maitland once said, is that "What is now in the past was once in the future." To put it another way, we simply cannot know with much certainty what the long-term consequences of particular antitrust decisions are going to be. Often the losers of the case turn out to be winners over the long haul, and vice-versa. In the landmark cases of Standard Oil and American Tobacco in 1911, for example, the companies lost and were forcibly split up; but both became more efficient over the long run. Conversely, U.S. Steel won its prolonged case in 1920 (in large part because it had stopped competing as fiercely as its constituent company Carnegie Steel had done). But by 1938 it had lost about two-thirds of the market share it had held at the time of its formation in 1901.[35]

In the case against IBM that began in 1969, antitrust pressures forced the company, over time, to alter its monopolistic practices. Had that not occurred, it seems unlikely that innovation in information technology would have grown at the blinding speed we now take for granted. The same is true of the 1984 breakup of AT&T under antitrust pressures. At the time of that breakup, many economists believed it to be a tragic mistake—some because it endangered (and ultimately killed) Bell Labs, one of the nation's finest centers of R&D. Yet we now

35. F. M. Scherer has often pointed out this pattern of unexpected consequences from wins and losses in big antitrust cases, including most of those mentioned here. On U.S. Steel, see also Thomas K. McCraw & Forest Reinhardt, *Losing to Win: U.S. Steel's Pricing, Investment Decisions, and Market Share, 1901-1938*, 49 J. Econ. Hist. 593-619 (1989).

know that for IBM, its competitors, and AT&T's successor firms, the long-term consequences of antitrust pressures unleashed immense entrepreneurial energy that otherwise might have remained dormant. That energy produced exactly the types of innovations that we most identify with Joseph Schumpeter.

A similar historical uncertainty emerges when we apply the "what is now in the past was once in the future" test to the related subject of deregulation. During the three decades since that movement began in the 1970s, the unanticipated consequences have been almost as numerous as the intended ones. In the case of airlines, the results have been painful but mostly positive; for railroads and trucking, clearly positive; for telecommunications, very positive; for electric utilities, mixed but on balance likely negative; for financial institutions, numerous innovations (complex derivatives, structured investment vehicles, credit default swaps), but some of them potentially catastrophic for the national economy.

These judgments themselves, of course, must be tentative and premature. Only in the long term can we be more certain. And Schumpeter almost always thought in the long term. This characteristic could hardly be more conspicuous than in the quoted passages from *Capitalism, Socialism and Democracy*, in which he writes of the "meaningless job" of drawing economic conclusions "*ex visu* of a point of time," about "a situation that is sure to change presently." Judges and juries must inevitably draw economic conclusions in antitrust cases, but it is not what Schumpeter chose to do. He almost never expressed an opinion of how pending legislation should be decided, and it is very hard to imagine his taking part in any case as a consultant or expert witness.

"Two kinds of people I distrust: architects who profess to build cheaply, and economists who profess to give simple answers."

Schumpeter had been trained at the University of Vienna as a lawyer as well as an economist, but he had left the practice of law in 1908—a step that tells us a great deal about his preferred way of thinking. In the area of competition policy, his main fear during the 1930s and 1940s was of what he called "*indiscriminate* trust busting." No such eventuality came to pass, as we now know, despite some unwise Supreme Court decisions during the 1940s. From the vantage point of our own time, indiscriminate trust busting seems the precise opposite of what has occurred since the 1940s.[36]

In 1943, a year after the appearance of *Capitalism, Socialism and Democracy*, Schumpeter wrote in his diary, "Two kinds of people I distrust: architects who profess to build cheaply, and economists who profess to give simple answers." So

36. Schumpeter used this phrase not only in CAPITALISM but also in his presidential address to the American Economic Association in December 1948. See his discussion of the monopoly question in Joseph Schumpeter, *Science and Ideology*, 29 AM. ECON. REV. 347-349 (1949). Italics in original.

it would be quite an irony if his name became attached to a particular approach to antitrust. Economists and others are free to invoke his name in specific cases, of course, but in doing so they should tread carefully—very carefully.[37]

37. One of the best articles on this question is Michael Katz & Howard Shelanski, "*Schumpeterian" Competition and Antitrust Policy in High-Tech Markets*, 14 COMPETITION 47 ff. (2005). Less impressive is Michael Carrier, *Two Puzzles Resolved: Of the Schumpeter-Arrow Stalemate and Pharmaceutical Innovation Markets*, 93 IOWA L. REV. 393 (February 2008), which, like many of the articles cited in note 34 above, makes the mistake of identifying Schumpeter's position as favoring industry concentration as a means to innovation.