

AN ECONOMIC APPROACH TO PRICE FIXING

LOUIS KAPLOW*

I. INTRODUCTION

The rule against price fixing is the least controversial prohibition in competition law. There is, however, less consensus than meets the eye on what constitutes price fixing and on how legal regimes should identify and remedy it. Moreover, prevailing views and existing doctrine are not grounded in systematic economic analysis of the problem.

This article aims to provide an analytical foundation for designing policy. It begins by stating the nature of the challenge posed by oligopoly pricing and articulating a welfare-based approach that specifies the benefits and costs of attempts to control it.¹ With this framework in mind, the next step is to consider how coordinated price elevation by oligopolistic firms might best be

* Finn M.W. Caspersen and Household International Professor of Law and Economics, Harvard Law School; Research Associate, National Bureau of Economic Research. I am grateful to Jonathan Baker, Aaron Edlin, Oliver Hart, Carl Shapiro, Steven Shavell, Kathryn Spier, the editors and referees, and participants at a Harvard Law School workshop and at Northwestern University's Searle Center Research Symposium on Antitrust Economics and Competition Policy for helpful comments; Shobitha Bhat, Michael Doore, Ariel Fox, Joseph Gay, Jason Green-Lowe, James Ianelli, Daniel Jacobsen, Noam Lerer, Lauren Merrell, William Milliken, Balaji Narain, Cody Rockey, Krysten Rosen, and Michael Sabin for research assistance; and the Harvard Law School's John M. Olin Center for Law, Economics, and Business and the Kauffman Foundation for financial support. This article is the second part of a larger project, the first of which—Louis Kaplow, *On the Meaning of Horizontal Agreements in Competition Law*, 99 CALIF. L. REV. (forthcoming 2011)—offers a critical assessment of the law on agreement (concerted action), and the third of which—Louis Kaplow, *Direct Versus Communications-Based Prohibitions on Price Fixing*, J. LEGAL ANALYSIS (forthcoming)—compares the sort of regime for price fixing developed here to the conventionally advocated approach. Seeds for some of the ideas in this article appear but were not developed in PHILLIP AREEDA & LOUIS KAPLOW, ANTI-TRUST ANALYSIS (4th ed. 1988), and Louis Kaplow & Carl Shapiro, *Antitrust*, in 2 HANDBOOK OF LAW AND ECONOMICS 1073 (A. Mitchell Polinsky & Steven Shavell eds., 2007). I have benefited greatly from prior interchanges with my co-authors, but neither they nor others who have offered comments should be held responsible for any views expressed herein (some of which diverge from what I have suggested previously).

¹ Discussion is confined to coordination on price. The broad analytical framework developed here is applicable to nonprice coordination, but there are also important differences. For example, many forms of nonprice coordination are more difficult, and detection of some types of

identified, with attention to the error costs associated with different means of proof. Various forms of sanctions are also explored, emphasizing the deterrent role of remedies.

These subjects have received limited attention in the literature. Instead, commentary focuses on penalizing certain sorts of interfirm communications that implement or facilitate oligopolistic price elevation.² This emphasis—in lieu of an economic assessment that begins with first principles—is attributable in significant part to a view of the law’s agreement requirement. This subject is examined in detail in a companion article,³ whereas the present inquiry is devoted exclusively to the question of what regime maximizes welfare. For present purposes, suffice it to say that the language of statutes and treaties, prior cases, and de facto practices of courts are substantially more amenable to a broad interpretation of horizontal agreements than is generally recognized.⁴ Interestingly, Donald Turner’s seminal article, which argues that the law’s prohibition should be circumscribed, defends his stance as a matter of policy, not legal doctrine—believing that it is more plausible to read Sherman Act Section 1’s language as encompassing loose, implicit understandings: “there are far better grounds for saying that though there may be ‘agreement’ it is not unlawful agreement.”⁵

coordinated behavior (e.g., firms abstaining from entering others’ markets) using market-based evidence may well be more challenging.

² See generally 6 PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* (2d ed. 2003).

³ See Louis Kaplow, *On the Meaning of Horizontal Agreements in Competition Law*, 99 CALIF. L. REV. 683 (2011).

⁴ To elaborate: Most state the rule instead as requiring some sort of express or explicit agreement or concert of action, although it seems that a selective communications prohibition is a reasonably close concrete approximation to what is meant. Second and related, those stating the law offer neither a reasonably clear operational statement of the rule nor a series of crisp examples from which the contours may readily be inferred. They differ on whether tacit agreements are included, but do not in any case define the concept, and in other respects they employ key terms that are ambiguous or even susceptible to opposite interpretations. If, instead, one considers the actual language of typical prohibitions, often involving terms like conspiracy or concerted action, one is led to a broad prohibition of interdependent oligopolistic behavior because these terms focus our attention on the existence of a meeting of the minds, not on the use of particular communications that might be employed to achieve such a mutual understanding. Indeed, the statutes, the general law of conspiracy, and numerous authoritative, earlier Supreme Court decisions support this view. Moreover, the broader interpretation that encompasses interdependent oligopoly behavior is more consistent with the actual practices of lower courts regarding their use of so-called plus factors to find a conspiracy, their instructions to juries on liability, and their method of determining damages. In addition, important elements of litigants’ behavior as well as practitioners’ advice and firms’ actions suggest that many hold a broad de facto view of the scope of liability.

⁵ Donald F. Turner, *The Definition of Agreement Under the Sherman Act: Conscious Parallelism and Refusals to Deal*, 75 HARV. L. REV. 655, 671 (1962). Earlier, he states that “It is not novel conspiracy doctrine to say that agreement can be signified by action as well as by words.” *Id.* at 665. Turner concludes on the matter as follows:

Even recognizing prevailing views of legal doctrine, it is striking that most of the questions investigated here have been so neglected. As stated, the present approach involves the direct application of economic understandings to the problem at hand. In recent decades, courts and commentators have increasingly embraced the view that competition law should be grounded in economic substance rather than formalistic distinctions. The U.S. Supreme Court's decisions in *Sylvania*, *Matsushita*, and *Leegin* come to mind,⁶ and other jurisdictions have also devoted greater attention to the economic underpinnings of competition law. Moreover, Richard Posner's 1969 article and book famously endorsed a more economically based strategy for addressing oligopoly pricing,⁷ yet subsequent commentary, even when citing his writings, does not significantly engage their content.⁸

The present analysis can usefully be understood as starting on a clean slate, given our current state of economic knowledge. From this perspective, it asks what would be the best choice of policy toward coordinated oligopoly pricing.⁹ In doing so, it goes well beyond Posner's analysis and differs in many respects, which can hardly be surprising in light of the decades of theoretical and empirical literature in industrial organization in the interim. Nevertheless, the spirit of his work as well as some of his key ideas provide important illumination.

Part II begins by examining core economic principles relevant to coordinated oligopoly pricing and its regulation. From the outset, it is notable that

I also find considerable appeal, as a general matter, in defining "agreement" for purposes of Sherman Act law in terms of interdependence of decisions, if for no other reason than that it seems to me to be a clearer and more workable standard than any other standard, of acceptable scope, which requires something more. Once one goes beyond the boundaries of explicit, verbally communicated assent to a common course of action—a step long since taken and from which it would not seem reasonable to retreat—it is extraordinarily difficult if not impossible to define clearly a plausible limit short of interdependence.

Id. at 683. It is ironic, therefore, that the U.S. Supreme Court in *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 554 (2007), cites *Turner*, *supra*, at 672, for the proposition: "mere interdependence of basic price decisions is not conspiracy."

⁶ *Continental T.V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36, 58–59 (1977); *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574, 587, 596 (1986); *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 551 U.S. 877, 887–89 (2007).

⁷ See RICHARD A. POSNER, *ANTITRUST LAW*, ch. 3 (2d ed. 2001) [hereinafter *ANTITRUST*] (the first edition was published in 1976 with the subtitle "An Economic Perspective"); Richard A. Posner, *Oligopoly and the Antitrust Laws: A Suggested Approach*, 21 *STAN. L. REV.* 1562 (1969) [hereinafter *Oligopoly*].

⁸ One of the few exceptions is a review of the first edition of Posner's book, which devotes a handful of pages to the subject at hand. See F.M. Scherer, *The Posnerian Harvest: Separating Wheat from Chaff*, 86 *YALE L.J.* 974, 981–84 (1977). Another (Richard Markovits) is discussed in note 232.

⁹ A sequel compares the approach developed here to the conventional one. See Louis Kaplow, *Direct Versus Communications-Based Prohibitions on Price Fixing*, *J. LEGAL ANALYSIS* (forthcoming).

none of the pertinent theory directly distinguishes between successful oligopolistic price elevation due merely to recognized interdependence (firms refrain from price cutting because of an expectation of retaliation derived from a shared appreciation of their circumstances) and that resulting from classic cartel behavior (firms meet secretly in hotel rooms to discuss prices and the consequences of cheating), or various cases in between. The theory of coordinated oligopoly behavior, based on the modern theory of repeated games, emphasizes whether firms have incentives to cheat by undercutting elevated oligopoly prices and how uncertainty may produce price wars even when no firm has defected. The harm from oligopoly pricing in terms of allocative inefficiency or loss in consumer welfare depends most directly on the extent and duration of supracompetitive pricing, not on the means of reaching or maintaining the heightened price. From a dynamic perspective, price elevation may also cause production inefficiency on account of excessive entry. There are some settings in which additional entry could be efficient due to insufficient product variety or other difficulties in recovering fixed costs; such benefits might sometimes justify some price elevation, but—crucially for present purposes—do not directly distinguish the means by which it may be accomplished.

An economic approach to regulating oligopolistic price elevation through legal liability—like the economic approach to law enforcement more generally—seeks to determine liability and apply sanctions based primarily on the deterrence benefits that result as well as any chilling of desirable behavior that may arise, in addition to the costs of operating the regime.¹⁰ In assessing the cost of false positives, attention focuses on incidental negative behavioral effects, not on mistakes that are defined by reference to proxy legal standards and then given arbitrary weight. An example that will prove important involves sanctioning firms that actually charge elevated oligopoly prices, the prospect of which chills (deters) such behavior. This outcome is favorable in terms of social welfare but under some advanced legal standards would be deemed to be an undesirable error in cases in which the firms did not employ forbidden modes of communication. The optimal legal policy also depends heavily on empirical matters, such as the extent of oligopoly pricing in the economy and the potential success of various means of detecting it; some pertinent evidence will be presented, but important gaps in knowledge remain.

¹⁰ See, e.g., A. Mitchell Polinsky & Steven Shavell, *The Theory of Public Enforcement of Law*, in 1 HANDBOOK OF LAW AND ECONOMICS 403 (A. Mitchell Polinsky & Steven Shavell eds., 2007). For analysis that focuses particularly on optimizing the tradeoff of false positives and false negatives, see Louis Kaplow, *Burden of Proof*, 121 YALE L.J. (forthcoming) [hereinafter *Burden of Proof*], and Louis Kaplow, *On the Optimal Burden of Proof* (2011) (unpublished manuscript) [hereinafter *Optimal Proof Burden*].

Part III explores the problem of detection, the greatest challenge in the regulation of interdependent oligopoly pricing. Firms naturally seek to hide possibly illegal aspects of their behavior, and reliable indicators are not always readily obtainable to enforcers. One approach is to employ market-based evidence to infer successful oligopolistic coordination. Price elevation may be inferred from pricing changes over time, such as the observation of significant industry price increases not accompanied by corresponding changes in cost or of sharp price drops associated with price wars. Alternatively, markups might be determined from measures of price and marginal cost or inferred from the elasticity of firms' demand curves. Note that, as with most of the analysis in Part II, these inquiries do not depend on whether or not detected price elevation originated through classic cartel behavior (such as secret meetings in hotel rooms).

Also relevant to detection is the conduciveness of conditions to coordinated oligopoly pricing. Highly conducive conditions make inferences of successful interdependent pricing more credible whereas unconducive conditions cast doubt on its plausibility. However, due to the noisy empirical relationship between industry structure and performance as well as the possibility that conditions are highly conducive yet oligopolistic pricing is effectively deterred, conducive conditions do not in themselves strongly indicate coordinated price elevation—whereas highly unconducive conditions do significantly negate the inference. This particular feature of sound detection strategy importantly differs from what many associate with a more traditional focus on the existence of particular interfirm communications because, under certain assumptions, more conducive conditions reduce the likelihood that such activity occurred even though they increase the magnitude of expected social harm.¹¹

Another route to detection looks for internal evidence of whether or not coordinated oligopolistic behavior is taking place. That is, in addition to attempts to observe behavior directly or to infer it from market activity, one can examine firms' own internal understandings (as reflected in agents' thinking and actions) about what they were doing. Yet another, more familiar form of evidence, often deriving from similar sources, concerns interfirm communications, the existence of which may likewise indicate what firms actually did. These sources of information are complementary to each other and to market-based techniques. In deciding whether to assign liability in a particular case, all such evidence on detection should be considered in light of the decision theoretic approach articulated previously. Some sorts of proof are more reliable than others and give rise to different risks of particular types of errors. For example, some market-based techniques that attempt to determine firms' marginal costs could result in adverse incentive effects in the case of underestima-

¹¹ For extensive treatment of this difference, see Kaplow, *supra* note 9, § 3.4.

tion, making it optimal to find liability only if the measured price elevation is substantial or other confirming evidence is present.

The analysis of sanctions in Part IV focuses primarily on deterrence. In many instances (reflecting current practice), the most important instruments are fines levied by government enforcers and damages collected by injured parties. If the probability of sanctions (determined by the intensity of enforcement effort, methods of detection, and the burden of proof) and their magnitude are sufficient, most coordinated price elevation will be deterred. A major challenge in setting monetary sanctions is determining the extent of price elevation, although this magnitude will often be indicated by much of the evidence on detection considered in Part III. The measurement problem is in principle the same whether price elevation was accomplished through secret meetings, mere recognition of interdependence, or in any other manner. The threat of imprisonment as well as fines assessed against individual actors can be a useful supplement, particularly in light of agency problems within firms. Injunctions are also considered. Although they have tended to be a fixation of much academic commentary, it is not evident that they are important in controlling coordinated oligopoly pricing.

Three additional considerations are the subject of Part V. Institutional issues are examined because the prospect of errors and possibly significant administrative costs are important considerations that in turn may depend on the type of system employed to enforce competition rules. Optimal detection criteria at the investigation stage may differ from those appropriate to adjudication. These in turn may depend on the mode of adjudication, for example, whether by an expert government agency or through private suits before generalist judges and lay juries. Another question concerns the role of liability for attempts. Finally, attention is devoted to the relevance for present purposes of the possibility of price elevation in oligopolistic industries that is due to the unilateral exercise of market power rather than coordination.

This article analyzes the regulation of coordinated oligopoly pricing by considering the nature of the problem, how to detect its presence, and what remedies to apply. In the course of this investigation, the commonly advocated approach—which most commentators also contend describes current law in leading jurisdictions—of attacking only express and perhaps also tacit agreements, variously defined, barely arises. That is, it does not emerge from a systematic consideration of how best to address coordinated oligopolistic price elevation. Relatedly, there is little overlap in subject matter between the present analysis and most prior work on the legal regulation of price fixing.

The conclusion to the article emphasizes two respects in which the present analysis is incomplete. First, the focus throughout is not on the question that has preoccupied much prior discussion, “How should we define the term

‘agreement’?,” but instead on the question, “What approach toward oligopoly pricing best promotes social welfare?” To answer the latter, it is natural to proceed by examining the nature of the problem and then determining how to identify its presence and to remedy it. Modern competition law emphasizes economic substance over form, has an open-ended, flexible formulation, and could in principle be amended. In any event, even if prevailing doctrine imposes significant constraints, it is best to start by trying to determine what in principle is the most sensible way to address oligopolistic price elevation.

Second, optimal policy depends greatly on empirical evidence in realms where existing knowledge is incomplete. One set of issues concerns the extent of oligopoly pricing that would prevail under various regimes. Another involves the manner in which such oligopoly pricing is achieved (for example, with resort to what forms of communication), and more broadly how much can be detected, by which methods, and at what error cost. Without further knowledge, it is difficult to identify an optimal regime with much confidence. However, the present framework not only guides that decision in the interim but also sharpens the research agenda so that better strategies might be devised in the future.

II. SOCIAL PROBLEM

Section A reviews those aspects of the theory of coordinated oligopoly behavior that are most relevant for the analysis of detection in Part III. Section B examines the welfare implications of oligopoly pricing. Static effects are fairly familiar but dynamic effects are less well known and more subtle, and both are important in understanding the possible adverse consequences of false positives. Section C presents a framework for assessing the effects of a competition regime on oligopoly pricing. Both benefits and costs in terms of behavior are primarily *ex ante*, through deterrence and the chilling of desirable activity. Section D discusses how this analysis provides the lens through which one can view evidence pertaining to the prevalence of oligopoly pricing in the economy, a question of clear importance since the empirical magnitude of the problem bears on how aggressive the policy response should be.

A. COORDINATED OLIGOPOLY BEHAVIOR¹²

Economists study coordinated oligopoly behavior using the theory of repeated games.¹³ This theory aims to explain and predict oligopoly behavior

¹² For more extensive discussion, see Kaplow, *supra* note 3, § IV.A. As the section heading signifies, attention is confined to coordinated oligopolistic price elevation; the relevance of unilateral market power in oligopolistic industries is considered in section V.C.

¹³ The seminal article is James W. Friedman, *A Non-cooperative Equilibrium for Supergames*, 38 REV. ECON. STUD. 1 (1971) [hereinafter *Supergames*]. On modern game theory in general, see DREW FUDENBERG & JEAN TIROLE, *GAME THEORY* (1991), and MARTIN J. OSBORNE & ARIEL

that had previously been the subject of rough, intuitive, and not entirely satisfactory accounts.¹⁴ The relevant branch of the theory is that addressed to what are termed noncooperative games, which are those in which binding agreements (such as might be enforced by a third party) are unavailable.¹⁵ Note that this branch encompasses explicit, classic cartels (even written agreements) because competition law renders them legally void; parties must rely on themselves to enforce their arrangements. Unfortunately, the terminology can be confusing. Within noncooperative *games*, it is common to refer to both cooperative and noncooperative *outcomes*, a distinction that more closely tracks broader usage. Successful coordinated oligopoly pricing is a cooperative outcome; rivalrous competitive pricing is a noncooperative outcome.¹⁶ Importantly, cooperation in this sense describes the character of the result, not the process by which the parties' minds (speaking metaphorically for firms) came to their mutual understanding. Thus, the focus of the theory is on what price levels can be sustained, using what forms of punishment and so forth, making no distinction based on what means of communication the parties may have employed. That is, the same analysis is applicable to classic cartels and to the most informal interdependence.¹⁷

The analysis of firms' interactions employs repeated games because they can capture the sort of strategic interaction that makes successful coordinated oligopoly pricing possible. It is familiar that, otherwise, price competition will yield the competitive outcome in a simple setting involving firms selling homogeneous products because firms that charge above marginal cost will be

RUBINSTEIN, *A COURSE IN GAME THEORY* (1994); on the application to oligopoly theory, see JAMES W. FRIEDMAN, *GAME THEORY WITH APPLICATIONS TO ECONOMICS* (1986) [hereinafter *GAME THEORY*], JEAN TIROLE, *THE THEORY OF INDUSTRIAL ORGANIZATION*, ch. 6 (1988), XAVIER VIVES, *OLIGOPOLY PRICING: OLD IDEAS AND NEW TOOLS* 301–23 (1999), and Carl Shapiro, *Theories of Oligopoly Behavior*, in 1 *HANDBOOK OF INDUSTRIAL ORGANIZATION* 329 (Richard Schmalensee & Robert D. Willig eds., 1989); and on the connection to antitrust, see MICHAEL D. WHINSTON, *LECTURES ON ANTITRUST ECONOMICS*, ch. 2 (2006), and Louis Kaplow & Carl Shapiro, *Antitrust*, in 2 *HANDBOOK OF LAW AND ECONOMICS* 1073, 1103–21 (A. Mitchell Polinsky & Steven Shavell eds., 2007).

¹⁴ See, e.g., Kyle Bagwell & Asher Wolinsky, *Game Theory and Industrial Organization*, in 3 *HANDBOOK OF GAME THEORY WITH ECONOMIC APPLICATIONS* 1851, 1872–73 (Robert J. Aumann & Sergiu Hart eds., 2002); Friedman, *Supergames*, *supra* note 13, at 11. Although the underlying story has long played a prominent role in antitrust commentary and court opinions, the modern theory that makes it more precise has been virtually absent in these arenas. See Kaplow, *supra* note 3, at 701 n.24.

¹⁵ Cooperative games are those in which externally enforceable binding agreements are possible, and the analysis focuses on what agreement parties would reach.

¹⁶ See, e.g., FRIEDMAN, *GAME THEORY*, *supra* note 13, at 20; FUDENBERG & TIROLE, *supra* note 13, at xviii. For further discussion, see Kaplow, *supra* note 3, § IV.A.1.

¹⁷ See, e.g., KAI-UWE KÜHN & XAVIER VIVES, *INFORMATION EXCHANGES AMONG FIRMS AND THEIR IMPACT ON COMPETITION* 43 (1995) (report to the European Commission); Robert E. Hall, 45 *J. ECON. LIT.* 1066, 1067 (2007) (review of WHINSTON, *supra* note 13).

undercut by rivals.¹⁸ Unfortunately for society, firms sometimes are able to charge supracompetitive prices, despite the absence of legally binding agreements.

The basic idea is that starting, let us suppose, at the industry-profit-maximizing (monopoly) price, P_m , no firm will wish to cut its price to steal its rivals' business if it expects this act to induce its rivals to cut price as well, perhaps matching the first firm's lower price and perhaps undercutting it. Whether, when the dust settles, the price is somewhat lower than P_m or all the way down to the competitive price, P_c , the prospective price cutter will be worse off. As long as it does not expect to profit sufficiently in the short run (before rivals cut their prices as well) to make up for the sacrifice of profits in the long run, it will adhere to the price of P_m . Note, importantly, that this logic is equally applicable regardless of whether each firm's expectation about others' reactions arises from their mutual appreciation of their situation or as a consequence of direct discussion of the matter.

Similar logic can explain how the price might rise to P_m in the first place, whether it starts at P_c or at some intermediate level. A firm may well be willing to brave a price increase if it expects (again, whether by conjecture or as a result of explicit discourse) that its rivals will reciprocate. If its rivals indeed cooperate by matching the price increase, the firms will all be better off thereafter, supposing that, by the logic of the preceding paragraph, the price increase can be sustained. As long as the first firm does not lose much profit in the interim due to any delay in others' reactions, the long-run gain will make the venture worthwhile. Moreover, the firm will expect its rivals to follow quickly because they understand (again, either because of their grasp of the circumstances or through prior dialogue) that delay will be taken as defection, leading the initiator quickly to drop its price back to the preexisting level.

The modern game theoretic literature makes the foregoing intuition more rigorous and extends it along a number of dimensions.¹⁹ Perhaps the most important for present purposes concerns the problem that oligopolistic firms face in detecting defections in markets where each firm's prices are not readily observed—a challenge first explored in depth by Stigler.²⁰ In this setting, firms might be able to infer that cheating occurred from a fall in sales of their own products, and in simple models this information would be sufficient. But suppose as well that buyers' demand fluctuates in ways firms cannot observe, raising the possibility that a decline in a firm's sales might have been due to

¹⁸ For more formal discussion that raises subtleties not pertinent to the present task, see, for example, TIROLE, *supra* note 13, at 212–18, and SHAPIRO, *supra* note 13, at 344–46.

¹⁹ See generally sources cited *supra* note 13.

²⁰ See George J. Stigler, *A Theory of Oligopoly*, 72 J. POL. ECON. 44 (1964); see also A. Michael Spence, *Tacit Co-ordination and Imperfect Information*, 3 CAN. J. ECON. 490 (1978).

cheating or instead to less favorable market conditions. In the presence of such uncertainty, firms need to choose a strategy that trades off rapid, sufficiently harsh punishment of actual cheating—in order to deter it effectively—and avoidance of price wars when there was no actual cheating but just a period of unusually low demand. In models of this problem, oligopoly pricing may still be possible, but it is less effective; there will be occasional price wars even when no cheating has occurred. It has also been suggested that this more complicated scenario more closely accords with what has actually occurred in some markets characterized by oligopoly pricing.²¹

This extension and many others show how repeated games can be used to provide more complete depictions of oligopolistic markets, which aids in examining such questions as whether conditions are conducive to successful coordinated oligopoly pricing and whether it is occurring—both important inputs into inferences about the existence of successful oligopolistic coordination, the subject of Part III. Observe, as has been noted throughout, that there is no direct relationship between the analysis of this subsection and whether firms' success is achieved through pure interdependence, highly explicit communications of a particular type, or any other variation. Both extreme, old-fashioned cartels (stripped only of the ability to legally enforce their agreements) and plain interdependence, along with everything in between, are all analyzed as noncooperative games.²² Whether a pair of strategies constitutes an equilibrium for two gas stations engaged in price signaling with their price postings or having a discussion in a smoke-filled room leading them to charge the monopoly price depends on precisely the same calculation that compares the gain from defection with the lost future profits due to the other firm's response.²³ Firms are assumed to act in their own best interests, to maximize profits, in either case. If the gains from cheating exceed the costs, it is supposed that a firm will cheat, and otherwise not.²⁴

²¹ See, e.g., Edward J. Green & Robert H. Porter, *Noncooperative Collusion Under Imperfect Price Information*, 52 *ECONOMETRICA* 87 (1984); Robert H. Porter, *A Study of Cartel Stability: The Joint Executive Committee*, 14 *BELL J. ECON.* 301 (1983); Robert H. Porter, *Optimal Cartel Trigger Price Strategies*, 29 *J. ECON. THEORY* 313 (1983); Jonathan B. Baker, *Identifying Cartel Policing Under Uncertainty: The U.S. Steel Industry, 1933–1939*, 32 *J.L. & ECON.* S47 (1989); Timothy F. Bresnahan, *Competition and Collusion in the American Automobile Industry: The 1955 Price War*, 35 *J. INDUS. ECON.* 457 (1987); Margaret C. Levenstein, *Price Wars and the Stability of Collusion: A Study of the Pre-World War I Bromine Industry*, 45 *J. INDUS. ECON.* 117 (1997); Shapiro, *supra* note 13, at 373–79; see also JEFFREY M. PERLOFF, LARRY S. KARP & AMOS GOLAN, *ESTIMATING MARKET POWER AND STRATEGIES* 104 (2007) (finding the evidence more mixed); Margaret E. Slade, *Strategic Pricing Models and Interpretation of Price-War Data*, 34 *EUR. ECON. REV.* 524 (1990) (examining alternative price-war models).

²² For further elaboration, see Kaplow, *supra* note 3, § IV.A.

²³ See WHINSTON, *supra* note 13, at 21.

²⁴ For qualifications and comments on their relevance, see Kaplow, *supra* note 3, § IV.B.3.

B. SOCIAL WELFARE

Subsection 1 examines static effects of oligopolistic price elevation, noting some subtleties that may be relevant in formulating policy. Subsection 2 addresses dynamic effects that may be important both with regard to the core deterrence benefit of enforcement and the possible cost of chilling beneficial conduct. Just as with the descriptive theory of coordinated oligopoly behavior, implications for social welfare do not fundamentally depend on the means by which oligopoly pricing is accomplished, that is, whether it arises purely through recognized interdependence or results from secret meetings at which detailed plans are formulated.

1. *Static Effects*

The core objection to oligopoly pricing is that prices are higher—higher than the competitive level and also higher than is ordinarily necessary to induce producers to supply goods and services to consumers (a topic examined further in subsection 2). Such high pricing is generally regarded to be undesirable in itself, which makes sense if the objective of competition policy is the maximization of consumers' welfare. Suppose instead that the objective is overall efficiency or social welfare (i.e., also including producers' profits, which ultimately are enjoyed by individuals, perhaps owners or workers). In that case, oligopoly pricing is still objectionable because, despite firms' gain in profits, the excess of price over marginal cost destroys value, resulting in what is referred to as deadweight loss. This loss is associated with sales that would have occurred at the lower, competitive price but are choked off by oligopolistic price elevation. For such forgone sales, buyers' valuations, which by assumption are above the competitive price, exceed marginal cost; hence, what buyers lose exceeds the costs sellers avoid. Thus, from both perspectives, oligopoly pricing is taken to be undesirable because it involves supracompetitive prices.

Commentators disagree about whether competition policy should maximize total welfare or just that of consumers.²⁵ It is worth identifying some of the unappreciated differences in the implications of these competing views (the

²⁵ On the choice between consumer and total welfare as objectives of competition policy, see Louis Kaplow, *On the Choice of Welfare Standards in Competition Law*, in GOALS OF COMPETITION LAW (Daniel Zimmer ed., forthcoming 2011), and Kaplow & Shapiro, *supra* note 13, at 1165–69 (examining the choice with regard to the use of the efficiencies defense in horizontal merger assessment). Note that, even if consumers are to be favored on distributive grounds (perhaps because they are on average less well off than beneficiaries of firms' profits), it does not make sense to ignore producers' surplus entirely, and in any event it tends to be advantageous to achieve distributive objectives more directly, through redistributive taxation and transfers. See, e.g., Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. LEGAL STUD. 667 (1994).

significance of which will occasionally be revisited).²⁶ First, for any given price increase, the loss in consumer surplus exceeds the loss in total surplus because the latter subtracts the additional profits accruing to sellers. Second, even fairly small price elevations are a significant concern if only consumer surplus counts. For example, consider linear demand that has a slope of negative one and intersects a supply curve with a constant marginal cost of 100 at a quantity of 100. A 1% price elevation, from the competitive price of 100 to a price of 101, reduces consumer surplus by 99.5 but total surplus only by 0.5. (Consumers pay 101 rather than 100 for the 99 units they still consume, and there is an additional loss of consumer surplus, equal to the deadweight loss, of 0.5 on the unit no longer consumed, the area of a triangle with a base and height of 1.) For a 5% elevation, price is 105, quantity is 95, lost consumer surplus is 487.5 and total surplus falls by 12.5.²⁷ And at the monopoly price, which is 150 in this example, quantity is 50, lost consumer surplus is 3750, and total surplus falls by 1250. Therefore, the ratio of lost consumer to lost total surplus is nearly 200 to 1 for the 1% elevation, nearly 40 to 1 for the 5% elevation, and only 3 to 1 for the 50% elevation. A corollary of the present analysis is that the incremental loss in consumer surplus as price rises is falling whereas the incremental loss in total surplus tends to be rising.

The basic lessons are that a competition policy concerned with consumer rather than total surplus should optimally be more aggressive and, in a relative sense, this prescription holds particularly for small price increases. Put another way, once one recognizes that there are costs associated with enforcement, as will be discussed in sections II.C and V.A, policy based on total welfare would be especially forgiving of small price elevations, whereas if consumer surplus is the focus, intervention may be optimal for much smaller price elevations.²⁸

Static effects of oligopolistic price elevation are not necessarily limited to allocative inefficiency (deadweight loss) and transfers from consumers to producers. When oligopolistic firms raise price and accordingly reduce output, it need not be the case that output is allocated efficiently among them. To see

²⁶ See Kaplow, *supra* note 25, § 3 (presenting a diagrammatic analysis). It is widely accepted that the difference between the two views is important in certain settings, such as mergers, where those concerned only with consumer surplus would disallow an otherwise on-balance efficient merger unless enough efficiency benefits are passed on to consumers so that prices do not rise at all. It is not generally discussed, however, that focusing only on consumer surplus has implications with regard to many other areas of competition policy, including price fixing.

²⁷ At a quantity of 95 and an elevation of 5, consumer surplus of 475 is transformed into seller profits and deadweight loss is 12.5 (the area of a triangle with a base and height of 5).

²⁸ A subtle but possibly important qualification is that the weight on each dollar of lost consumer surplus (under the view that only consumer surplus matters) need not equal the weight on each dollar of lost total surplus (under the view that total welfare matters). Nevertheless, the point in the text about relative significance holds; that is, the relative importance of large price elevations versus small ones is greater under the total surplus view.

this point, begin with a competitive equilibrium. Each firm sells additional units until the point at which its marginal cost equals price. Since all firms face the same, competitive price, all operate such that their last unit produced has the same marginal cost as the last unit produced by other firms; hence, there is ordinarily production efficiency in the sense that reallocating output among the firms could not reduce and (with rising marginal costs) would instead increase total production costs for the given level of output. When oligopolistic firms elevate price, however, they would each like to sell more output than at the competitive price, but instead they must sell less, for otherwise the elevated price could not be maintained. How much of the smaller industry output is supplied by each firm will depend on the circumstances. Perhaps they share output equally, as hypothesized in some simple models. If products are differentiated, consumers will allocate themselves based on how they value particular products. In any event, there is no guarantee that production will be allocated efficiently, which is a further source of reduced welfare.

An interesting related point is that well-organized cartels have the advantage over looser arrangements that some of this loss in production efficiency can be avoided. For example, firms might agree to let their more efficient members produce more than they otherwise would permit, in exchange for transfer payments.²⁹ Accordingly, if oligopolistic price elevation is to occur to a given extent, classic cartels may entail lower efficiency costs than those that arise under pure interdependence or other less formal schemes.

2. *Dynamic Effects*

Oligopolistic price elevation and deterrence thereof may also have important dynamic effects. As with static consequences, these effects do not on their face depend on the means by which oligopoly pricing is achieved. The dynamic effects considered here pertain to entry and inducing investment more broadly. The latter consideration is also central to understanding the rationale underlying competition law's tendency to prohibit price fixing by groups of firms while (ordinarily, in most jurisdictions) permitting unilateral price elevation by a monopolist, a distinction that has led to some confusion in academic commentary on the appropriate breadth of price-fixing prohibitions.

a. *Entry*

Regarding entry, the literature identifies two competing effects.³⁰ First, to the extent that prices are elevated above marginal cost, there tends to be too

²⁹ See Richard Schmalensee, *Competitive Advantage and Collusive Optima*, 5 INT'L J. INDUS. ORG. 351 (1987).

³⁰ See, e.g., N. Gregory Mankiw & Michael D. Whinston, *Free Entry and Social Inefficiency*, 17 RAND J. ECON. 48 (1986).

much entry on account of a business-stealing effect: firms that enter obtain profits in part by diverting customers from incumbents, whose lost surplus is not taken into account by entrants. Therefore, resources wasted on excessive entry are an additional cost of price elevation. Second, to the extent that products are differentiated, there tends to be too little entry because firms offering new products do not capture all of the surplus generated by their contribution to product variety: firms consider only the revenue they obtain, ignoring infra-marginal consumers' surplus. If this effect is sufficiently large, the additional entry induced by price elevation is socially beneficial.

It is useful to elaborate these points separately for industries selling homogeneous goods and those with differentiated products. Homogeneous goods industries are particularly important because coordinated oligopoly pricing is thought to be less difficult when products are homogeneous (see subsection III.A.2), and empirical evidence on price-fixing prosecutions indicates that most serious violations have involved homogeneous products (see section II.D). In these industries, only the first effect is present, so price elevation would always seem to induce excessive entry.

Consider a simple example of an industry in which actual and prospective entrants face a fixed entry cost and, once that cost is incurred, produce at a common, rising marginal cost. If pricing is competitive, that is, price equals marginal cost, firms will enter the industry until the point at which an additional entrant will no longer be able to cover its fixed costs. This state of affairs will be efficient. For the last firm that enters, its entry lowers prices, which benefits consumers, while the business it captures from others involves output that would have sold at marginal cost and hence did not generate a surplus to the sellers.³¹ A subsequent firm would not enter because, at the lower price it would induce, it could not cover its fixed costs. This unprofitability implies that consumers' further gain is less than the production cost of generating that gain.

Suppose instead that pricing will be at the monopoly level regardless of how many firms enter; that is, coordinated oligopolistic pricing (however achieved) is perfectly effective. In this case, the higher price induces more

³¹ The exposition in the text, following *id.* at 50–54, oversimplifies in ignoring the integer constraint (i.e., that in reality, each firm is a discrete unit, so there cannot be an equilibrium with a fractional firm operating). As they show, when the integer constraint is imposed explicitly, it is possible to have one firm too few, rather than too many. (One way to think about part of the implication is that, in moving from zero firms to one firm, the variety available to consumers does increase; hence, the second effect is operative and in that event implies that too little entry is possible.) They note, however, that in simulations reported in Martin K. Perry, *Scale Economies, Imperfect Competition, and Public Policy*, 32 J. INDUS. ECON. 313 (1984), this qualification is important only if the number of firms would be quite small. This case is considered further later in this subsection.

firms to enter, and more will enter until the additional entrant could not cover its costs (because its share of the monopoly output would be too small, even given that price exceeds its marginal cost). All of the supplemental entry is a social waste. There is no consumer benefit—for price is assumed to be unaffected—while further production costs are incurred.³² Indeed, in this case, essentially all of the oligopolists' profits are transformed into production inefficiency.³³ The conclusion is that the total welfare loss from price elevation in homogeneous goods industries could be even larger than the static inefficiency in cases in which there are not effective barriers to entry—to the extent that such entry does not undermine the price elevation. It is also useful to note that the magnitude of the incentive for excess entry is determined by the magnitude of price elevation; when price equals marginal cost, there is no incentive for excess entry.

There is a caveat to the tendency toward excessive entry, one that is most likely to be important in industries that might otherwise have difficulty supporting one or two firms, say, due to relatively high fixed costs.³⁴ One possibility is that even a one-firm industry in which the monopoly price is charged would not yield enough profit to cover fixed entry costs. Then there would be too little entry if total surplus would make it efficient to have a single producer. This problem arises because even monopolists often cannot capture inframarginal consumer surplus (because they cannot perfectly price discriminate). In this instance, a subsidy would be required; permitting price fixing would not help since it is assumed that even the monopoly price is insufficient to induce the entry of a single firm.

Consider also the case in which a monopoly price is sufficient to induce one firm to enter, but that a second firm is deterred from entering because it is supposed that, conditional on entry, competition would be sufficiently aggressive that the prospective entrant's share of the remaining profit would be inadequate to cover fixed costs. The result would be monopoly (and might warrant direct regulation, particularly if the market cannot support a second firm). Again, subsidized entry might be advantageous, in this case because the induced competition may reduce allocative inefficiency enough to justify the

³² This example is analogous to one presented in Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813, 1871–73 n.200 (1984) (showing how prohibiting de facto horizontal merger to monopoly among competing patents, with the possibility of new entrants, tends to result in lower prices, which improves not only static, allocative efficiency but also dynamic, productive efficiency by reducing entry that would only serve to increase production costs).

³³ In a simple example with linear demand, Gregory Mankiw and Michael Whinston report that in the limiting case the entry effect dissipates 50 percent of the total welfare potential in the market. See Mankiw & Whinston, *supra* note 30, at 53 n.10. In contrast, the deadweight loss from overpricing would be 25 percent of total potential surplus in this example.

³⁴ See, e.g., Kaplow & Shapiro, *supra* note 13, at 1079–80, 1089, 1093–94, 1098.

additional production costs. Freely permitting price fixing, however, would be a poor solution because then the additional fixed costs would be incurred but there would be no price reduction and thus no offsetting benefit. If no subsidization were possible, it would sometimes be better than nothing to allow a supracompetitive yet less-than-monopolistic price, which again might require direct regulation. Hence, although in some cases entry might otherwise be insufficient, permitting price fixing is not likely to be an attractive solution.³⁵

Turn now to industries with differentiated products.³⁶ With product heterogeneity, as mentioned, additional entry not only influences price, creating the business diversion effect just discussed, but also enhances product variety, which consumers tend to find valuable. Moreover, since the total value of enhancing variety includes inframarginal surplus that sellers do not capture (in the absence of perfect price discrimination), the incentive to add varieties is socially suboptimal. Combining these effects, entry may be excessive or inadequate relative to the level that maximizes social welfare.

³⁵ Some have suggested to me that a prohibition on interdependent price elevation might be undesirable because of the problem in covering fixed costs. One could imagine that this point might hold in some cases, but it is difficult to see it as a general argument in favor of the position. Whenever interdependent behavior cannot succeed without elaborate communications, such as through a classic cartel (not all of which succeed in any event), this argument entails allowing explicit price fixing. Perhaps such would be desirable, assuming that price would only be elevated moderately, but no one seems to favor it. If exceptions are required, it is generally thought best to allow them through special legislation, perhaps authorizing price regulation—in part for the reason noted in the text, that price elevation to the monopoly level ordinarily induces excessive entry. In any case, whatever permission, if any, should be offered to cartels would naturally be extended to looser forms of interdependence, which, when successful, may or may not result in a degree of price elevation that produces results superior to those under an unpermissive regime. The core idea is that the fixed-cost problem does not in itself distinguish among the means by which successful oligopolistic coordination may be accomplished. Perhaps there might be cases that could be identified in which pure independence would result in modest price elevation, on average—a level that was in the neighborhood of “just right”—whereas a stronger cartel would cause excessive elevation, and likewise there might be others (ones with less conducive conditions) in which it is the stronger cartel that provided approximately the right inducement whereas pure interdependence would be insufficient. More broadly, a permissive approach tends to allow significant price elevation in some industries, those particularly conducive to collusion, and little or none in others—not an intermediate level in all industries and certainly not an intermediate level in just those industries where such may be desirable.

The view that the difficulty of covering fixed costs justifies legal permission of interdependent price elevation is also inconsistent with standard merger policy, wherein the prospect that mergers may render coordinated price increases more likely is uniformly viewed as a reason for prohibition, never as a defense. Similarly, mergers that significantly increase unilateral market power are condemned for that reason, again with no exception for cases in which price elevation will help cover fixed costs. Such a policy is also applied to facilitating practices, examined further in subsection III.A.1.c.

³⁶ In addition to Mankiw & Whinston, *supra* note 30, see TIROLE, *supra* note 13, ch. 7; VIVES, *supra* note 13, ch. 6; Avinash K. Dixit & Joseph E. Stiglitz, *Monopolistic Competition and Optimum Product Diversity*, 67 AM. ECON. REV. 297 (1977); and Michael Spence, *Product Selection, Fixed Costs, and Monopolistic Competition*, 43 REV. ECON. STUD. 217 (1976).

It is useful to consider two of the factors that influence which case is likely to prevail. The first concerns the price level. As price elevation becomes larger, entry and thus variety increase. Note in particular that, starting at price equal to marginal cost, the business diversion effect is unimportant (as explained above), so the only effect would be increased variety. Some price elevation will accordingly be desirable in terms of total welfare. In this regard, recall from subsection 1 that deadweight loss is negligible as price increases just above marginal cost, so the overall tradeoff in terms of total welfare would be favorable. (However, the marginal loss to consumer welfare from price elevation is greatest at this point, so the tradeoff would not necessarily be favorable if the social concern is only with consumer welfare.) When price elevation is substantial, the business diversion effect becomes larger; moreover, the marginal benefit from additional variety would tend to be smaller. This combination indicates that significant price elevations will more likely be associated with socially excessive entry (in addition to the allocative inefficiency that results from price being above marginal cost).

The second factor concerns the degree of product heterogeneity, that is, the magnitude of consumers' benefit from product variety. Clearly, the greater this benefit, all else equal, the more likely entry will be insufficient rather than excessive. Put another way, incremental entry, at any given point, will be more valuable. In this regard, it is important to recall that, when differentiation is substantial, successful oligopolistic coordination is less likely, making this case of less practical interest when analyzing price fixing. Of most relevance will be cases with only modest differentiation, where results seem unlikely to diverge substantially from those with homogeneous products, where entry tends to be excessive.³⁷ This point is particularly apt with regard to more informal modes of oligopolistic coordination to the extent that the feasibility of such coordination depends even more on there being little or no product heterogeneity. In other respects, however, the dynamic effects of oligopolistic price elevation on entry are largely independent of whether such elevation is achieved purely through interdependence or through more explicit sorts of communication.

b. Investment

Entry is but one form that investment may take. Viewed more broadly, elevated prices tend to reward whatever activities lead firms to be in a position to charge prices in excess of marginal cost and capture a larger share of the market when doing so. The prospect that the allocative inefficiency of

³⁷ Cf. Spence, *supra* note 36, at 234 ("Given monopolistically competitive pricing, high own price elasticities and high cross-elasticities create an environment in which monopolistic competition is likely to generate too many products.").

supracompetitive pricing may often be outweighed by the dynamic gains from investment induced by the prospect thereof constitutes the classic justification for awarding intellectual property rights. More broadly, competition regimes, such as that in the United States, tend to tolerate monopoly pricing as long as the monopoly position was obtained and is maintained through efficient behavior—producing products that consumers value and selling them at sufficiently attractive prices—rather than through exclusionary practices.³⁸ This permissive approach is not without costs; ex ante incentives can be excessive³⁹ and ex post payoffs, with accompanying allocative inefficiency, may be more than is necessary to induce investment. Hence, intellectual property rights tend to be limited, and direct regulation of monopoly is sometimes imposed. However, the judgment behind general competition regimes is that, in other settings, monopolies should be permitted to price as they wish, although the degree to which this view is accepted varies across jurisdictions.

Price elevation achieved by interdependent behavior among oligopolists is qualitatively different in this regard. Such price elevation does not reward firms to the extent that they outperform their competitors but instead bestows profits whose magnitude depends on firms' success in refraining from competition.⁴⁰ To be sure, the prospect of such rewards will induce a variety of ex ante investment behavior. In addition to entry, already discussed, firms might expend more to increase their production capacities if they expect such capacities to have a positive effect on their share of oligopoly rents, and this expansion would be wasteful if such additional capacity will remain idle. Similarly, some activities, like advertising, may be zero-sum (to an extent). However, other types of investment may be socially valuable, such as product improve-

³⁸ See, e.g., *United States v. Aluminum Co. of Am.*, 148 F.2d 416, 430 (2d Cir. 1945) ("The successful competitor, having been urged to compete, must not be turned upon when he wins."); *United States v. Grinnell Corp.*, 384 U.S. 563, 570–71 (1966) (referring to "the willful acquisition or maintenance of that [monopoly] power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident").

³⁹ They may also be inadequate: firms, even monopolies, cannot ordinarily engage in perfect price discrimination (due to difficulties of identifying consumers' valuations and the problem of arbitrage), so they do not capture all consumer surplus, which makes it possible that investments with social benefits to direct consumers that exceed investment costs will be forgone because even the prospect of (non-price-discriminating) monopoly profits is insufficient. (Compare the discussion in subsection (a) of insufficient variety.) When there are positive externalities, investment also may be insufficient (but, with negative externalities, there is an additional reason that it may be excessive).

⁴⁰ A classic statement of the ex ante investment benefits of the prospect of market power and their inapplicability to coordinated oligopolistic pricing is offered by Harold Demsetz:

To destroy such power when it arises may very well remove the incentive for progress.

This is to be contrasted with a situation in which a high rate of return is obtained through a successful *collusion* to restrict output; here there is less danger to progress if the collusive agreement is penalized.

Harold Demsetz, *Industry Structure, Market Rivalry, and Public Policy*, 16 J.L. & ECON. 1, 3 (1973).

ments and cost reductions. Even regarding the more useful activities, the reward from oligopoly pricing will substantially reflect the firms' ability to abstain from competition rather than the merits of their own prior investments. Thus, while permitting oligopolistic price elevation may produce some dynamic gains, there is insufficient nexus to justify price fixing.

Furthermore, oligopolistic pricing reduces potential dynamic benefits. When some oligopolists are more efficient or offer products superior to those of others, successfully coordinated pricing tends to dampen the tendency of better firms to serve an increasing share of consumers. In addition, firms may have less incentive to become more efficient and innovative in the first place because they may not greatly benefit from such activity unless they will be willing to defect from the interdependent arrangement, which they will be reluctant to do if the oligopoly profit margin is substantial. Also, economies of scale are not fully realized by successful oligopolists. A related point, mentioned in subsection 1, is that successful interdependent oligopoly pricing can be worse than old-fashioned explicit cartels because the latter might be able to rationalize production and thus achieve some efficiencies that will not result from mere coordinated price elevation. Competition laws often allow competing firms to merge or enter into joint ventures precisely because of the potential to realize efficiencies; such actions are only permitted subject to review designed to determine whether the gains are sufficient to justify any anticompetitive effects. Oligopoly pricing incurs the costs without producing these benefits. Finally, note once again that none of the foregoing analysis (except the production rationalization point on explicit cartels) suggests any direct distinction between unaided albeit successful oligopolistic interdependence and price elevation achieved through more explicit communication.⁴¹

Some commentators have argued that plain interdependent oligopoly pricing should be permitted because analogous price elevation by monopolists is legal.⁴² In light of the foregoing, this claim is surprising. Competition law explicitly distinguishes and subjects to tough sanctions the efforts by groups of firms to eliminate competition among themselves, in contrast to the price-

⁴¹ A speculation is that the possible investment benefits from price elevation may tend to be greater when appropriability is more difficult, which would tend to be true in less concentrated industries, and these are the settings in which more elaborate communications are often thought more likely to be necessary.

⁴² See, e.g., AREEDA & HOVENKAMP, *supra* note 2, at 8–9 (offering reasons for attacking price fixing, most of which do not distinguish monopoly pricing); *id.* at 232 (arguing against prohibition of pure interdependence because such would be inconsistent with rules on monopoly); *id.* at 272 (same); Turner, *supra* note 5, at 668 (“It would make no sense to deprive lawful oligopolists—those who have achieved their position by accidental events or estimable endeavor—of the natural consequence of their position if the lawful monopolist is left with his.”). *But see* John E. Lopatka, *Solving the Oligopoly Problem: Turner’s Try*, 41 ANTITRUST BULL. 843, 854–55 (1996) (criticizing this view).

elevating behavior of monopolists. Moreover, this differential approach is well founded, as just explained. A final reason the argument is puzzling is that the same logic would allow classic cartels—indeed, even legally enforced cartel prices, for a monopolist can legally force its employees to charge the price it commands—yet those advancing the argument roundly condemn express price fixing.

C. FRAMEWORK FOR DECISION-MAKING

1. *Elaboration*

Detection of successful coordination (examined in Part III) is inevitably imperfect. Accordingly, a central question—probably the most difficult and important one regarding coordinated oligopoly pricing—is deciding how much of what sorts of evidence in various contexts should be deemed sufficient for a finding of liability. It is essential, therefore, to specify the proper framework for making this determination. Although the method is largely generic, it is worth spelling it out in order to focus subsequent analysis and because it has not been much elaborated in legal or economic literature, particularly in light of the fact that error costs are primarily in terms of effects on *ex ante* behavior, which itself is endogenous to the legal regime.⁴³

First, consider situations in which firms have in fact engaged in coordinated oligopolistic price elevation. In such cases, the primary benefit of assessing liability is deterrence. That is, the social benefit is from such behavior occurring less often in the first place, and the prospect of sufficient sanctions for such behavior (taking into account their probability and magnitude) will tend to deter it. That the actual application of sanctions after the fact is not in itself central can be appreciated by considering a perfectly operating regime that succeeds in deterring all price fixing. No successful prosecutions would occur, yet all the benefits of preventing harmful behavior would be obtained. In contrast, if there was no deterrence, price fixing would be widespread; there might be many successful prosecutions, but to no avail. Suppose, for example, that sanctions were sufficiently low that even detected price fixing remained profitable.

Against this background, it is clear that the social cost of false negatives—failures to identify price elevation that has in fact occurred—lies in the loss of deterrence. That is, deterrence will tend to fall the more one insists on stronger proof, whether by requiring that particular means be employed or by insisting that the degree of confidence, however obtained, be higher. The extent of the deterrence reduction will depend on a number of factors. Suppose,

⁴³ The analysis in this subsection draws on the unpublished papers cited in note 10.

for example, that some essentially random fraction of cases of actual price fixing result in detection failure—whether because no one notices the price elevation or because it cannot successfully be demonstrated in adjudication. In that event, expected sanctions for all oligopolistic price elevation would fall. If sanctions were sufficiently high that such activity was still unprofitable, there would be no welfare loss. Relatedly, if it was possible to raise sanctions to make up for the detection deficit, there again would be no difficulty. Because there are often limits to how high sanctions may be (firms will be judgment proof beyond a certain point, for example) and because there are also costs in trying to identify and prosecute acts of price elevation, it seems likely that insisting on greater certainty of proof will involve some loss in deterrence due to the greater portion of false negatives that result.⁴⁴

Of additional concern, the failure to detect oligopolistic price elevation will not be random. Most means of detection will be more effective in some settings than in others. Moreover, firms are likely to have some sense of these differences *ex ante*. Accordingly, if some methods of proof are readily allowed but others are not (they may be disallowed or subject to high proof standards), it may well be that certain groups of firms will be deterred and others not. Further raising the magnitude of sanctions may primarily relate to the former, where there may be little or no deterrence deficit, while having little effect on the latter.⁴⁵ Raising deterrence for these undeterred firms may accordingly require permitting more encompassing means of proof.

In assessing the social welfare consequences of deterrence failure, it is necessary to consider not only how frequently oligopolistic pricing will arise but also the magnitude of price elevations. It is socially more important to deter significant price elevations than small ones. Both deadweight loss and consumer welfare reductions are greater when price is further above marginal cost. And, as the numerical illustrations in subsection B.1 show, the rate of increase in loss in total surplus (deadweight loss) is rising in the extent of price elevation; hence, if total surplus is the social objective, large price increases are disproportionately of concern. Likewise, regarding dynamic efficiency, any marginal benefits from price elevation in terms of increased product variety or otherwise are probably falling in the extent of the elevation, whereas costs of excess entry are rising. (Recall that the business diversion effect depends on the extent to which price exceeds marginal cost.) Hence, it

⁴⁴ Some of the evidence in section D and in subsection IV.A.3 suggests that many regimes currently fall significantly short in deterring price fixing; however, the extent to which such could be remedied with stiffer penalties rather than a higher frequency of applying sanctions is unclear.

⁴⁵ This point is of particular relevance to commonly advocated narrower approaches that *de facto* or *de jure* exonerate oligopolistic price elevation in certain settings. Obviously, raising penalties will not generate deterrence for groups of firms that are effectively immune from them.

is probably much more socially important to deter significant price elevations than small ones. Moreover, as will be elaborated in section III.C, attempting to deter small elevations probably entails relatively greater risks in terms of false positives, the subject considered next.

Second, consider firms that have not engaged in oligopolistic price elevation. The prospect of sanctions may tend to chill (deter) beneficial activity in settings in which firms anticipate that their actions generate a substantial risk of false positives. By analogy, if certain medical procedures involve a high risk of malpractice liability even when doctors behave properly, doctors may refrain from such procedures even when they are beneficial. Commentators, courts, and enforcement agencies that have been reluctant to take too aggressive an approach toward coordinated oligopoly pricing seem to have this sort of concern in mind. The point is especially sharp if evidence that firms engaged in similar or identical pricing behavior is taken to be proof of price fixing, or close to it, for such behavior is the norm even among perfect competitors. Logically, such evidence does not even begin to make the case for oligopolistic price elevation.⁴⁶ What is less obvious is just what chilling effects would look like when sensible but imperfect approaches to identifying oligopolistic price elevation are employed. Accordingly, this subject is examined further in subsection 2 and revisited in section III.C.

Because the most important benefits and costs are in terms of ex ante behavior, it follows that observing the relative extent of erroneous outcomes in adjudicated cases (if such were possible) provides a highly misleading indicator of whether a system is well designed and effectively functioning. To illustrate this point, consider some simple numerical examples. Suppose that the world has 1000 potential cases. Under one regime, 90 percent of true instances of price elevation are detected and punished, whereas liability is mistakenly assessed in 2 percent of cases in which there is no price elevation. Finally, assume that deterrence is quite effective in this regime, so that oligopoly pricing occurs in only 10 of the 1000 markets. Since 90 percent of those 10 cases are subject to sanctions, there will be 9 convictions consisting of true positives, and, correspondingly, one false negative. In the 990 markets with no price elevation, the 2 percent error rate suggests that approximately 20 false positives will arise. The ratio of 20 false positives to 9 true positives, more than 2 to 1, may look quite disturbing, even though deterrence is sub-

⁴⁶ Pricing patterns are explored further in subsection III.A.1.a. Note as a preliminary matter, though, that identical or very similar pricing and price movements are not even suspicious in many instances. What actually requires more explanation are prices that do not move together, for neither perfect competitors nor well-coordinated oligopolists would so behave. Price differences may reflect some combination of product differentiation, cost differences (including cost shocks that affect firms differently), different information, or perhaps deviation from an implicit or explicit price agreement.

stantial and the 2 percent false positive rate may entail little chilling of desirable conduct.

Suppose next that the high false-to-true positive ratio prompts a tightening of the burden of proof. Now false positives are only 1 percent, and in cases in which price elevation does occur, the detection rate is only 80 percent. That is, the false positive rate is halved and the false negative rate doubles. Furthermore, imagine that the lower punishment rate significantly erodes deterrence, such that there are now 100 instances of price elevation. Given the 80 percent punishment rate, there are 80 true positives (and 20 false negatives). Of the 900 instances with innocent behavior, the 1 percent false positive rate translates into 9 false positives.⁴⁷ The ratio of false to true positives is now 9 to 80, or about 1 to 9—equivalently, 2 to 18, which is twenty times more favorable than before (when it was over 2 to 1).⁴⁸ Put another way, initially almost 70 percent of instances in which sanctions were applied were mistakes; now that percentage is barely over 10 percent, so the fraction of mistaken convictions in the total has fallen to nearly 1/7 of its prior level.

But is the second scenario substantially better? Or little better? Or even possibly much worse? Deterrence has fallen greatly; there are 100 rather than 10 settings with oligopoly pricing, so harm is on the order of ten times higher. How about the cost of chilling effects? The number of false positives has fallen from 20 to 9, on a base of nearly 1000, so there is some gain, but one that may well be small, perhaps even tiny in terms of the welfare loss from the chilling of desirable behavior.⁴⁹ Whichever regime turns out to be best, it is clear that the ratio of false to true positive findings—or more broadly an ex post assessment of cases that actually arise—provides a significantly distorted picture of the relative desirability of the two regimes. Instead, the primary determinants are the magnitudes of the favorable and undesirable ex ante effects on behavior.

The approach outlined in this subsection assesses the importance of errors, both false positives and false negatives, in terms of their effects on social welfare. There are two important respects in which this method deviates from most prior analyses of the subject. First, errors are usually viewed almost as if

⁴⁷ That the number of false positives falls when the burden of proof in adjudication is increased need not hold in general. In a more flexible model, the resulting decline in deterrence may lead enforcers (public or private) to be more aggressive, sweeping more cases (including more true negatives) into the system, so even if adjudication produces a lower false positive rate per case it processes, the total number of false positives could rise. For further analysis, see Kaplow, *Burden of Proof*, *supra* note 10, and Kaplow, *Optimal Proof Burden*, *supra* note 10.

⁴⁸ More precisely, it was 20 to 9, which is 2 to 0.9, and 18 is twenty times 0.9.

⁴⁹ Note also that, although not the focus of this section, administrative costs may well be higher under the latter scheme with its tougher burden of proof: because of the resulting decline in deterrence, the total number of positives, true and false, rises from 29 to 89.

they are intrinsically bad, and the manner of specifying their weight is mysterious. Even if one sets aside the previously described problem—that relative error frequencies *ex post* provide a highly misleading indication of the effectiveness of a legal system—it remains unclear how most commentators imagine the importance of either type of error to be determined.

Second, errors are often understood by reference to a formal legal criterion, even when that criterion itself is chosen because it is a proxy indicator of which behavior should be sanctioned. Most relevant for present purposes, many commentators endorse a legal standard that limits punishment to cases in which it can be established that certain forms of explicit communication were employed. Hence, the application of punishment in a case where this standard is not met is regarded as a false positive, which type of error it is taken to be important to avoid. However, if certain classes of such erroneous findings entail the assignment of liability in cases in which there was coordinated oligopolistic price elevation—but without the requisite communications, or at least without proper proof thereof—these false positives would be desirable, not detrimental, in terms of their effect on social welfare. Clearly, the prospect of such false positives (defined by reference to the formal legal rule) will deter undesirable behavior, not chill desirable behavior. Thus, when errors are viewed in a vacuum, there is the danger that, not only might the weights be wrong, but the directional implication will sometimes be the opposite of what is denoted by designating the outcomes as erroneous. This hazard provides a further motivation for carefully attending to the actual nature of chilling effects, the topic of the next subsection, rather than thinking abstractly in terms of false positives.

2. *Chilling Effects*

The chilling of desirable behavior is a concern that strongly motivates past discussions of price-fixing rules but almost always remains implicit,⁵⁰ which makes it difficult to assess its importance and how the magnitude of the problem is influenced by the nature of the rules that are adopted. To begin the

⁵⁰ For example, although Posner offers the only substantial direct assessment of competing regimes for addressing oligopoly pricing, devotes significant attention to the problem of detection, and is obviously concerned about false positives, his latest treatment (like the others) barely mentions what the costs of such a regime might be. See POSNER, *ANTITRUST*, *supra* note 7, at 98 (mentioning without further elaboration that “[a] subtle objection to my suggested approach is that it might discourage entry into monopolistic, duopolistic, or other highly concentrated markets” and failing even to make clear whether this supposed problem is a possible cost of successfully deterring actual oligopolistic behavior or arises only if false positives are anticipated to be likely). Turner’s seminal article advocating a narrow price-fixing prohibition devotes scant attention to the actual costs of a more encompassing approach. See Turner, *supra* note 5, at 669–71. Even more notable, Phillip Areeda and Herbert Hovenkamp’s 300-page treatise volume mentions but does not significantly elaborate the direct or indirect costs of a broader prohibition of the sort Posner advocates. See AREEDA & HOVENKAMP, *supra* note 2, at 227–34.

analysis, it is useful to state two potentially detrimental behavioral effects of enforcement, although they are not the focus here. First, as explored previously, some price elevation may be desirable—for example, by usefully contributing to product variety—so successful deterrence will sometimes be disadvantageous. As a matter of convenience and clarity, this is subsumed under deterrence, the net benefit of which is accordingly reduced (unless the problem can be avoided by allowing exceptions). This effect does not depend on making errors in the detection of coordinated price elevation; rather, it is caused by the prospect of true positives.

Second, although false positives that have a truly random character are detrimental, part of their cost should also be grouped with deterrence. Taking the extreme case, if there is some probability that a firm will bear a price-fixing sanction without regard to how it acts, then the net penalty from actually engaging in oligopolistic price elevation is reduced: this expected random sanction must be subtracted from the expected sanction attributable to true positives in order to determine the net expected legal cost of the activity. Problems of this sort are most likely to arise if adjudication puts heavy weight on whether industry conditions are conducive to collusion and only modest weight on whether successful oligopolistic coordination is actually taking place. Such a strategy, it is clear, would undermine deterrence; insisting on proof of socially detrimental behavior rather than the existence of an opportunity is important, as will be elaborated in subsection III.A.2. Note, however, that the prospect of essentially random sanctions (or certain ones that are a function merely of background conditions) does not tend to discourage particular behaviors,⁵¹ the subject to which we now turn.

Whether and what sorts of chilling effects may arise depend on what forms of proof are employed. Suppose, for example, that sudden, substantial price increases are considered as possible evidence of coordinated oligopolistic price elevation. The inference would, of course, be negated if there had just been a corresponding change in conditions, such as an increase in a common cost. For example, a freeze in Florida that destroys much of the crop of oranges may lead to a corresponding increase in wholesale and retail prices for oranges. If no such explanation is available, the price increase may give rise to an inference of oligopoly pricing. This prospect, in turn, would discourage such price increases, but this effect is desirable: it is the deterrence achieved when true positives are anticipated.

False positives may arise in such cases, however, particularly when one considers that, in response to the foregoing mode of detection, oligopolistic

⁵¹ Random sanctions do, however, discourage entry, supposing that a prospective entrepreneur who never starts a business will escape sanctions.

firms wishing to raise prices might instead increase them gradually, making it more difficult to identify the price change as oligopolistically induced.⁵² Gradual price increases are common, and it may be hard to determine whether they are in response to true cost shocks or instead constitute oligopolistic opportunism. This ambiguity makes false negatives and false positives more likely. The former reduce deterrence, encouraging oligopolistic firms to raise prices. The latter produce chilling effects.

To pursue this second effect, suppose that firms in an industry see their costs rising and anticipate that they may continue to do so for some time. Furthermore, imagine that this is an industry in which collusion is plausible but is not taking place. Firms might fear that passing on their cost increases through price hikes would produce some risk of (erroneous) liability, particularly if their cost increases are hard to document. (Rising prices of tangible inputs, like labor, materials, and electricity would tend to be easy to establish, but perhaps the need to increase reserves for subsequent maintenance that has become more difficult due to heightened regulations would be more difficult to prove.) How might such firms react?

Since price increases are assumed to be the basis for the risk of mistaken liability determinations, firms may raise prices to a lesser extent than they would otherwise. This moderation will tend to be inefficient, recalling the supposition that prices were competitive, not elevated, to begin with. In an industry with rising marginal costs, the need to restrain price increases would induce firms to reduce output relative to the level they would produce if they could increase prices appropriately. If all firms behave in this manner, industry output will be less than total demand, which is unsurprising since price is being set below the competitive equilibrium level. Necessarily, consumers' loss in surplus from the sales reduction exceeds the production cost for these units, so there is a net social loss. Note that consumer surplus obtained on units still sold would rise on account of the pricing restraint, this gain being a transfer from producers. Another effect of the prospect of such pricing restraint is that entry in the industry would be less attractive because there would be less producer surplus to cover fixed costs. Although entry can be excessive—particularly in homogeneous goods industries, as subsection B.2.a explores—such excess arises when price is elevated above marginal cost. When prices are instead suppressed, entry is too little from a social perspective, and all the more so if there is also lost product variety.

Thus, the prospect of false positives from truly cost-justified price increases, when prices are otherwise at competitive levels, may have chilling effects—price and quantity reductions and reduced entry—that are socially

⁵² See *infra* note 78 and accompanying text.

detrimental. Nevertheless, as long as the degree of induced price suppression is small, these losses will be insubstantial. However, if marginal cost was constant rather than rising, the problem could be much worse. The reason is that price suppression, beginning at a competitive level at which price equals marginal cost, would render production entirely unprofitable, inducing exit. At the point at which only a single firm remained, or only one entered in the first place, this problem would vanish because a single firm cannot be guilty of joint price elevation. However, as a monopolist, this lone firm may not merely cover costs, including any that may be difficult to document, but also charge a monopoly price.⁵³

It is familiar that constant-marginal-cost industries, when there are fixed costs, pose a potential problem of natural monopoly, as do industries in which firms have declining marginal costs. Moreover, this problem exists aside from the present concern with false positives in the enforcement of an anti-price-fixing regime, as subsection B.2.a notes. The social response sometimes is public utility regulation, but it generally is not to permit private price fixing. A *laissez-faire* approach is not entirely without virtues, for as more firms enter, prices may come down somewhat as price elevation is more difficult to maintain, and it is possible that the benefit from such price reduction would outweigh the waste from excessive entry (although it may not). In any case, it is hardly clear that an optimal response would be to permit whatever price elevation can be achieved without certain forms of explicit communications but to disallow all other price elevation.⁵⁴

Consider next a case in which one or some—but not nearly all—firms in an industry enjoy sustained high profits as a consequence of charging prices that are well above average cost. It might appear that coordinated oligopoly pricing has occurred. But another possibility (a false positive if liability were based on such evidence) is that such firms enjoy cost or other advantages over their rivals. If there is no oligopolistic price elevation, then such firms may be selling at marginal cost, despite appearances to the contrary: they may have rising marginal costs such that the marginal cost of much of their production is below the market price, which would explain why average cost is notably less than price. There may also exist other firms with smaller market shares or with significant sales at higher average cost. The prospect that the more efficient firms might be found liable would tend to discourage *ex ante* investments in cost reduction or quality improvement.

To avoid such chilling effects, it seems important that assessments of cost focus on marginal rather than average costs. Moreover, in the setting just de-

⁵³ As subsection B.2.a discusses, this situation would often be better than one with many firms charging an equally elevated price because in the latter case resources are wasted on entry.

⁵⁴ See *supra* note 35.

scribed, firms' costs may differ. If the market were in fact competitive, all firms will be equating the common market price to marginal cost, so marginal costs would be identical. However, since marginal cost may be difficult to measure, leading to the use of various proxies that perhaps more nearly indicate average cost, then costs would appear to be different. In this case, adjudicators should look to the higher-cost firms. If their marginal costs, as best can be determined, more nearly equal the market price, then there may be little concern of coordinated oligopoly pricing. The more profitable firms would owe their success to efficiency, not price fixing.⁵⁵ This suggestion also poses dangers of false negatives, however, particularly when most sales are by a number of firms with lower costs—who may be pricing well above marginal cost due to successful oligopolistic coordination—and some sales are made by a competitive fringe of less efficient firms. The problem is that, if such cases are to be identified and condemned, there may sometimes be false positives, giving rise to the prospect of chilling desirable *ex ante* investment.

The most relevant lesson for present purposes is that detrimental chilling effects may arise as a consequence of the prospect of imposing sanctions for price elevation that is not in fact due to successful oligopolistic coordination but instead is in response to industry conditions (such as costs) that are difficult to ascertain. Thus, as will be discussed in section III.C, it tends to be optimal to adjust proof requirements accordingly—perhaps, for example, by crediting ambiguous evidence of higher marginal costs, particularly when the degree of price elevation that must thereby be explained is modest.

There are, of course, other means by which oligopoly pricing might be proved, and it remains to consider whether they too may involve chilling effects. Many such possibilities will be explored in Part III, which addresses various means of detection. Yet another consideration regarding chilling effects is the manner in which sanctions are determined, a subject taken up in section IV.A. Briefly, suppose that sanctions are a direct function of the determined severity of the violation. A consequence of this approach is that small price elevations will be met with low sanctions. Regarding chilling effects that arise when marginal cost is underestimated, the problem may often involve small estimated overcharges, which accordingly would result in small penalties and thus only modest effects on *ex ante* behavior. In contrast, an approach that makes the penalty largely independent of the overcharge (a tendency under guidelines for fines in the United States and the European Union, discussed in subsection IV.A.3) will levy relatively harsher sanctions in such cases, increasing the magnitude of chilling effects—and also producing less deterrence in cases with unusually large price elevations, which are less likely to involve false positives.

⁵⁵ See Demsetz, quoted *supra* note 40.

D. EMPIRICAL EVIDENCE

Empirical examination of the extent of oligopolistic price elevation in various industries in the United States peaked in the 1980s. A prominent survey concludes that “[t]here is a great deal of market power, in the sense of price-cost margins, in some concentrated industries” and “[o]ne significant cause of high price-cost margins is anticompetitive conduct.”⁵⁶ A number of the studies reported show industry price-cost margins in the neighborhood of 0.5, which indicates that price is double (100 percent above) marginal cost; other industries have much lower margins.⁵⁷ Other surveys provide further evidence of oligopoly pricing in many industries and in various local markets.⁵⁸

More detailed evidence is provided by studies of litigated—mostly government prosecuted—cases.⁵⁹ In these cases, there typically were highly explicit communications among the defendants, and oligopolistic pricing was usually achieved. Connor’s broad survey of the literature finds that median measured overcharges, averaged over the period of such price-fixing conspiracies, were approximately 25 percent, with a mean of about 40 percent; nearly two-thirds of the episodes had overcharges above 20 percent.⁶⁰ Overcharges were about a

⁵⁶ Timothy F. Bresnahan, *Empirical Studies of Industries with Market Power*, in 2 HANDBOOK OF INDUSTRIAL ORGANIZATION 1011, 1052–1053 (Richard Schmalensee & Robert D. Willig eds., 1989).

⁵⁷ *Id.* at 1051 tbl. 17.1.

⁵⁸ See, e.g., F.M. SCHERER & DAVID ROSS, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 426–47 (3d ed. 1990); LEONARD W. WEISS, ed., CONCENTRATION AND PRICE (1989); Robert E. Hall, *The Relation Between Price and Marginal Cost in U.S. Industry*, 96 J. POL. ECON. 921 (1988); Richard Schmalensee, *Inter-Industry Studies of Structure and Performance*, in 2 HANDBOOK OF INDUSTRIAL ORGANIZATION 951 (Richard Schmalensee & Robert D. Willig eds., 1989). Based on a review of historical evidence in the petroleum, automobile, tobacco, and airline industries as well as brief reports on many others, James Brock concludes: “Considered in this light, perhaps the inordinate degree of solicitude for oligopoly that has arisen in the courts and antitrust agencies is profoundly misplaced.” James W. Brock, *Antitrust Policy and the Oligopoly Problem*, 51 ANTITRUST BULL. 227, 280 (2006). As suggested by the analysis in section B (see especially note 40, quoting Demsetz), the fact that concentration is often associated with oligopolistic price elevation does not negate the idea that the creation of concentration through firm growth may reflect efficient behavior. See, e.g., Michael Salinger, *The Concentration-Margins Relationship Reconsidered*, in BROOKINGS PAPERS ON ECONOMIC ACTIVITY: MICROECONOMICS 287, 291, 310 (Clifford Winston & Martin Bailey eds., 1990). Concentration attributable to horizontal mergers may be viewed quite differently in this regard. See, e.g., Brock, *supra*.

⁵⁹ See, e.g., JOHN M. CONNOR, GLOBAL PRICE FIXING (2d updated & rev. ed. 2008) [hereinafter GLOBAL PRICE FIXING]; John M. Connor, *Price-Fixing Overcharges: Legal and Economic Evidence*, 22 RES. L. & ECON. 59 (2007) [hereinafter Survey]; John M. Connor & Robert H. Lande, *How High Do Cartels Raise Prices? Implications for Optimal Cartel Fines*, 80 TUL. L. REV. 513 (2005); George A. Hay & Daniel Kelley, *An Empirical Survey of Price-Fixing Conspiracies*, 17 J.L. & ECON. 13 (1974); Joseph E. Harrington, Jr., *How Do Cartels Operate?*, 2 FOUNDATIONS & TRENDS IN MICROECONOMICS 1 (2006); Margaret C. Levenstein & Valerie Y. Suslow, *What Determines Cartel Success?*, 44 J. ECON. LIT. 43 (2006).

⁶⁰ Connor, Survey, *supra* note 59, at 59, 90, 94–95. This survey covers 259 publications involving 279 markets and 512 episodes, with 770 observations on average overcharges. See *id.* at

quarter lower in the United States and a quarter higher in the rest of the world.⁶¹ Particularly notable cases involved international vitamins cartels in operation during the 1990s, with total overcharges of nearly \$9 billion (in 2005 dollars), with almost a third in the United States.⁶²

What do we learn, and fail to learn, from such studies? This question is best examined against a template indicating what we in principle need to know to design a legal regime in light of the analysis earlier in this Part.

Initially, we would like to know the extent of coordinated oligopoly pricing and how it varies with enforcement.⁶³ Even if there were no explicit enforcement, oligopoly pricing may not be rampant, for theory indicates that coordination is often difficult to effectuate even when sanctions are ignored, and history suggests that many markets seemed to behave fairly competitively despite the absence of competition law. It is worth keeping in mind that the mere fact that laws make cartel agreements legally unenforceable has some effect, and possibly a significant one, in reducing the extent of oligopoly pricing.

The existing empirical evidence offers some illumination. Given the extent of oligopolistic price elevation and the number of successful prosecutions of explicit price-fixing arrangements involving substantial overcharges, it seems safe to conclude not only that, in the absence of enforcement, there would be a significant problem but that, even with substantial modern enforcement and penalties much stiffer than had existed in the past, the existing level of deterrence may be insufficient. Indeed, even detected and punished cartels may suffer little or realize net gains.⁶⁴ Underdeterrence is compounded to the extent that detection is highly incomplete. A natural way to test this hypothesis and to estimate the probability of sanctions would be to see what portion of industries and markets in which elevated oligopoly prices identified independently of enforcement activity have been subject to prosecution. That is, much could be learned by combining the empirical industrial organization literature aimed at understanding oligopoly, without attention to the legal regime, with studies of the operation of that regime that focus only on prosecutions. Limited evidence to date suggests that the probability of detection is quite low.⁶⁵

79–82. See also Connor & Lande, *supra* note 59, at 540, 559–60 (reporting mean and median overcharges in all published scholarly economic studies of cartels of 49 percent and 25 percent, and for all final verdicts in U.S. antitrust cases of 31 percent and 22 percent).

⁶¹ See Connor, *Survey*, *supra* note 59, at 59, 90–92.

⁶² CONNOR, GLOBAL PRICE FIXING, *supra* note 59, at 338 tbl. 12.1.

⁶³ It is also necessary to know the resulting social cost: the deadweight loss from static allocative inefficiency, the extent of resources dissipated by excessive entry and other forms of rent-seeking, and also other effects, including positive ones, on firms' incentives. The same is true for chilling effects.

⁶⁴ This is a theme of CONNOR, GLOBAL PRICE FIXING, *supra* note 59.

⁶⁵ See *id.* at 394 n.1 (citing informal evidence from forensic economists and antitrust defense counsel speculating “that as few as 10% of all price-fixing conspiracies are investigated or prose-

To the extent that deterrence may currently be inadequate—and it is unlikely to be optimal in most competition regimes since the scope of laws, level of enforcement, and penalties vary widely—it is also necessary to know how the extent of oligopolistic price elevation varies with enforcement. Since enforcement, as just mentioned, differs across jurisdictions and there also have been important legal changes, particularly in penalties, within jurisdictions, it may be possible to identify incremental deterrence effects, although little such study has been undertaken.⁶⁶ One challenge in conducting such an investigation regards ambiguity in the law, particularly the law as actually enforced and as perceived some years beforehand by firms when deciding how to behave, both of which may diverge from hornbook statements.

Finally, a central focus of this article is on proof requirements—what must be demonstrated with what degree of confidence to establish a violation—which are also important determinants of deterrence. Like enforcement effort, these different dimensions are not perfect substitutes in their deterrent effect. For example, if some forms of behavior are difficult to detect or are deemed to be immune, only changes in the prerequisites for establishing liability can have an appreciable impact on deterrence with regard to the settings in question. Unfortunately, even if one ascertained which markets involved likely oligopolistic price elevation yet no enforcement, we would be unlikely to know the methods by which such price elevation was achieved or what evidence might be obtainable, so it would be difficult to gauge how various changes in the legal regime would influence deterrence in these settings.

Studies of prosecuted cases, involving detected explicit communications, do provide some valuable information about the markets involved. George Hay and Daniel Kelley conclude their investigation by observing:

cuted”); Peter G. Bryant & E. Woodrow Eckard, *Price Fixing: The Probability of Getting Caught*, 73 REV. ECON. & STAT. 531 (1991) (estimating the probability of a U.S. federal indictment to be at most 0.13–0.17 per year, although these estimated probabilities are only for those conspiracies eventually caught); John Connor, *The United States Department of Justice Antitrust Division’s Cartel Enforcement: Appraisal and Proposals* 8–9 & n.20 (Am. Antitrust Inst. Working Paper 08–02, 2008) (citing studies supplementing Bryant and Eckard and reaching similar conclusions for the United States during 1990–2004 and for the European Union’s prosecution of international cartels from 1969–2003); Connor & Lande, *supra* note 59, at 524–26 (noting Assistant Attorney General for Antitrust Ginsburg’s estimate in 1986 before the Sentencing Commission that the probability is at most 10 percent).

⁶⁶ See Julian L. Clarke & Simon J. Evenett, *The Deterrent Effects of National Anticartel Laws: Evidence from the International Vitamins Cartel*, 48 ANTITRUST BULL. 689 (2003) (finding smaller price elevations in jurisdictions with active and effective competition law regimes); cf. GEORGE SYMEONIDIS, *THE EFFECTS OF COMPETITION CARTEL POLICY AND THE EVOLUTION OF STRATEGY AND STRUCTURE IN BRITISH INDUSTRY* (2002) (analyzing the effects of the introduction in the United Kingdom of the 1956 Restrictive Trade Practices Act); Michael Kent Block, Frederick Carl Nold & Joseph Gregory Sidak, *The Deterrent Effect of Antitrust Enforcement*, 89 J. POL. ECON. 429 (1981) (finding that increased enforcement intensity and private class actions reduce markups in the bread industry).

A brief summary of our empirical results would be that conspiracy among competitors may arise in any number of situations but it is most likely to occur and endure when numbers are small, concentration is high and the product is homogeneous. We suspect these results will conflict with at least some previously held opinions on the expected locus of conspiracy, and conversely on the ability of oligopolists to regularly attain monopoly profits through tacit collusion.⁶⁷

As with any evidence drawn from prosecuted cases, however, it does not tell us about other cases. One possibility is that there may be more settings with larger numbers of firms and less conducive conditions, although this possibility does not seem particularly likely unless such cases involve little prolonged overcharging, in which event their omission would be less important.⁶⁸

Another, more important possibility is that there are a number of industries in which conditions are highly conducive to coordinated oligopolistic price elevation—small numbers, homogeneous products, and so forth—yet there are also many unprosecuted cases, either because the behavior was undetected or it could not be prosecuted due to limited evidence or on account of a view that only more informal coordination occurred and such is legal. Consider, for example, industries or markets identified in the empirical literature in which it appears that oligopoly pricing is present but there have not been price-fixing challenges. If indeed there was coordinated oligopolistic price elevation, we do not know how it was achieved, such as by pure interdependence or highly explicit communication. Nor do we know what evidence might have come to light if there had been a serious investigation. Accordingly, it is difficult to know the extent to which the features of prosecuted criminal cases—that were selected (in the United States) to meet a “beyond a reasonable doubt” proof standard and are probably thought to require direct and highly probative evidence of explicit, detailed, face-to-face communications—are representative of the larger, submerged portion of the iceberg.

Turning to chilling effects, the situation with regard to empirical knowledge is far worse. Essentially nothing is known about existing regimes or how changes in enforcement instruments would influence the nature and magnitude of such effects. Nor would it be easy to learn about these questions. It would be necessary to study firms that do not elevate prices but might be confused for firms that do, in order to identify how their behavior may be affected by their perceived probability of mistakenly being subject to sanctions. Some illumination might come from examining areas in which there

⁶⁷ Hay & Kelley, *supra* note 59, at 26–27.

⁶⁸ *See id.* at 24 n.15 (“It should be kept in mind that most of the conspiracies were found in concentrated markets even though *a priori* these are the conspiracies which are most likely to escape detection. Thus if there is a bias, it should result in the underreporting of conspiracies in markets with high concentration.”).

have been government prosecutions or private cases alleging price fixing but where there does not appear in fact to have been oligopolistic price elevation.

III. DETECTION

The central challenge in addressing coordinated oligopolistic price elevation is detection.⁶⁹ Because firms have incentives to hide their behavior to the extent that it may be illegal, it will be difficult to identify instances of successful oligopolistic coordination. Relatedly, because an aggressive approach may well be necessary, sometimes false positives will occur, resulting in undesirable chilling effects. As a consequence, it is important to consider all pertinent means of inference and to figure out how to employ them in complementary ways.

Section A focuses on evidence derived from the observation of market conditions and behavior. Most directly, this entails considering different ways of inferring successful oligopolistic coordination. Attention is also devoted to the relevance of the conduciveness of industry conditions. Section B explores evidence from firms—whether records of strategy or internal indicators of what is actually taking place in the market—that bears on these same issues. It also considers the relevance of and evidence on interfirm communications. Taken together, the analysis will suggest that confident identification of coordinated oligopolistic price elevation will sometimes be possible and other times the available evidence will be quite murky. In intermediate cases, of which there may be many, section C considers further how liability should be assessed in light of the decision-making framework articulated in section II.C. The trade-off of deterrence and chilling effects will depend on the types of evidence available, and some suggestions will be made about how to adjust methods of inference and proof requirements accordingly.

Note at the outset that this Part is not concerned with the question whether successful oligopolistic price elevation is brought about by one or another means of interfirm communications, although as just noted such communications may be evidence bearing on whether successful coordination has occurred. The reason is that the social consequences of oligopoly pricing do not depend directly on this matter, as section II.B discusses. Whether some methods of proof should be privileged above others should reflect analysis of the pertinent inferences and application of section II.C's framework. Accordingly, any special or exclusive relevance of interfirm communications, or any other indicator of oligopoly pricing for that matter, is not to be determined a priori but instead needs to emerge from the appropriate functional analysis.

⁶⁹ See POSNER, *ANTITRUST*, *supra* note 7, at 98–99; Posner, *Oligopoly*, *supra* note 7, at 1578, 1583, 1593.

A. MARKET-BASED EVIDENCE

1. *Means of Inferring Successful Oligopolistic Coordination*

There are myriad means of inferring the existence of successful oligopolistic coordination,⁷⁰ many of which are not independent of each other; that is, some are likely to be simultaneously present or may otherwise be mutually reinforcing. A converse is equally important to keep in mind: regarding many factors that imply coordination, their absence may often negate the existence of coordination, where absence should be understood as not merely constituting ambiguity and difficulty of proof but rather demonstration of non-existence.

Despite the overlap and interrelationship, it is useful to group indicators into two clusters—pricing patterns and price elevation—which will be explored in turn. Certain pricing patterns, such as sudden price increases unrelated to cost shocks or the outbreak of price wars, may be indicative of oligopolistic coordination rather than competitive behavior. As will be emphasized, mere similarity of price movements and other parallel activity are not suspicious in this regard, a point about which there has been some confusion. Two caveats should be noted at the outset. One is that price elevation is not a sufficient condition for the existence of oligopolistic *coordination* because a competing hypothesis is the exercise of unilateral market power (although price elevation may be a sufficient condition for many of the harms identified in section II.B). This possibility is deferred until section V.C. Additionally, the absence of price elevation indicates that there was no *successful* coordination but not that none was attempted; as section V.B explores, it may well be optimal to punish even unsuccessful attempts, particularly if they can be identified with confidence.

This section also discusses a third type of indicator, which involves practices that may facilitate successful coordination. These may be further indicators of coordination but also may be given independent legal significance. In particular, one may wish to regulate facilitating practices themselves, without

⁷⁰ Prior treatments (which differ, notably by not focusing on chilling effects) include PERLOFF ET AL., *supra* note 21; POSNER, ANTITRUST, *supra* note 7, at 79–93; Jonathan B. Baker & Timothy F. Bresnahan, *Economic Evidence in Antitrust: Defining Markets and Measuring Market Power*, in HANDBOOK OF ANTITRUST ECONOMICS 1 (Paolo Buccirossi ed., 2008) [hereinafter *Economic Evidence*]; Jonathan B. Baker & Timothy F. Bresnahan, *Empirical Methods of Identifying and Measuring Market Power*, 61 ANTITRUST L.J. 3 (1992); Bresnahan, *supra* note 56; Timothy Bresnahan, *Testing and Measurement in Competition Models*, in 3 ADVANCES IN ECONOMICS AND ECONOMETRICS: THEORY AND APPLICATIONS, SEVENTH WORLD CONGRESS 61 (David M. Kreps & Kenneth F. Wallis eds., 1997); Joseph E. Harrington, Jr., *Detecting Cartels*, in HANDBOOK OF ANTITRUST ECONOMICS 213 (Paolo Buccirossi ed., 2008); Kaplow & Shapiro, *supra* note 13, at 1087–95; and Robert H. Porter, *Detecting Collusion*, 26 REV. INDUS. ORG. 147 (2005).

requiring demonstration in a particular case that their use resulted in oligopolistic price elevation.

Before proceeding, note that many market-based means of inferring successful oligopolistic coordination do not reveal how success was achieved, notably, the existence, extent, and nature of any interfirm communications. There may be direct evidence of such interactions and also indirect evidence and inferences, some of which may be based on the factors considered here. The point is that, as an initial matter—without often substantial further and possibly complicated effort—one cannot ordinarily tell what sorts of behavior contributed to the mutual understanding among the firms that produced price elevation or the conditions preceding the outbreak of a price war, for example.⁷¹

a. Pricing Patterns

Certain pricing patterns may indicate successful oligopolistic coordination or a breakdown that implies its prior existence. Given a long history of confusion on this subject,⁷² it is best to begin by emphasizing that the presence of parallel (i.e., common) pricing and other behavior is not usually a symptom for the simple reason that ordinary competitive interaction also has this character. Indeed, when competition is vibrant, most pricing and other behavior are parallel.⁷³ When firms' costs increase, their prices rise and quantities fall.

⁷¹ See, e.g., Alexis Jacquemin & Margaret E. Slade, *Cartels, Collusion, and Horizontal Merger*, in 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 415, 452 (Richard Schmalensee & Robert D. Willig eds., 1989) (in discussing certain economic methods of identifying collusion, they state: “[I]t is impossible to distinguish pure tacit collusion from . . . explicit cartel agreements. What matters for the empirical estimates is the outcome and not the cause of noncompetitive pricing.”). It is sometimes suggested that one occasionally can tell whether elaborate, explicit communications have occurred, particularly when behavior is especially sharp and precise (e.g., secret bids that are identical down to many digits). Even this point is overstated, for often the opposite inference might be made instead because, the more explicit were the communications, the more readily firms could have orchestrated their behavior so as to avoid leaving clear, visible tracks.

⁷² Such confusion may seem surprising because the point is so obvious from the most elementary economic analysis—at a simple enough level that not even introductory study is required—and also because, in the United States, the Supreme Court emphatically rejected the conflation in *Theatre Enterprises v. Paramount Film Distributing Corp.*, 346 U.S. 537 (1954). Poor use of language may be partly responsible. See *infra* note 73. A group of economists filed an amicus brief in *Twombly* that is devoted substantially to the point that parallel behavior is commonplace, so that punishing it would be problematic. See Brief of Amici Curiae Economists in Support of Petitioners, *Twombly*, 550 U.S. 544 (No. 05-1126). (In *Twombly* itself, the dispute involved the standard for surviving a motion to dismiss, which question is intimately related to the substantive legal test for liability but involves separate issues as well. See *infra* section V.A.)

⁷³ The term “conscious parallelism” is often used. To the extent that it refers to parallel behavior of which firms are conscious, it is little different, for each competitor is ordinarily aware that, when it, for example, must increase price because of a rise in the cost of some widely employed input, its competitors will behave similarly. But sometimes this terminology, or even parallelism alone, is used as a shorthand for interdependence, which is quite different. Likewise, the term

When demand increases, firms' prices and quantities both rise. When technology changes, consumers' locations or tastes shift, regulations are modified, and so forth, competitors react similarly, even identically. No doctor would view the fact that a patient is breathing as evidence of bronchitis (although it is a necessary condition); only atypical breathing (wheezing, coughing, shortness of breath) would be even *prima facie* symptomatic, warranting further examination. Likewise, it makes no sense to deem parallel pricing or other commonly undertaken behaviors as even indicative of oligopolistic price elevation since such activity is ubiquitous and, in particular, characterizes innocent competitive activity.

The goal is to distinguish successful oligopolistic interdependence from competitive, independent, rivalrous behavior. As a logical matter, traits shared by both categories, such as parallel pricing, are not useful in drawing the distinction.⁷⁴ Instead, analysis should focus on behavior that is consistent with oligopolistic interdependence and inconsistent with competition, which favors liability, and behavior consistent with competition but inconsistent with interdependence, which disfavors liability. In short, we are concerned with the observational differences between competition and oligopolistic price elevation.

In considering pricing patterns that may support inferences of successful oligopolistic coordination, it is useful to consider three phases: raising prices from a competitive to a supracompetitive level (or further escalating prices), maintaining oligopolistic prices, and price drops, notably, as a consequence of price wars.⁷⁵ Initiation or enhancement of oligopolistic price elevation may be marked by a sharp price increase. Sudden price increases, however, can also occur in competitive markets, most obviously when there is a cost shock, such as a sudden increase in the price of oranges for grocery retailers or in the price of oil for sellers of gasoline and other petroleum-based products. Accordingly,

"independent" is usually used to mean the opposite of interdependent but is sometimes used to include interdependence. See Kaplow, *supra* note 3, §§ I.B & II.A.

⁷⁴ Traits that appear to be common to neither are perplexing, unless they are shared by some third, omitted category. Because competition and oligopoly are often the only two important hypotheses under consideration (putting to the side unilateral market power, the subject of section V.C), evidence seemingly inconsistent with both may warrant further investigation to ascertain whether it may, after all, be consistent with one or another category of behavior (or even both). See also *supra* note 46 (suggesting that nonparallel behavior is more in need of explanation).

⁷⁵ The best methods may differ when detecting collusion in auctions, such as when the government puts construction contracts or procurement out to bid. See, e.g., Patrick Bajari & Lixin Ye, *Deciding Between Competition and Collusion*, 85 REV. ECON. & STAT. 971 (2003); Porter, *supra* note 70, at 159–62. On the difficulty of detecting bid-rigging in auctions, in part because cartel members can submit phantom bids (for all but the member designed to win) that are devised to elude detection by announced techniques, see Robert H. Porter & J. Douglas Zona, *Detection of Bid Rigging in Procurement Auctions*, 101 J. POL. ECON. 518 (1993).

it is also necessary to check for concurrent cost increases or other changes, like sudden shifts in demand, that may explain the price increase.⁷⁶ When price increases are large and sudden and other simultaneous changes of corresponding magnitude seem unlikely, coordinated oligopolistic price elevation seems plausible.

Note that independent explanations must match the changes that have actually transpired. For example, under competition, increases in demand would ordinarily be associated, at least in the short run, with higher prices and firms each producing more (unless they are capacity constrained). In contrast, oligopolistic price elevation involves higher prices but reduced quantities. Accordingly, it is important to consider changes in production as well as in price when attempting to infer whether price increases are competitive or oligopolistic.

Now consider cost shocks. With competition, these ordinarily result in higher prices and reduced quantities and thus appear more similar to oligopolistic price increases. Cost increases may be identifiable through examination of input markets and also the markets for other goods. Moreover, some cost shocks will not have common effects across firms and thus should be reflected in differential quantity reactions. For example, some firms may use more of certain inputs than others (transportation costs for certain inputs or for final products are an obvious example) or be subject to different changes in input prices (labor costs and rental prices tend to differ geographically).

There are, however, a number of reasons that oligopolistic price increases may not be sudden and sharp, making this means of detection less useful. One is that oligopolists may increase their prices in smaller steps because they do not fully trust one another. Perhaps a price leader is uncertain whether rivals will follow its moves quickly and completely and does not want to risk losing significant market share, which may be costly to recoup, if they do not. Or firms may be uncertain of what others believe to be the best price, which may be true when coordination is accomplished with little explicit, direct communication. In such cases, oligopolistic price increases may be harder to detect, but this need not be true. Instead, signaling and jousting about price, such as through sequential price changes, where firms await others' reactions and then proceed, may sometimes provide an even stronger basis for inferring oligopolistic elevation.⁷⁷ Even if each price adjustment is smaller, there will be more of them, and it may be even clearer that cost or other exogenous changes cannot explain observed patterns.

⁷⁶ Demand shifts often are not sudden but they can be, such as in response to sharp price increases of substitutes (perhaps themselves attributable to cost shocks in other markets).

⁷⁷ Closely related are advance price announcements, considered further in subsections (c) and B.2.

A greater difficulty is that sophisticated firms, aware of what inferences may be drawn from their price increases, may instead adjust prices strategically in order to disguise their oligopolistic behavior.⁷⁸ Of course, gradual but substantial price increases or occasional smaller jumps in fairly close proximity may also be suspicious. But by spreading out the changes and perhaps timing them with other events—perhaps exogenous changes that plausibly affect cost, but in amounts that may be difficult to determine with much precision—firms may to an extent render detection through this approach more difficult. Observe that more complex behaviors of this sort might require more elaborate (and thus more explicit) communications, which themselves may leave more traces, and more internal evidence of the sort considered in subsection B.1 may also be generated, so this strategy is not without risk to the firms employing it. To the extent that such camouflaging strategies are net helpful to firms, an implication is that interdependent behavior involving little or no direct interfirm communication may be easier to distinguish from competitive behavior than is interdependent behavior implemented after more explicit interchanges.

In maintaining oligopolistic prices, it may appear that firms' pricing patterns would differ little from those of aggressive competitors: in the absence of any exogenous changes, prices would remain constant, and in response to changes, prices would change accordingly, rising, for example, if cost or demand increases. As will be discussed in subsection (b), economic theory and econometric methods can, in principle and sometimes in practice, distinguish price responses to changes in cost or demand as a function of the nature of competitive interaction among firms. These analyses typically imagine immediate responses by oligopolists. For the moment, consider a less subtle point that may more readily be detectable in pricing patterns, namely, the possible lack of response or delayed or muted response of oligopolists, in contrast to competitors, to exogenous changes in market conditions.

The reasonably familiar point is that oligopoly prices, and perhaps also market shares, will tend to be stickier over time than are those of competitors.⁷⁹ The main reason is that, coordination being difficult, frequent fine-tun-

⁷⁸ See, e.g., Joseph E. Harrington, Jr., *Optimal Cartel Pricing in the Presence of an Antitrust Authority*, 46 INT'L ECON. REV. 145 (2005); Joseph E. Harrington, Jr. & Joe Chen, *Cartel Pricing Dynamics with Cost Variability and Endogenous Buyer Detection*, 24 INT'L J. INDUS. ORG. 1185 (2006); cf. Robert C. Marshall, Leslie M. Marx & Matthew E. Raiff, *Cartel Price Announcements: The Vitamins Industry*, 26 INT'L J. INDUS. ORG. 762 (2008) (finding that the vitamins cartel adjusted its pattern of price changes to minimize buyer resistance).

⁷⁹ See, e.g., Rosa M. Abrantes-Metz, Luke M. Froeb, John F. Geweke & Christopher T. Taylor, *A Variance Screen for Collusion*, 24 INT'L J. INDUS. ORG. 467 (2006) (using a variance screen to identify price coordination by gasoline stations); Dennis W. Carlton, *The Rigidity of Prices*, 76 AM. ECON. REV. 637, 655 (1986) (presenting empirical evidence that price rigidity is greater in more concentrated industries); Levenstein, *supra* note 21, at 122 (finding that prices

ing may be avoided. If accomplished through explicit negotiations, the likelihood of detection will rise. If accomplished through price signaling behavior, there may be misunderstandings. For example, a firm decreasing price in response to a perceived softening in demand or decline in costs may mistakenly be viewed by others as a cheater, triggering a price war.⁸⁰ Consequently, firms may be more reluctant to respond promptly to changes. In contrast, competitors do tend to respond quickly, and if changes are frequent, often.

This difference is not always easy to ascertain because there are other explanations for sticky prices, such as menu costs (referring to the cost of changing prices *per se*). But in some settings, such costs may be negligible or insufficient to explain the extent of stickiness. Likewise, menu costs tend to apply symmetrically to price increases and decreases, and their existence may lead firms to change prices at somewhat different times, whereas sluggishness due to problems in coordinating oligopoly pricing does not have these features. Additionally, oligopolists often tend to stabilize cooperating firms' relative market shares for fear that, if they do not, firms with growing shares will be seen as guilty of secret price cutting or other forms of cheating. Competitors also may have stable shares, but, depending on cost structures, they may be less stable.

Sudden and sharp price reductions are as suspicious as sudden and sharp price increases, again, in the absence of corresponding changes in cost or demand. Oligopolists do not ordinarily wish to drop their price, but such is sometimes unavoidable. Price wars arise to punish cheaters or, as explained in section II.A, when firms experience a loss of customers and thus must act on the assumption that cheating occurred even though a decline in demand that is not yet evident might have been the cause. If prices fall substantially, without any exogenous change of corresponding magnitude, it follows that the preexisting price involved oligopolistic price elevation at least to that extent. Note further that price wars that do not permanently end oligopoly pricing may well be followed by price increases, which themselves may be detected in the manner described previously.

Using price wars to detect oligopolistic price coordination raises an important danger that is not present when targeting price increases. If the direct

changed much less frequently during cooperative periods in the bromine cartel); Martin Pesendorfer, *A Study of Collusion in First-Price Auctions*, 67 REV. ECON. STUD. 381 (2000) (finding more stable market shares in bid-rigging cartel that did not use side-payments compared to one that did).

⁸⁰ This notion may also help to explain the sometimes-observed phenomenon that prices rise more promptly following cost increases than they fall following cost decreases. *See, e.g.*, Felipe Balmaceda & Paula Soruco, *Asymmetric Dynamic Pricing in a Local Gasoline Retail Market*, 56 J. INDUS. ECON. 629 (2008).

effect of penalizing price increases is to discourage them, the result is precisely the deterrent effect that is desired. However, if penalizing price decreases discourages them, we may be concerned that the consequence will be higher, not lower prices. This depiction of a possible problem does, fortunately, involve a crucial omission: oligopolists need price wars to be feasible in order to establish and maintain supracompetitive prices in the first place, as section II.A explains. That is, the prospect of punishment is a necessary condition to successful interdependent oligopoly pricing. Hence, if enforcement makes price wars difficult, oligopoly pricing may be discouraged after all.

Although this response provides a more complete picture, the problem is still more complicated. For example, if price wars are rendered more costly, it may be they would still be employed but with a less sensitive trigger, in which case society may suffer from oligopoly pricing in any event but receive somewhat less frequent relief because price wars are rarer or shorter. (A shorter punishment phase may accomplish similar discouragement of cheating because the prospective cheater will reckon that it suffers not only from the periods of reduced profit but also the possibility of legal sanctions.) Clearly, the issue is more complex than it may initially appear, and how useful detection through price wars turns out to be may depend on subtle factors.⁸¹

Nevertheless, this conundrum might be viewed as presenting an opportunity rather than an obstacle. For oligopoly pricing detected through price wars, one might provide leniency, immunity, or even some sort of reward to the initial cheater, preserving full legal sanctions for those firms administering the punishment in response to the defection. Some interesting arguments and proposals along these lines have been offered.⁸² For present purposes, it seems fair to say that the optimal legal approach is not obvious, but it seems likely that a

⁸¹ See, e.g., Philippe Cyrenne, *On Antitrust Enforcement and the Deterrence of Collusive Behavior*, 14 REV. INDUS. ORG. 257 (1999) (analyzing a model of the phenomenon in which the penalty does not depend on the extent of oligopoly pricing and is assumed to be insufficient ever to deter oligopoly pricing or the price wars required to sustain it, and finding that in this case the penalty will not affect price or output and will only shorten the duration of price wars).

⁸² See, e.g., Cécile Aubert, Patrick Rey & William E. Kovacic, *The Impact of Leniency and Whistle-Blowing Programs on Cartels*, 24 INT'L J. INDUS. ORG. 1241 (2006) (favoring rewards over mere leniency to help break otherwise stable cooperative behavior, and emphasizing how this tactic can usefully help create principal-agent conflicts in firms); Joseph E. Harrington, Jr., *Optimal Corporate Leniency Programs*, 56 J. INDUS. ECON. 215 (2008); Nathan H. Miller, *Strategic Leniency and Cartel Enforcement*, 99 AM. ECON. REV. 750 (2009) (providing empirical evidence that introduction of a leniency program in 1993 in the United States enhanced deterrence); Massimo Motta & Michele Polo, *Leniency Programs and Cartel Prosecution*, 21 INT'L J. INDUS. ORG. 347 (2003); Giancarlo Spagnolo, *Leniency and Whistleblowers in Antitrust*, in HANDBOOK OF ANTITRUST ECONOMICS 259 (Paolo Buccirossi ed., 2008). Two interesting proposals that are related in spirit have been advanced. See William Bishop, *Oligopoly Pricing: A Proposal*, 28 ANTITRUST BULL. 311 (1983); Guy Sagi, *The Oligopolistic Pricing Problem: A Suggested Price Freeze Remedy*, 2008 COLUM. BUS. L. REV. 269.

thoughtful enforcement strategy can be devised that uses detection through behavior involved in price wars as a weapon against oligopoly pricing.

b. Price Elevation⁸³

Most of the discussion in subsection (a) concerns price elevation, but the primary means of inferring the existence of successful oligopolistic coordination is to examine major price changes—when substantial price elevations come into existence or evaporate quickly—as well as nonresponsiveness to changes in conditions in the interim. This subsection looks at means of inferring whether existing prices are elevated by examining the prices themselves, typically under the assumption that they do reflect and respond to existing market circumstances.

A logically straightforward way to determine whether price exceeds a competitive level, marginal cost, is to compare price and marginal cost directly.⁸⁴ While price is often easy to determine,⁸⁵ marginal cost may be quite difficult to measure in many settings. This challenge is familiar in competition law enforcement, often being confronted when measuring market power⁸⁶ and also at issue in predatory pricing disputes⁸⁷ and in some other contexts.

Leading difficulties involve the determination of which costs are variable over what time period and the allocation of joint or common costs to particular products.⁸⁸ It also may often be easier to measure average cost, even though marginal cost is more relevant for determining whether pricing is

⁸³ Keep in mind that assessment of elevation per se does not distinguish coordinated from unilateral elevation, as elaborated in section V.C.

⁸⁴ Closely related is to measure profits to see whether they are above a competitive (risk-adjusted) return on capital. In addition to difficulties similar to those associated with measuring marginal cost—see, e.g., Franklin M. Fisher & John J. McGowan, *On the Misuse of Accounting Rates of Return to Infer Monopoly Profits*, 73 AM. ECON. REV. 82 (1983); Franklin M. Fisher, *On the Misuse of the Profits-Sales Ratio to Infer Monopoly Power*, 18 RAND J. ECON. 384 (1987)—profits reflect average cost, which the text to follow explains may provide a misleading indicator of supracompetitive pricing in some settings.

⁸⁵ There may be complications when goods are bundled or customized and also limits of data, such as when many transactions are privately negotiated at off-list prices, although the latter information may well be available to a competition authority even if not to academic researchers. A different challenge is that, when products are differentiated, there is no single price (and no single marginal cost), but rather one for each product.

⁸⁶ If measuring *existing* (that is, exercised) market power, the problem is tantamount to measuring price-cost margins. See, e.g., Louis Kaplow, *Why (Ever) Define Markets?*, 124 HARV. L. REV. 437, 444–47 (2010). Measuring unexercised or prospective market power may be more difficult.

⁸⁷ It is common to inquire whether price is below cost (although the notion of cost is not always defined). See, e.g., *Brooke Grp. Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209 (1993). Note that if true marginal cost is above measured cost, which as the text to follow indicates is the more likely error, predation tests are too lax but determinations of the existence of oligopolistic price elevation are too strict.

⁸⁸ See, e.g., Kaplow & Shapiro, *supra* note 13, at 1087–88.

supracompetitive. An example of where problems may arise is the case in which a firm's marginal cost is low through much of its output range but rises rapidly as production approaches full capacity. If the firm is producing near capacity and pricing near marginal cost, its price may be significantly above average cost and thus may appear to be elevated well above marginal cost. Thus, when firms produce close to capacity, one must be particularly cautious about making inferences of supracompetitive pricing.⁸⁹ Indeed, in the short run, firms at capacity could not behave so as to reduce industry price significantly in any event.⁹⁰

Underestimating marginal cost and thus producing false positives can occur as a consequence of too readily excluding costs as fixed that are really variable (perhaps equipment could be rented or sold), ignoring common costs (which may actually be raised on account of the need to support additional quantity of the product in question), or failing to recognize that marginal costs may rise steeply when close to capacity, as just described. Responses would include some combination of erring in the opposite direction (treating more costs as variable costs of the product in question, for example) and setting a higher threshold, that is, requiring demonstration of a more significant elevation of price above measured marginal cost. Whether this direct approach to measuring the gap between price and marginal cost will often be useful is hard to say. For some products, presumably, marginal cost will be fairly easy to measure or the magnitude of elevation will be sufficiently great that there is little doubt of supracompetitive pricing. Often, significant uncertainty will remain, and inquiries of this sort may provide little illumination. In all cases, it will be useful to look to other indicators of price elevation as well.

Another approach relies on comparing prices for the same product sold to different purchasers. Prices might be compared across markets; notably, geographical price differences may indicate supracompetitive pricing in regions with higher prices.⁹¹ The logic is that price is at least as high as marginal cost in all markets, so higher prices must to that extent be supracompetitive. The limitation to such inferences is that markets may exhibit cost differences as

⁸⁹ Compare the discussion in subsection II.C.2 of the case in which one or more firms have costs below those of others, at least over much of the production range, but nevertheless charge a price equal to the higher marginal cost at the actual level of output. It was suggested that looking to industry marginal cost, by examining competitors, was appropriate. As noted previously, an important caveat is that, if a group of leading firms significantly elevates price, there may well be a competitive fringe, with different products or technology, that is selling at marginal cost, and this fact would not negate the existence of oligopolistic price elevation.

⁹⁰ In the long run, if average cost is far below marginal cost and price, one would expect the firm to expand production capacity, competitors to expand, new entry to take place, or some combination—that is, unless firms are coordinating their capacity decisions.

⁹¹ It is also sometimes useful to compare prices across products that may have similar cost structures but face different degrees of competition.

well. Indeed, the prices of many competitively supplied products differ across regions. Accordingly, it is necessary to establish that cost differences are relatively small or to control for cost differences when making comparisons. The latter recreates the prior problem of measuring marginal cost directly, but the problem may be less severe since attention is limited to factors that make marginal cost for a given product differ across regions.⁹²

Similarly, different prices charged to different customers in a single market—that is, the presence of price discrimination—imply the existence of market power.⁹³ The lower price must be at least as high as marginal cost, for otherwise those units would not be sold, so it follows that price elevation with regard to any higher prices must at least equal the price-cost gap—although there are complications, including that not all differential pricing constitutes price discrimination in the sense just described.⁹⁴ Here, confounding cost dif-

⁹² This method of estimating price elevation was employed in the FTC's challenge to the merger of Staples and Office Depot. *See, e.g.*, Orley Ashenfelter, David Ashmore, Jonathan B. Baker, Suzanne Gleason & Daniel S. Hosken, *Empirical Methods in Merger Analysis: Econometric Analysis of Pricing in FTC v. Staples*, 13 INT'L J. ECON. & BUS. 265 (2006). Note that cross-market techniques can be combined with others. For example, if there is a common cost shock (see the discussion in the text to follow) but prices react differently in different markets, inferences about noncompetitive pricing may be possible.

⁹³ "It is well known that price discrimination is only feasible under certain conditions [including that] firms have short-run market power. . . ." Lars A. Stole, *Price Discrimination and Competition*, in 3 HANDBOOK OF INDUSTRIAL ORGANIZATION 2221, 2226 (Mark Armstrong & Robert H. Porter eds., 2007). "[W]hen markets are perfectly competitive and firms have neither short-run nor long-run market power, the law of one price prevails and price discrimination cannot exist." *Id.* at 2224; *see also id.* n.2 ("It is straightforward to construct models of price discrimination in competitive markets without entry barriers in which firms lack long-run market power . . . providing that there is some source of short-run market power that allows price to remain above marginal cost . . .").

⁹⁴ This logic applies not only to posted price differentials but also to unposted prices, including prices determined through buyers' haggling. In such contexts, it is often overlooked that, if some buyers are thus able to obtain significant discounts, other buyers are at least to that extent paying a price in excess of marginal cost. Similar analysis applies when some buyers obtain nonprice concessions. (Note that offering nonprice concessions that are less valuable to buyers than the cost of providing direct kickbacks may indicate oligopoly pricing, as sellers attempt to circumvent the agreed price without visibly cheating.) Some variation in the price paid by different buyers may, however, be explained by buyers' imperfect information about prices rather than, or in addition to, oligopolistic price elevation. *See, e.g.*, Kenneth Burdett & Kenneth L. Judd, *Equilibrium Price Dispersion*, 51 ECONOMETRICA 955 (1983); Steven Salop & Joseph Stiglitz, *Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion*, 44 REV. ECON. STUD. 493 (1977); Yuval Shilony, *Mixed Pricing in Oligopoly*, 14 J. ECON. THEORY 373 (1977); Hal R. Varian, *A Model of Sales*, 70 AM. ECON. REV. 651 (1980); Louis L. Wilde & Alan Schwartz, *Equilibrium Comparison Shopping*, 46 REV. ECON. STUD. 543 (1979). Indeed, if oligopolistic pricing is highly successful, such as when a cartel succeeds almost completely in preventing secret discounts, price dispersion may be smaller than under more competitive conditions with imperfectly informed purchasers. *See* John M. Connor, *Collusion and Price Dispersion*, 12 APPLIED ECON. LETTERS 335 (2005); *see also* Severin Borenstein & Nancy L. Rose, *Competition and Price Dispersion in the U.S. Airline Industry*, 102 J. POL. ECON. 653 (1994) (finding greater price dispersion on more competitive routes, grounding the explanation in a model with differentiated products). *But see* Kristopher S. Gerardi & Adam Hale Shapiro, *Does*

ferences may exist but would be more circumscribed, as they are limited to differential costs of serving different customers, which often would be negligible. Note that if the existence of price discrimination was known by firms to be used by competition law enforcers as an indicator of oligopolistic price coordination, price discrimination may be discouraged without deterring (average) price elevation. Static, allocative efficiency may rise or fall as a consequence.⁹⁵ Producers' surplus would fall, which may often but not always be socially advantageous as a consequence of the reduced incentive for entry and other rent-seeking investments, as discussed in subsection II.B.2.

A different method of identifying successful oligopolistic coordination is to examine whether ordinary pricing behavior (that is, aside from episodes of commencing or terminating coordination) responds to changes in demand and cost in the same manner as would pricing by competitors or somewhat (or entirely) like pricing by a monopolist.⁹⁶ Such analysis sets aside the point in subsection (a) that, by comparison to monopolists, oligopoly pricing may be sticky, which itself differentiates oligopolistic pricing patterns from competitive ones. Indeed, a substantial body of econometric research since the 1980s is designed to measure the exercise of market power using such a strategy.⁹⁷

To illustrate some of the intuition behind these methods, consider an upward shift in demand in an industry with common, constant marginal cost. Perfect competitors would expand output, but price would not rise (it would continue to equal the preexisting marginal cost). A monopolist generally

Competition Reduce Price Dispersion? New Evidence from the Airline Industry, 117 J. POL. ECON. 1 (2009) (finding that increased competition reduces price dispersion, and showing that Borenstein and Rose's contrary result was due to omitted-variable bias); Alberto A. Gaggero & Claudio A. Piga, *Airline Market Power and Intertemporal Price Dispersion* (Loughborough University Dept. of Econ. WP 2009-10, 2009) (finding that competition reduces fare dispersion on routes between Ireland and the United Kingdom). Yet another possible source of price dispersion arises when firms set prices in advance (that cannot later be modified) when the magnitude of future demand is uncertain. See James D. Dana, Jr., *Equilibrium Price Dispersion Under Demand Uncertainty: The Roles of Costly Capacity and Market Structure*, 30 RAND J. ECON. 632 (1999); Benjamin Eden, *Marginal Cost Pricing When Spot Markets Are Complete*, 98 J. POL. ECON. 1293 (1990).

⁹⁵ See, e.g., Hal R. Varian, *Price Discrimination*, in 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 597, 619-22 (Richard Schmalensee & Robert D. Willig eds., 1989).

⁹⁶ See, e.g., Bresnahan, *supra* note 56, at 1012; Bresnahan, *supra* note 70, at 71.

⁹⁷ In addition to the sources cited in note 70, see, for example, Timothy F. Bresnahan, *The Oligopoly Solution Concept is Identified*, 10 ECON. LETTERS 87 (1982); Hall, *supra* note 58; Lawrence J. Lau, *On Identifying the Degree of Competitiveness from Industry Price and Output Data*, 10 ECON. LETTERS 93 (1982); see also Kenneth S. Corts, *Conduct Parameters and the Measurement of Market Power*, 88 J. ECONOMETRICS 227 (1999) (arguing that the common method of estimating a conduct parameter, which indicates where an industry falls between perfect competition and monopoly, can be misleading in certain settings); Aviv Nevo, *Identification of the Oligopoly Solution Concept in a Differentiated-Products Industry*, 59 ECON. LETTERS 391 (1998) (explaining that the conduct parameter is difficult to identify in differentiated products industries, although it is possible to compare particular models of industry behavior).

would increase price—although not necessarily, because it would depend on how the elasticity of demand changed as well, but this would be a typical response, and demand elasticity would also be estimated. Therefore, if price rises in an oligopolistic industry with constant marginal cost, this would be evidence of oligopolistic price coordination. Cost shifts allow similar analysis. With constant marginal costs, a common cost increase will be fully passed on by perfect competitors, and idiosyncratic cost changes will cause large shifts in production across competitors. But if there is coordinated oligopoly pricing, the passing on of a common cost increase will depend on the curvature of demand, which would be estimated, and firm-specific cost shifts would involve less reallocation of production.

More broadly, econometric techniques entail simultaneous estimation of firms' costs and demand. Such methods do not assume that firms' marginal costs are directly measurable; these costs are implicitly estimated from the data.⁹⁸ This work represents a significant advance in market power measurement, but it is hardly a panacea. Such techniques often require strong assumptions about the structure of demand and cost that it may not be possible to test directly.⁹⁹ The central question for present purposes concerns the reliability of such techniques given the data that would be available when investigating a particular group of firms. These methods are increasingly used to predict effects of horizontal mergers, where agencies are often attempting to determine whether a price elevation of at least a few percent is likely to result.¹⁰⁰ Allowing for some (perhaps significant) margin of error, a similar approach could be employed to detect oligopolistic price elevation.¹⁰¹

Entry also provides a window on the existence of successful oligopolistic coordination. First, entry itself may be a clue. For example, if incumbent firms have unused capacity yet new firms enter the industry (or fringe firms expand), it is plausible that prices are elevated. Recall the discussion in subsection II.B.2.a of how oligopoly pricing induces (often socially excessive) entry. Second, incumbents' response to entry may be revealing. If pre-entry behavior

⁹⁸ See, e.g., PERLOFF ET AL., *supra* note 21, at 5, 42; Bresnahan, *supra* note 56, at 1012.

⁹⁹ See, e.g., PERLOFF ET AL., *supra* note 21, at 42, 70–71, 91.

¹⁰⁰ See, e.g., Jonathan B. Baker & Timothy F. Bresnahan, *The Gains from Merger or Collusion in Product-Differentiated Industries*, 33 J. INDUS. ECON. 427 (1985); Roy J. Epstein & Daniel L. Rubinfeld, *Merger Simulation: A Simplified Approach with New Applications*, 69 ANTITRUST L.J. 883 (2001); Roy J. Epstein & Daniel L. Rubinfeld, *Merger Simulation with Brand-level Margin Data: Extending PCAIDS with Nests*, 4 ADVANCES ECON. ANALYSIS & POL'Y, Mar. 2004, art. 9, 1, <http://www.bepress.com/cgi/viewcontent.cgi?article=1212&context=bejeap>; Gregory J. Werden & Luke M. Froeb, *Unilateral Competitive Effects of Horizontal Mergers*, in HANDBOOK OF ANTITRUST ECONOMICS 43 (Paolo Buccirossi ed., 2008).

¹⁰¹ As mentioned previously with regard to direct measurement of marginal cost, one might require demonstration of larger price elevations the more uncertain the method of estimation appears to be.

is competitive, a new entrant will draw sales from other firms, prices will tend to fall, and the other firms' quantities will fall. The same may occur if the industry is oligopolistic, but it is also possible in the latter case that incumbents' output would stay constant or even rise, depending on how the operative understanding about industry price is affected by entry. A move to a substantially more competitive environment, indicated by a steep decline in price, would evidence preexisting supracompetitive pricing much as would the outbreak of a price war.¹⁰²

The foregoing list is hardly exhaustive.¹⁰³ It is clear that successful oligopolistic coordination resulting in price elevation has a number of features that distinguish it from ordinary competition and that there exist techniques for identifying these differences. On the other hand, these differences will not always be apparent, and sometimes alternative explanations will be available that themselves may be difficult to assess. Accordingly, it is important also to consider other sources of information that bear on whether oligopolistic coordination is taking place.

c. Facilitating Practices

Facilitating practices are acts that make it easier to engage in oligopolistic coordination. Typically, they address the challenges of determining a mutually acceptable oligopoly price and deterring cheating through the prospect that defection will be detected and punished quickly. Because coordinated oligopoly pricing is often difficult to accomplish, facilitating practices may be important and thus have long been a focus of policy analysis of the oligopoly problem.¹⁰⁴ Indeed, direct interfirm communications, taken by many to be determinative of what constitutes a price-fixing agreement, can also be viewed as facilitating practices.¹⁰⁵

Facilitating practices may be relevant in two ways. First, their use provides a basis for inferring the existence of oligopolistic coordination. This inference is sensible when there exists no other explanation for the practice or when

¹⁰² See, e.g., Stephen A. Nye, *Can Conduct Oriented Enforcement Inhibit Conscious Parallelism?*, 44 ANTITRUST L.J. 206, 220 (1975).

¹⁰³ See, e.g., sources cited *supra* note 70.

¹⁰⁴ See, e.g., POSNER, ANTITRUST, *supra* note 7, at 86–87, 88–89, 91–93; Paolo Buccirossi, *Facilitating Practices*, in HANDBOOK OF ANTITRUST ECONOMICS 305 (Paolo Buccirossi ed., 2008); see also CARL KAYSER & DONALD F. TURNER, ANTITRUST POLICY: AN ECONOMIC AND LEGAL ANALYSIS 150–52 (1959) (offering an early discussion).

¹⁰⁵ See Kaplow, *supra* note 3, § II.B.2 (exploring this perspective on interfirm communications); Kai-Uwe Kühn, *Fighting Collusion by Regulating Communication Between Firms*, 16 ECON. POL'Y 169, 180 (2001) (“Enforcing primarily against such ‘secondary’ activities or ‘facilitating practices’ may then be a much more powerful enforcement strategy than an attempt to infer behaviour directly. The activities that seem to be most closely associated with collusion tend to centre around communication.”).

other explanations can be shown to be implausible in a particular case. In contrast, practices that may facilitate oligopolistic interdependence but would likely be employed regardless are not directly probative.

Second, facilitating practices may themselves be made a basis for liability.¹⁰⁶ One rationale would be that it often makes sense to punish attempts even when they are unsuccessful (a subject explored in section V.B). Another of particular importance in the present context is that punishing attempts is good policy when it is difficult or costly to determine whether they are successful. As when using facilitating practices as a diagnostic for the existence of coordinated oligopoly pricing, one must determine whether there are other plausible explanations for the practice. Furthermore, if there are, any benefits must be weighed against the competitive risks. The remainder of this subsection considers a number of facilitating practices without much being concerned with whether any given practice should be viewed primarily as a symptom of oligopoly pricing or also as an act that might be directly condemned.

An important way to facilitate oligopolistic interdependence is through improved mutual understanding of relevant circumstances. Most obvious are direct interfirm communications about price, which will be explored further in subsection B.2. Advance price announcements that on their face are directed at the world are also important. Firms hoping to increase the industry price may wish to convey their beliefs without taking the risk of raising prices unilaterally, possibly alienating customers and losing interim sales without being confident whether and how quickly rivals will follow. Perhaps more important, if a firm believes that the industry price should be reduced because it perceives falling demand or because of cost reductions, it faces the danger that a unilateral price cut will be perceived as a defection, thereby triggering a price war.

Both concerns can readily be overstated and may be unimportant in some markets. For example, price changes might be initiated during periods when sales are unlikely, when the moves will be quickly observed by competitors, and where they can be rescinded almost immediately if not promptly matched. Nevertheless, advance price announcements, which may be followed by rivals' responsive announcements and further modifications by the initiator, in as many rounds as necessary, may reduce risks attendant with changing

¹⁰⁶ If liability is to be found under Sherman Act Section 1 or EU Article 101 (formerly 81), there must be some sort of agreement, and as Kaplow, *supra* note 3, § II.B.2, indicates, commentators are often inconsistent on this point—in that, when finding an agreement on facilitating practices to be present, they implicitly employ a notion of agreement that contradicts the view of the law that they advance with regard to price fixing.

prices, thereby facilitating oligopoly pricing.¹⁰⁷ Such announcements could be quite detailed, indicating price moves, dates, and so forth, or they may be more vague, such as a statement that the firm believes that demand has fallen, after which its price decrease may be more likely to be interpreted as an industry-wide invitation rather than a unilateral defection.

Firms might also share information about prices and other matters through trade associations. Possibilities include broad discussions about industry trends as well as precise, exhaustive information about actual prices, such as by having the association audit invoices, reporting transaction prices to other members.¹⁰⁸ In this regard, it is interesting to note that the enactments of anti-price-fixing regimes in the United States and in Britain were each followed by a surge in trade association activity.¹⁰⁹ Furthermore, in some notorious price-fixing cartels, secret discussions were held in conjunction with trade association gatherings, which provided a cover for the relevant players regularly happening to be at the same hotels on the same dates.¹¹⁰ Also, in those prosecuted price-fixing cases that do not involve a fairly small number of firms, use of trade associations seems common.¹¹¹

¹⁰⁷ It is usually overlooked that, if advance price announcements are permissible, they may serve as cover for secret interfirm communications that are outlawed. Suppose, for example, that firms agree at a clandestine meeting to increase price to 100 at the stroke of midnight a week later. Such simultaneous activity by itself may be highly suspicious, serving as an indicator of the meeting. However, the firms could further plan that, shortly after the meeting concludes, one of them will publicly announce a supposed unilateral intention to start pricing at 100 at the prescribed time. Others could publicly respond, indicating their similar (also purportedly unilateral) intentions, or they could simply act likewise and, if investigated, point to the first announcer's statement. If successful oligopolistic coordination subjects firms to sanctions, such action may tend to trigger liability. But if it must be proved that there were secret meetings, the inference would be made more difficult as a consequence of the intervening public statements.

¹⁰⁸ See, e.g., *Am. Column & Lumber Co. v. United States*, 257 U.S. 377 (1921). A rarely mentioned but possibly important alternative means of obtaining information about rivals' pricing, including cheating, is for each firm to investigate rivals' prices carefully and continuously, including possibly through the use of secret buying agents. See MICHAEL E. PORTER, *COMPETITIVE STRATEGY* 96 (1980).

¹⁰⁹ See, e.g., J.B. Heath, *Some Economic Consequences*, 70 *ECON. J.* 474, 475 (1960) (more than 150 price exchange agreements arose within a few years of Britain declaring overt price fixing to be illegal, and a similar rise in trade associations occurred in the United States in the decades following passage of the Sherman Act). In Britain, it appears that information sharing arrangements were often successful, especially before they were subject to stricter scrutiny. See DENNIS SWANN, DENIS P. O'BRIEN, W. PETER J. MAUNDER & W. STEWART HOWE, *COMPETITION IN BRITISH INDUSTRY: RESTRICTIVE PRACTICES LEGISLATION IN THEORY AND PRACTICE* 158-63 (1974).

¹¹⁰ See CONNOR, *GLOBAL PRICE FIXING*, *supra* note 59, at 11, 143-44, 295-96.

¹¹¹ See Arthur G. Fraas & Douglas F. Greer, *Market Structure and Price Collusion: An Empirical Analysis*, 26 *J. INDUS. ECON.* 21, 39-42 (1977); Hay & Kelley, *supra* note 59, at 21-22; Margaret C. Levenstein & Valerie Y. Suslow, *International Cartels*, in 2 *ISSUES IN COMPETITION LAW AND POLICY* 1107, 1123 (Wayne Dale Collins ed., 2008).

A major reservation regarding such information-sharing activities is that they may also have redeeming virtues. These benefits might provide an innocent explanation, thereby negating the inference of oligopolistic coordination, and also indicate that it would be socially costly to prohibit or chill the activities. While these points are important, they are subject to a significant qualification that is not always appreciated: in determining whether a possibly ambiguous facilitating practice should be viewed positively or negatively, it is necessary to consider the cooperating firms' incentives. That is, it is necessary to consider whether, under relevant alternative hypotheses, firms in fact gain from sharing the information. As a general matter, one may note that sellers' interests will often be opposed to buyers' and society's interests, for sellers are better off the more successful are their attempts at price elevation.¹¹²

For example, regarding the common suggestion that firms engage in various open pricing practices in order to help buyers, it must be asked why they would have an incentive to do so. If firms are colluding, making it easier for buyers to find discounts—i.e., firms that are cheaters—will tend to undermine oligopoly pricing. In contrast, making it easier for sellers to identify cheaters will make punishment swift and thereby deter defection. Hence, if firms in an industry conducive to coordinated price elevation act to make pricing more open, the logical inference would be that the latter, socially undesirable effect, is greater. If firms secretly share the information, they avoid any tradeoff.¹¹³ However, if that is difficult (at least without being caught by enforcers) and if public sharing is viewed as innocuous, they may well prefer public sharing to no sharing. An individual buyer in the short run will be happy to learn where

¹¹² Regarding sellers' motives, it is noteworthy that, in Michael Porter's text on competitive strategy, giving buyers advance notice is not on his long list of strategic reasons for advance announcements, which is dominated by concerns with communication between competitors. See PORTER, *supra* note 108, at 76–80. But see DAVID DRANOVE & SONIA MARCIANO, KELLOGG ON STRATEGY: CONCEPTS, TOOLS, AND FRAMEWORKS FOR PRACTITIONERS 139 (2005) (“[p]reannouncements often help customers plan their future purchases, such as [by] allowing contractors to better budget”). Porter emphasizes that making statements publicly can be a way to enhance the effective commitment. See PORTER, *supra* note 108, at 103–04; see also THOMAS C. SCHELLING, THE STRATEGY OF CONFLICT 29–30 (1960) (discussing how publicity rather than secrecy is necessary for reputation to be effective). Another view is that firms may have their cake and eat it too by making statements that are public to each other yet less likely to be noticed by buyers. See Oliver P. Heil, George S. Day & David J. Reibstein, *Signaling to Competitors*, in WHARTON ON DYNAMIC COMPETITIVE STRATEGY 277, 281 (George S. Day & David J. Reibstein with Robert E. Gunther eds., 1997) (“Announcing plans is particularly effective when it is conducted outside the public eye (in a trade journal, for example). This way, the announcements will not delay consumers' purchases or damage consumers' willingness to pay.”). Another forum might be meetings with stock analysts that cover an industry, the subject of one of the FTC's recent actions against Valassis Communications for inviting its competitor to collude. See Press Release, Fed. Trade Comm'n, FTC Consent Order Protects Competition in the Market for Free-standing Newspaper Inserts, <http://www.ftc.gov/opa/2006/03/valassis.shtm> (Mar. 14, 2006).

¹¹³ This point is familiar from situations like that in *United States v. Container Corp. of Am.*, 393 U.S. 333 (1969), in which firms verified with competitors various assertions by buyers that they had been offered lower prices.

lower prices might be obtained. But, in the long run, buyers as a whole are worse off if the result is more elevated prices with fewer, if any, discounts.¹¹⁴

The point in this example about firms' incentives is reinforced by considering a more (but not perfectly) competitive setting. When it is costly for buyers to learn about the prices offered by different sellers, there tends to be price dispersion, with buyers engaging in costly search activity to discover lower prices but also finding it optimal not to search indefinitely, often settling for prices above the lowest that may be available.¹¹⁵ Likewise, unsophisticated customers may pay higher prices than others do. In this setting, firms developing a mechanism to make price information readily available to buyers may be socially desirable. The problem is that the firms' joint incentive may be the opposite. Hence, we should be skeptical of group information dissemination that is claimed to assist buyers in generally obtaining lower prices from the firms involved.¹¹⁶

Now consider what may appear to be a more innocuous situation that does not involve pricing information: firms at trade association meetings share information about future demand because demand is uncertain and difficult to predict.¹¹⁷ If all firms can better plan capacity decisions, inventories, and the like, production will tend to be more efficient, a social gain. An important question, however, is whether individual firms gain, and what other consequences may flow from such sharing. Recall from section II.A that uncertain demand is a significant threat to oligopoly pricing in settings where transaction prices are secret, which is typical in many industries selling intermediate goods to large buyers, where sales are individually negotiated. In that case, firms' inability to distinguish sales lost to demand fluctuations from those lost to cheating can trigger price wars and thus may constitute the greatest threat to the sustainability of coordinated oligopoly pricing. Accordingly, in such

¹¹⁴ A further complication regarding buyers' incentives is that, if they are intermediate purchasers rather than final, individual consumers, they may suffer little from common price elevation because their competitors are in the same situation. In contrast, a buyer who is able to obtain a concession unavailable to others benefits, and one unable to obtain concessions received by others suffers.

¹¹⁵ See *supra* note 94, including the sources cited therein.

¹¹⁶ This skepticism does not mean that there cannot be exceptions, particularly regarding particular firms (as distinguished from joint activity). Firms introducing a new product or operating in a new location may wish to attract customers and may enjoy scale economies in advertising; note that, in this case, the other firms (selling existing substitutes or operating in traditional locations) would not wish to join in their rivals' efforts. Likewise, certain firms may wish to advertise prices to draw business from incumbents, which themselves may attempt to suppress such efforts. See, e.g., Lee Benham, *The Effect of Advertising on the Price of Eyeglasses*, 15 J.L. & ECON. 337 (1972).

¹¹⁷ Exchange of cost information may be different. See, e.g., Carl Shapiro, *Exchange of Cost Information in Oligopoly*, 53 REV. ECON. STUD. 433 (1986) (finding, in a model with Cournot behavior, that oligopolists have an incentive to share cost information, which raises their profits, reduces consumer surplus, and increases total welfare).

industries there may be reason for society to fear information exchanges that reduce future demand uncertainty.¹¹⁸ Moreover, firms may have a strong incentive to share this information precisely when this threat to their elevated prices is present but not otherwise.¹¹⁹ It might be socially advantageous for more information to be shared if it is possible independently to prevent it from being employed to facilitate price elevation, but that raises different questions.¹²⁰

Another set of facilitating practices is addressed to product standardization. The preliminary analysis of coordinated oligopoly behavior in section II.A imagines that firms produce a single, homogeneous product, and frequent references throughout this article are made to the familiar idea that successful coordination is most likely when products are homogeneous or nearly so. Heterogeneity raises a number of challenges. First, it is more difficult to coordinate on an industry price. Indeed, if heterogeneity is substantial or even if goods differ only modestly but vary along a single dimension that consumers rank similarly, it may be necessary for an array of prices to be charged. Moreover, this set of prices needs to be changed over time as industry conditions evolve, resulting in a much more complex problem and one that may give rise to more disagreement than the adjustment of a single price for a homogeneous good.

Accordingly, standardization facilitates collusion. If firms could create or gravitate toward a common template indicating relative values, then perhaps they would only need to adopt—and, over time, adjust—a single index, with other prices following in line. To illustrate, it is often suggested that basing-point pricing facilitates collusion. The problem arises when transportation costs are nontrivial and rivals' plants are significantly dispersed. It might be

¹¹⁸ See, e.g., Levenstein & Suslow, *supra* note 111, at 1121–23 (describing efforts of international cartel members to avoid price wars by sharing information about market conditions, including through trade association activity). Likewise, inventories can be used strategically (for example, build-ups can enable more profitable defection), so firms desiring to facilitate oligopolistic price elevation may benefit from sharing inventory information and coordinating inventory decisions. See Julio J. Rotemberg & Garth Saloner, *The Cyclical Behavior of Strategic Inventories*, 104 Q.J. ECON. 73 (1989).

¹¹⁹ This result is obtained, for example, in the model in Richard N. Clarke, *Collusion and the Incentives for Information Sharing*, 14 BELL J. ECON. 383 (1983).

¹²⁰ See *id.* Note that if firms have incentives to share even if there is no price elevation, and expected sanctions for price elevation are sufficiently high in the industry in question, then it would not make sense to infer successful oligopolistic coordination from voluntary information sharing. And if firms do not have such incentives, it may be optimal for the government to provide for or subsidize information, but perhaps only if enforcement against price elevation is sufficiently strong. Consider that the Danish competition authority gathered and published prices of concrete, motivated by a belief in the desirability of price transparency, with the result that prices increased due to eased oligopolistic coordination. See Svend Albæk, Peter Møllgaard & Per B. Overgaard, *Government-Assisted Oligopoly Coordination? A Concrete Case*, 45 J. INDUS. ECON. 429 (1997).

difficult for all firms to agree on prices at each of many locations, but perhaps they could limit themselves to one or a few basing points, each with a single price, with prices at all other locations determined by adding some standardized freight rate to the price at the (nearest) basing point.¹²¹ In other settings, where heterogeneity concerns quality differences, firms might attempt to suppress information about this variation so they can utilize a uniform price. A firm with a slight quality advantage might go along because it prefers to sell at an oligopoly price rather than enjoying a somewhat higher market share at a more nearly competitive price.

A second problem is that one form of cheating on the oligopoly price consists of improving quality or offering additional services or better terms, while adhering to the nominal coordinated price. Oligopolists therefore wish to suppress all manner of quality competition as well.¹²² Note that, in all of these instances, a competitive firm that took its rivals' actions as given would have an incentive to differentiate its product, emphasize the advantages of its own wares over those of competitors, adjust its price to reflect differences in quality, and otherwise act in ways that divert business from rivals to itself.

A rather different type of facilitating practice is aimed at altering firms' direct payoffs from raising or cutting prices. One means is cross-ownership: if firm A owns a fraction of firm B (setting to the side mergers or other transfers of control), A has less incentive to cheat if the result will hurt B's profits.¹²³ Another mechanism is the use of side payments, for example, to share industry profits more equally when some firms need to make greater sacrifices or when it turns out *ex post* that one firm sold more than its allotted share of output (whether due to cheating or luck; rivals may be unable to tell which). Side payments are widely accepted as evidence of coordinated oligopolistic price elevation, for why else would a competitor make a payment to a rival for no consideration. A more subtle form of the practice involves cross-purchases. That is, a firm that sold more than its allotted share might buy from firms that sold less; if such purchases are at the elevated oligopoly price, compensation will have been accomplished.¹²⁴

A more direct way to reduce incentives to provide selective price reductions is through most-favored-customer clauses that bind a seller to give all of

¹²¹ See, e.g., SCHERER & ROSS, *supra* note 58, at 502–08.

¹²² See, e.g., *Catalano v. Target Sales*, 446 U.S. 643 (1980) (holding illegal an agreement not to offer credit to buyers).

¹²³ See, e.g., David Gilo, Yossi Moshe & Yossi Spiegel, *Partial Cross Ownership and Tacit Collusion*, 37 RAND J. ECON. 81 (2006); Robert J. Reynolds & Bruce R. Snapp, *The Competitive Effects of Partial Equity Interests and Joint Ventures*, 4 INT'L J. INDUS. ORG. 141 (1986).

¹²⁴ This practice has been employed by firms in some prosecuted cartels. See CONNOR, *GLOBAL PRICE FIXING*, *supra* note 59, at 143 (citric acid conspiracy); Harrington, *supra* note 59, at 57–62 (discussing forms of compensation, including buy-backs, in a number of conspiracies).

its customers retroactive price cuts if it is found to have sold to anyone else at a lower price.¹²⁵ For such clauses to be effective, there must be a sufficient chance that, at least eventually, other customers would detect the secret price cut. Additionally, it is necessary for firms to coordinate on the adoption of such restrictions, for, if some firms abstain, their incentive to cut price is increased because contractually obligated rivals may find punishment too costly.¹²⁶

It has also been suggested that an industry might employ resale price maintenance to facilitate collusion. When it is in force, upstream price cuts cannot be passed along to consumers, so they do not generate as much gain in market share. This view, however, requires qualification. Price cuts still induce retailers to steer customers, perhaps aggressively, toward the price-cutter's product. Moreover, since the retailer is prevented from passing any of the reduction to consumers, the upstream price cut may be less visible to rivals.

A different approach is to use meeting-competition clauses to reduce rivals' incentives to cut prices.¹²⁷ These might promise customers that the offeror will meet or undercut by some margin any lower price of a competitor. Such clauses may be effective particularly when customers have some switching costs: if they are offered a price reduction by a rival, they would prefer to stick with their existing supplier, using the rival's offer to extract a lower price. It follows that rivals would be less inclined to reduce prices in the first place. Note, however, that even without any price-matching promise, a customer can always convey the rival's price cut to its regular supplier, and if the price cut involves cheating, as we are currently supposing, that supplier would be inclined to respond even without being obligated to do so.

As suggested at the outset, it often will be ambiguous whether a posited facilitating practice is being employed (and how broadly in the industry), what are its actual effects, and therefore how strongly it supports an inference of oligopolistic price elevation or a conclusion that the practice is undesirable and thus should be prohibited. Some practices are more probative or problematic than are others, and for most of them the conclusion will depend very

¹²⁵ See, e.g., Thomas E. Cooper, *Most-Favored-Customer Pricing and Tacit Collusion*, 17 RAND J. ECON. 377 (1986); Steven C. Salop, *Practices that (Credibly) Facilitate Oligopoly Coordination*, in NEW DEVELOPMENTS IN THE ANALYSIS OF MARKET STRUCTURE 265 (Joseph E. Stiglitz & G. Frank Mathewson eds., 1986).

¹²⁶ A further, underexplored complication is that, even if adoption is industry wide, such contracts nevertheless make punishment, not just cheating, more costly. See Buccirosi, *supra* note 104, at 340.

¹²⁷ See, e.g., Aaron S. Edlin, *Do Guaranteed-Low-Price Policies Guarantee High Prices, and Can Antitrust Rise to the Challenge?*, 111 HARV. L. REV. 528 (1997); Aaron S. Edlin & Eric R. Emch, *The Welfare Losses from Price Matching Policies*, 47 J. INDUS. ECON. 145 (1999); Sridhar Moorthy & Ralph A. Winter, *Price-Matching Guarantees*, 37 RAND J. ECON. 449 (2006); Salop, *supra* note 125.

much on the circumstances of the particular case. Another variable concerns how reliable is the implication of the facilitating practice in comparison to that of other sorts of evidence bearing on the existence of successful coordination. Once again, different types of proof are best employed in a complementary fashion.

2. *Conduciveness of Conditions*

Assessment of the degree to which industry conditions in a particular setting are conducive to successful oligopolistic coordination can sharpen the accuracy of inferences on the ultimate question of whether such coordination is occurring. This subsection begins by elaborating on this idea because of its importance and because the relevance of conduciveness under commonly advocated approaches may be radically different (arguably the opposite in many settings).¹²⁸ Then attention turns to factors bearing on conduciveness to collusion. Last, the subsection reflects on the weight that should be given to these structural factors in light of doubt cast on the structure-conduct-performance paradigm by many decades of empirical research.

The social cost of errors and thus the importance of their reduction is emphasized in subsection II.C.1, and the (sometimes significant) uncertainty about inferences regarding whether coordinated oligopolistic price elevation is taking place was the subject of the preceding subsection and will recur in section B. It remains to consider precisely how evidence bearing on the conduciveness of conditions improves the inference process. First, in a crude sense conduciveness is a necessary condition to success.¹²⁹ In the extreme case, if oligopoly pricing is impossible under the circumstances, we can confidently infer that it is not taking place. Similarly, if conditions are unconducive, success seems unlikely.

One reservation to this argument concerns ill-conceived attempts. As will be discussed in section V.B, it is often helpful to punish attempts even if they fail—or, importantly, when it is difficult to know whether they have succeeded. Nevertheless, the deterrence benefit of such punishment is likely to be limited when considering realms in which success is difficult or impossible in

¹²⁸ Specifically, under the approach that focuses on explicit, direct, typically secret communications, it is commonly believed that conditions highly conducive to successful coordination—rather than merely moderately conclusive—favor a finding of no liability because the likelihood that explicit communications were used is considered to be low when they are less essential.

¹²⁹ For this reason, conduciveness may be a candidate for a procedural screen (*see infra* section V.A), just as market power is a prerequisite for some competition law violations. One difficulty with this approach, just as with market power, is that substantial evidence and analysis may be required to make the assessment, and it may be as much or more uncertain than the substantive test for liability.

any event, and we should also be more skeptical that an attempt occurred in unconducive circumstances.

The more important qualification is that the investigator or adjudicator may be mistaken about how unconducive the conditions actually are. Particularly in large industries with significant stakes, the firms involved will ordinarily have a better grasp of industry conditions and of their own ability to succeed in spite of them than will an outsider. Thus, if one sees clear attempts to coordinate or strong evidence that oligopolistic coordination is successful, the better inference is that the mistake lies not with the firms' analysis of conditions but rather with the enforcer's.

On reflection, it should be clear that inferences about conditions should influence inferences about successful coordination and vice versa. If we are highly confident about one set of inferences, then we should adjust, perhaps significantly, our inferences about the other. Thus, if conditions seem quite unconducive, we should be more skeptical of evidence that may otherwise lead us to believe that oligopolistic price elevation is taking place.¹³⁰ However, if evidence on the latter is quite strong, then we should doubt evidence that conditions are highly unconducive. Often, there will be nontrivial uncertainty about both matters. In addition, even when there is reasonable confidence regarding conduciveness, it may be that it is at an intermediate level, readily admitting the possibility that coordination would succeed and that it would fail. In such cases, we would need to rely primarily on the strength of evidence bearing on success to determine whether oligopolistic coordination is taking place.

To round out consideration of the relevance of conduciveness, it is also important to emphasize that although, at some level, it is a necessary condition to success, it is not a sufficient condition, and this is so even when conditions are extremely conducive. First, the characterization may be incorrect. Moreover, high conduciveness is no guarantee: even tasks properly assessed to be easy may not be doable by certain actors in some settings.

There is, however, a more serious reservation on account of the fact that deterrence may be successful. Even the easiest task will not be undertaken if the expected penalty for doing so is sufficiently great. In a well-functioning regime, coordinated oligopolistic price elevation may be deterred in most instances, a point noted in subsection II.C.1's elaboration of the framework for decision-making. Put another way, unless one believes that the legal system is substantially unsuccessful, one cannot assume that highly conducive conditions in and of themselves imply a significant likelihood of oligopolistic price elevation.

¹³⁰ See POSNER, *ANTITRUST*, *supra* note 7, at 69.

Turning next to the conditions themselves, the subject has received extensive attention in the literature and is not particularly controversial.¹³¹ First, numbers and market concentration, two closely related factors, are important for many reasons. When the number of firms is larger, coordinating on a common price and punishment strategy tends to be more difficult, cheating may be harder to detect because there will be smaller firms whose defection may be more difficult to identify, and smaller firms may also find cheating more attractive because, if their capacity is sufficiently greater than their existing output, they have more to gain relative to what they lose from punishment. There may also be more asymmetry when numbers are larger, which also inhibits coordinated price elevation. One reason of immediate relevance is that, as just mentioned, firms with smaller shares have a greater incentive to cheat, and the greater the asymmetry in firm size, the greater this problem will tend to be. Accordingly, it is unsurprising that the number of firms in the industry was low and concentration high in most prosecuted cases.¹³²

Firms' capacities are also relevant. Greater capacity is two-edged: a cheater may be able to grab more of the market if its capacity is larger, but other firms have a greater ability to punish, keeping in mind that the industry price can only be driven down if firms as a whole can supply a sufficient quantity.¹³³ When there are more than a few firms, it would seem that greater capacity would tend to help cheaters more since there are more firms that can contribute to the quantity increase required for punishment, implying that each does not need substantial additional capacity. Asymmetries in capacity make collusion more difficult: a firm with substantially more excess capacity may gain much from cheating, whereas the others with limited additional capacity may find it difficult to impose punishment.

Another important factor is the transparency of prices and the structure of the buyer side of the market, two dimensions that are often interrelated. Cheating is easiest to detect when prices are open.¹³⁴ When there are large numbers of small buyers (consider markets for many consumer goods), public

¹³¹ See, e.g., CONNOR, *GLOBAL PRICE FIXING*, *supra* note 59, at 32–42; POSNER, *ANTITRUST*, *supra* note 7, at 69–79; VIVES, *supra* note 13, at 306–10; Hay & Kelley, *supra* note 59, at 14–17; Kaplow & Shapiro, *supra* note 13, at 1108–21; Levenstein & Suslow, *supra* note 59, at 57–75.

¹³² See, e.g., Hay & Kelley, *supra* note 59, at 22 tbl. 2, 23–24 & tbl. 3, 27. Concentration was extremely high for the large international cartels examined in CONNOR, *GLOBAL PRICE FIXING*, *supra* note 59; see *id.* at 123, 174, 249–52, 261.

¹³³ See, e.g., William A. Brock & Jose A. Scheinkman, *Price Setting Supergames with Capacity Constraints*, 52 *REV. ECON. STUD.* 371 (1985).

¹³⁴ Furthermore, recall from the discussion of facilitating practices in subsection 1.c that many of these pertain to making price information more available. More generally, most facilitating practices have the characteristic that they aim to operate on one of the factors identified in this subsection in such a way as to make conditions more conducive to collusion; conversely, for any factor noted here that is subject to firms' control, one can imagine corresponding facilitating practices.

pricing is common, which should make collusion easier. In contrast, when there are a few large buyers that have the ability to make infrequent large orders, the terms of which are privately negotiated, collusion tends to be more difficult. (In this respect, the evidence that many prosecuted price-fixing cases involve sales of intermediate goods¹³⁵ may seem surprising, although there are other factors that may explain this fact.¹³⁶) An interesting and important exception concerns public auctions, where governments ordinarily publish winning bids; this practice makes cheating particularly easy to detect, which may help to explain why so many price-fixing prosecutions involve bid rigging.¹³⁷ Yet another relevant aspect of the buyer side of the market concerns changes in demand over time. As noted in section II.A, demand uncertainty inhibits coordination by making cheating more difficult to identify when prices are not public. Additionally, anticipated changes in demand can affect incentives to cheat by influencing the ratio of cheaters' current gains from defection to their future losses during punishment.¹³⁸

Product heterogeneity also makes oligopolistic coordination more difficult, which is consistent with the fact that most prosecuted cases involve homogeneous products.¹³⁹ First, reaching consensus on price and other terms is more complex when the problem is multidimensional, as discussed in subsection 1.c. Relatedly, when there is significant potential for differentiation, cheating is more likely because it may be more difficult to police product quality, subtle terms of sale (such as being more prompt and helpful in dealing with buyers), and other features. When products are differentiated, the cheating calculus also differs in two opposed respects: on one hand, a small price cut is less attractive because it will draw fewer customers from rivals, but on the other hand, punishment from rivals' price cuts will be less painful.¹⁴⁰

¹³⁵ This point is apparent from examination of the lists of industries in Connor, *Survey*, *supra* note 59, at 136–53; Harrington, *supra* note 59, at 98–102; and Hay & Kelley, *supra* note 59, at 29–38.

¹³⁶ First, intermediate goods are more often homogeneous whereas consumer products are often differentiated (including by location and format of retail outlet). Second, if prosecutions have been limited to cases of explicit communications, it may be that such are less necessary or at least less frequent when prices are more open.

¹³⁷ See, e.g., Connor, *Survey*, *supra* note 59, at 79, 88–89. Note that even if bids are not published, knowing who won may well be sufficient, particularly if the firms arranged in advance that only one of them would put in a lowest bid (rather than all placing identical bids).

¹³⁸ In addition, cheating is more likely when it has a lasting impact on demand, such as when there are customer switching costs or network effects, although these factors also make it more difficult for the cheater to increase market share quickly.

¹³⁹ Examination of the lists of industries in Connor, *Survey*, *supra* note 59, at 136–53; Harrington, *supra* note 59, at 98–102; and Hay & Kelley, *supra* note 59, at 29–38, strongly suggests this to be true (see also *id.* at 24–25, 27), and it is certainly true of the large international cartels examined in CONNOR, GLOBAL PRICE FIXING, *supra* note 59.

¹⁴⁰ See, e.g., R. Deneckere, *Duopoly Supergames with Product Differentiation*, 11 ECON. LETTERS 37 (1983); Thomas W. Ross, *Cartel Stability and Product Differentiation*, 10 INT'L J. INDUS. ORG. 1 (1992).

Greater similarity on virtually any dimension is helpful to oligopolistic firms seeking to sustain elevated prices. It was already suggested that smaller firms and those with relatively more excess capacity are more inclined to cheat. Firms with shorter time horizons also are more prone to defection. (Consider a firm in temporary financial distress that needs additional cash quickly.) Firms with more similar production costs will more readily be able to agree on an oligopoly price and none will gain disproportionately from cheating. More extensive contacts and more similar backgrounds of key decision-makers may make mutual understanding easier and enhance trust. Such might include educational and social backgrounds, longevity in the industry, and participation in joint ventures and trade association activities (even if no secret discussion about pricing takes place).¹⁴¹ Multi-market contact may have similar benefits and also enhances the strategic space in which punishment may occur.¹⁴²

A final set of considerations involves group market power, which is indicated by the industry elasticity of demand (the ease with which consumers switch to outside products as price rises) and also the ability of others to supply the market through entry or expansion by fringe firms not part of the group understanding. In the extreme, if there is little or no ability for even a hypothetical monopolist to increase price, then an oligopoly will fare no better. However, if there is some industry-wide market power, albeit limited in degree, oligopolistic firms may well choose to elevate price to that extent. Observe that, when the oligopoly price is lower, both the gain from cheating and the cost of punishment are each proportionately lower as a first approximation. Hence, if an oligopolistic price was otherwise sustainable, the same tends to be true in a market where power is less. The major caveat is the existence of fixed costs of coordination. Some may involve time and effort. Others may relate to the structure of penalties, which will be elaborated in Part IV. In brief, for fines or damages that are proportional to the price elevation, lower profit potential implies correspondingly lower penalties, so the deterrence calculus would be largely the same. However, there may be fixed punishment costs, including the costs of responding to investigations and participation in adjudication and also some penalties that even if not fixed do not rise in proportion to the magnitude of the violation, which is common for

¹⁴¹ See, e.g., PORTER, *supra* note 108, at 90 (in discussing how the “*continuity of interaction* among the parties can promote stability since it facilitates the building of trust,” notes that such continuity “also is aided by a stable group of general managers of these competitors”); Edward L. Glaeser, David I. Laibson, José A. Scheinkman & Christine L. Soutter, *Measuring Trust*, 115 Q.J. ECON. 811, 834–35 tbl. 6 (2000) (in experiment on trust and trustworthiness, finding higher levels when participants knew each other longer); Christopher R. Leslie, *Trust, Distrust, and Antitrust*, 82 TEX. L. REV. 515 (2004).

¹⁴² See, e.g., B. Douglas Bernheim & Michael D. Whinston, *Multimarket Contact and Collusive Behavior*, 21 RAND J. ECON. 1 (1990).

prison terms and also for government fines in many jurisdictions. In that case, deterrence will be more powerful when industry market power is lower. A countervailing factor is that enforcement may be less likely for smaller elevations because they are harder for enforcers to detect and prove and also because enforcers may choose to devote fewer resources to such cases.

At this point, it is useful to reflect further on the importance of these conditions bearing on the feasibility of collusion in different markets. In addition to the already noted logical relevance of conduciveness to inferences about whether successful oligopolistic coordination is taking place, empirical evidence on the correlation between various indicia of conduciveness and industry profitability or price elevation is also pertinent. After all, the weight that should be accorded this type of evidence depends importantly on how predictive it is of the problem we seek to detect. To take as an extreme, an often mentioned view is that oligopolistic price coordination is inevitable, even through pure interdependence, when conditions are favorable;¹⁴³ but this view seems quite inconsistent with existing empirical work.¹⁴⁴

Since the 1980s (if not earlier), industrial organization economists have become quite skeptical of what is referred to as the structure-conduct-performance paradigm that, in its simplest form, holds that market structure dictates conduct which in turn dictates performance.¹⁴⁵ In the present setting, the proposition would be that conduciveness determines whether firms behave interdependently which in turn determines whether prices are elevated and thus whether firms realize supracompetitive profits.

Reservations about the relationship between structure (industry conditions) and performance are due in significant part to cross-sectional studies of different industries in which structural factors, notably industry concentration, explain little if any of the variation in profitability. In addition, even if there is a relationship, there are competing explanations, notably, that concentration is often due to one or a few firms being highly efficient, which is the basis for their profitability, rather than being caused by industry-wide price elevation. Evidence for this interpretation derives from some studies showing that higher profits are only enjoyed by the largest firms. That claim, in turn, is contested on the ground that smaller firms may operate largely in different market segments (strong product differentiation) or that lesser efficiency of fringe producers is not inconsistent with the large firms collectively elevating price. Resolving this debate has waned in importance as more studies have failed to

¹⁴³ See *infra* note 211.

¹⁴⁴ It is also inconsistent with the evidence presented in section II.D indicating the frequent use of explicit communications—thereby risking heavy sanctions—even when conditions are conducive.

¹⁴⁵ See, e.g., Baker & Bresnahan, *Economic Evidence*, *supra* note 70, at 24–26.

find a substantial relationship between concentration and profits in any event, although there is evidence of some correlation.¹⁴⁶

Stronger evidence consistent with the importance of structural conditions is that they, notably concentration, correlate with price, as distinguished from profits. Such findings are more common in intra-industry studies, often comparing price across regions where concentration varies. On one hand, such studies are more powerful because many problems in measuring profits are avoided, more factors are implicitly controlled since only a single industry is considered, and price is more relevant because a lack of high profits may arise even with elevated prices if nonprice competition erodes margins. There are also limitations, notably, that costs may vary across regions and in ways related to concentration; that is, higher costs may lead to both higher concentration and higher price. Accordingly, it is important to control for costs. More broadly, regressions relating concentration or other features of industry structure to price or to profits suffer from endogeneity, so greater effort is required to identify causation and measure the magnitude of any effects.¹⁴⁷

Another source of evidence not as often considered in the industrial organization literature is that discussed in section II.D deriving from enforcement against price fixing. As mentioned, these cases are typically in industries with high concentration, few firms, and homogeneous products, precisely the structural conditions believed to be conducive to successful oligopolistic coordination.¹⁴⁸

Taken together, there certainly is justification for paying attention to the conduciveness of industry conditions in examining particular cases. Nevertheless, empirical evidence gives us reason not to place excessive weight on their relevance.¹⁴⁹ It probably remains true that highly unconducive conditions

¹⁴⁶ Schmalensee offers the following view of the competing hypotheses:

At any rate, there is no support here for the use of the DEH [differential efficiency hypothesis], the DCH [differential concentration hypothesis], or the DEH/DCH hybrids as maintained hypotheses in policy analysis or the study of individual industries.

It would appear likely that the relative importance of collusion and differential efficiency vary considerably among industries and over time.

Richard Schmalensee, *Collusion Versus Differential Efficiency: Testing Alternative Hypotheses*, 35 J. INDUS. ECON. 399, 420 (1987); see also PERLOFF ET AL., *supra* note 21, at 33–34; sources cited *infra* note 149.

¹⁴⁷ See, e.g., William N. Evans, Luke M. Froeb & Gregory J. Werden, *Endogeneity in the Concentration–Price Relationship: Causes, Consequences, and Cures*, 41 J. INDUS. ECON. 431 (1993) (discussing the issues and implementing a solution for the airline industry, finding that an unbiased estimate of the effect of concentration on price exceeds the ordinary least squares estimates by 150–250 percent).

¹⁴⁸ Some caveats on the interpretation of such evidence are noted in section II.D.

¹⁴⁹ See, e.g., Bresnahan, *supra* note 56, at 1055; Schmalensee, *supra* note 58, at 952, 971, 976, 988; see also Rosa M. Abrantes-Metz & Luke M. Froeb, *Competition Agencies Are Screening for Conspiracies: What Are They Likely to Find?*, ECONOMICS COMMITTEE NEWSL. (ABA Sec-

strongly negate the plausibility that firms successfully elevate price. Given the noisy relationship, however, strong evidence of success should usually be regarded as convincing. On the other hand, high conduciveness is a weak basis for assuming success, particularly in light of the point emphasized at the outset of this section that, if deterrence is strong, oligopolistic price elevation should be infrequent even when conditions are ideal.¹⁵⁰

Conduciveness is also important with regard to the magnitude of deterrence benefits and chilling costs. When conditions are quite uncondusive, any successful oligopolistic coordination is likely to result in price elevations that are small and short-lived. On the other hand, chilling costs are likely to be greater in such situations because these industries are reasonably likely to be competitive. And when conditions are highly conducive, undeterred firms may succeed in significant, long-lasting price elevations, whereas chilling costs are less worrisome since highly competitive behavior is less likely.

B. INTERNAL EVIDENCE AND COMMUNICATIONS

1. *Internal Evidence of Oligopolistic Coordination*

The discussion throughout section A, like much prior literature, proceeds implicitly as if evidence of successful oligopolistic coordination will largely be external to the firm. However, for virtually any factor and channel of inference, there is corresponding internal evidence. In addition to hard evidence—for example, invoices or other records indicating prices charged and quantities sold—there is a variety of other information. Some of it may directly convey firms' thinking (strategy or decision memos, notes of meetings, internal policy pronouncements), and much more will convey aspects indirectly (cost and marketing data being suggestive of firms' beliefs about marginal cost and demand). This subsection first considers more fully how various sorts of internal evidence may be relevant to inferences about oligopolistic behavior, on the assumption that reliable internal information can be obtained, and then takes up two important reservations: the difficulty of discerning what a firm

tion of Antitrust Law), Spring 2008, at 10 (survey suggesting that recent efforts at screening have been largely unsuccessful); Paul A. Grout & Silvia Sondregger, *Structural Approaches to Cartel Detection*, in *EUROPEAN COMPETITION LAW ANNUAL: 2006—ENFORCEMENT OF PROHIBITION OF CARTELS* 83 (Claus-Dieter Ehlermann & Isabela Atanasiu eds., 2007) (advocating the use of structural screening by enforcement agencies as a component of cartel detection); Joseph E. Harrington, Jr., *Behavioral Screening and the Detection of Cartels*, in Ehlermann & Atanasiu, *supra*, at 51 (advocating greater use of behavioral rather than structural screening by enforcement agencies); Scherer, *supra* note 8, at 982–83 (offering a highly skeptical view of the prospects of using industry structure as a screen).

¹⁵⁰ A concern with empirical studies of the relationship between structural conditions and price elevation is that they do not take into account the deterrent effects of anti-price-fixing laws. If deterrence is substantial, this could help to explain the weakness of the observed relationships.

“knows” and the problem that firms may sanitize or otherwise pervert records of internal understandings in order to avoid liability.

Confining attention to larger firms,¹⁵¹ a substantial portion of the time of higher-level managers, as well as efforts throughout particular departments, is devoted to information gathering, analysis, and decision-making. Sometimes, there are outside consultants and industry-wide information agencies as well. These activities generate all manner of notes, reporting up and down chains of command, memos, communications, meetings (formal and informal), and so forth. They relate to what firms know or believe as well as to the reasoning behind their decisions.

It follows that there may often be significant internal evidence bearing on every consideration identified in section A. Consider pricing patterns. When a firm suddenly (or otherwise) increases its price, it will have reasons, usually supported by its own evidence. If there is a corresponding increase in cost, this will be seen in the firm's internal cost data and projections as well as in its decision-making process. When price suddenly drops, similar internal information and activity will also be involved. Whether the firm is a secret price cutter, is responding to perceived cheating by others, or is reflecting changes in cost or demand, the information and rationale will be reflected internally. The same can be expected for more modest price adjustments. More broadly, if firms are interacting interdependently, playing the sort of repeated game described in section II.A involving a consensus elevated price and punishment strategy, we would expect internal discussions and other traces to differ from the situation in which firms behave as competitors who take rivals' behavior as given.¹⁵²

Regarding the existing degree of price elevation, one would expect firms to have knowledge of their own prices and marginal cost and thus an estimate of price-cost margins.¹⁵³ Firms think about which costs are fixed and variable

¹⁵¹ If adjacent, owner-operated gasoline stations coordinate their prices, the only internal source of information about their thinking is likely to be the owners' own statements upon questioning (unless perhaps they keep journals, confide in spouses, or are talkative at bars after work). Large organizations are quite different in this regard.

¹⁵² For example, in some court cases (*see* Kaplow, *supra* note 3, at 747–48 n.158), there is evidence of firms' “understanding” with others, viewing competitors as “friends,” and so forth, all of which indicates interdependence. Interestingly, courts referring to such evidence tend to view it as indicative of express communications in contrast to pure interdependence, a logical error since pure interdependence involves a meeting of the minds and resulting mutual understanding that is equivalent to that existing in a classic cartel; the difference is in how that understanding came about. *See also* Jeffrey L. Kessler & Ronald C. Wheeler, *An Old Theory Gets New Life: How to Price Without Being a ‘Price Signaler,’* ANTITRUST, Summer 1993, at 26, 26 (“internal company documents about a price change in a concentrated market will quite naturally often speculate about the possible reaction of rivals to the price change”).

¹⁵³ Joseph Farrell and Carl Shapiro (economists who have spent time in competition law enforcement agencies) argue that, although “gross margins are hard to measure using public data

and how joint costs are properly allocated. They know when production is at or near capacity and if marginal cost is rising sharply. When they price discriminate or grant a price concession to a large buyer, they presumably are aware of their costs and their reasons for charging different prices to different customers. If their prices vary across geographic markets, they again have reasons and information on which their reasoning is based. In deciding how much of a cost shift to pass on to consumers or how to respond to demand fluctuations, they are thinking about whether their marginal costs are constant over the relevant output range, what is the elasticity of the demand they face, and possible interactions with competitors. If they have excess capacity, they have thought about using more of it, which probably involves reducing price, and presumably have decided against it, again, for a reason.

Facilitating practices may be adopted to ease oligopolistic coordination or, in many cases, for additional or alternative reasons. If firms announce prices in advance, they know why they have done so. If they voluntarily share sensitive information through a trade association or otherwise, they anticipate gaining something in return and have a sense of what the benefit will be. If a firm's product is superior to others, those in marketing and sales will want to take advantage of this fact, and if they are ordered not to do so, there will be reasoning behind this decision; likewise if they desist from offering new products, enhancing service to buyers, and so forth.¹⁵⁴ If they have adopted most-favored-customer clauses, price-matching guarantees, or resale price maintenance, they will have considered the pros and cons and reached a decision.

One particular facilitating practice, which will be elaborated in the next subsection, concerns interfirm communications. These will also tend to leave internal tracks. Not only may there be direct evidence of such communications—such as email, notes, and explicit mention thereof—but there also might exist other evidence indicating that such communications have been made, such as analyses or discussions premised on knowledge obtained thereby. In addition, evidence may clarify why they were made, which is par-

. . . and methods sufficiently standardized for cross-sectional studies[,] . . . firms have an incentive to keep track of their cost functions via managerial accounting tools, for instance to know how far they can profitably cut prices. Such information . . . normally is available to antitrust agencies and courts." Joseph Farrell & Carl Shapiro, *Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition*, 10 B.E. J. THEORETICAL ECON., Jan. 2010, art. 9, 1, 18, <http://www.bepress.com/cgi/viewcontent.cgi?article=1563&context=bejte>. Others are more skeptical. See, e.g., MASSIMO MOTTA, COMPETITION POLICY 116 (2004) ("Determining the impact of a marginal change in the quantity produced by a firm on the total cost of production is often beyond practical feasibility even with the best knowledge of the technological conditions under which a firm operates. Indeed, there might be large differences in the estimates of marginal costs even within the management of the same firm."); Jonathan B. Baker, *Market Definition: An Analytical Overview*, 74 ANTITRUST L.J. 129, 142–43 (2007); Gregory J. Werden, *Demand Elasticities in Antitrust Analysis*, 66 ANTITRUST L.J. 363, 394 (1998).

¹⁵⁴ See POSNER, ANTITRUST, *supra* note 7, at 97.

ticularly important when the competitive effects of the communications seem to be ambiguous: consider again the example of public price announcements.

Similar reasoning applies to the conduciveness of conditions. Firms know their production capacities and have at least an educated guess about those of competitors. Similarly, they have views about cost asymmetries and the degree of homogeneity of their products, including, importantly, how salient are any differences to buyers. They have a sense of how well they know and can trust counterparts at rival firms. And they have a view on industry market power, including demand elasticity and prospects for entry and fringe expansion.

In sum, regarding every question considered in section A on market-based evidence, there often will exist traces and sometimes a rich vein of pertinent material inside the firm. Note, importantly, that omissions can also be revealing. For example, a lack of new data or of discussion about changes in demand makes a sharp demand shift an unlikely explanation for a price move. In contrast, if a new product proposal is shot down because detailed studies show that the firm lacks production capability or that consumers are uninterested, suppressing heterogeneity is an unlikely explanation.

Such internal evidence is underplayed in the analysis of competition policy, including importantly that directed toward what might be provable in an enforcement action. It sometimes seems as if firms are personified to such an extent that they are treated as if they were a single person whose mind cannot be fathomed by an outsider. An individual might be called to testify, but it may be supposed that the person would lie and, lacking other evidence, perjury could not be deterred or overcome.

Consider now two important limitations in relying on internal evidence from firms. First, since a large firm is a group of many individuals with complex, overlapping, and sometimes conflicting duties, it can be difficult to determine what a firm knows or what reasoning explains its actions. Many decisions are based on soft information or are made in contradiction to available information, perhaps because the information is seen to be unreliable, because there are overriding considerations, or because of incompetence. There may exist internal disagreement or misunderstanding. Firm politics may also play a role, reflecting that employees are not perfect agents of the owners. Reasons may be clear in the minds of some, but they may not be the ones with decision-making authority.

In addition, there remains the question of how such views and intentions will be reflected in firms' records and other indicators; there may be omissions, incompleteness, and ambiguity. And individuals who testify when the firm is subject to an enforcement proceeding may have strong incentives to bend the truth, selectively recall events, and the like. Disgruntled employees

may tilt in the opposite direction, perhaps attributing nefarious motives for price changes that are innocuous. To top it off, in an adversary proceeding, in which an outsider—the government or a private plaintiff—seeks to reconstruct firms' knowledge and thinking, creating a mosaic that selects from myriad fragments, the potential for confusion and conscious obfuscation by both sides is vast.

That said, it seems plausible that some substantial and fairly reliable conclusions will often be possible to reach. It is not as if firms can operate without even an approximate sense of what are their costs, what they believe about demand, and why they act. If successful oligopolistic coordination significantly elevates prices for extended periods of time, it seems that there may well be substantial internal indicators. Similarly, if no such behavior occurs, there may be much evidence inconsistent with hypothesized coordinated price elevation.

A second concern is that lawyers or others will anticipate the liability implications of firms' knowledge, decision-making, and actions and therefore will sanitize or distort the various clues to be found within the firm. Some may believe that this sort of adjustment is not very difficult to accomplish, but it is not clear that experience bears this out.¹⁵⁵ Moreover, we should be skeptical that a cleansing process can be highly effective.

Initially, it is difficult for a team of lawyers or others to reach broadly and deeply into a large corporation, controlling how myriad individuals speak, write, email, and otherwise behave. Furthermore, any such process that is not limited to a handful of key agents may itself leave incriminating tracks. Relatedly, aggressive destruction of electronic and physical documents is often prohibited, difficult to implement (since so many copies will exist in so many places), and leaves its own traces that can be highly incriminating.

Furthermore, even if firms succeeded in insuring that no individuals made explicit reference to forbidden activity, most of the information described in this subsection would remain. That is, none of the foregoing depends on find-

¹⁵⁵ See, e.g., *Barry Wright Corp. v. ITT Grinnell Corp.*, 724 F.2d 227, 232 (1st Cir. 1983) (Breyer, J.) (“[I]f the search for intent means a search for documents or statements specifically reciting the likelihood of anticompetitive consequences or of subsequent opportunities to inflate prices, the knowledgeable firm will simply refrain from overt description.”). For example, regarding some aspects of the U.S. enforcement actions against Microsoft, it is notorious that the leader of the company (among others) sent (and retained record of) numerous emails that strongly indicated anticompetitive purpose. See, e.g., John Heilemann, *The Truth, The Whole Truth, and Nothing But The Truth: The Untold Story of the Microsoft Antitrust Case and What It Means for the Future of Bill Gates and His Company*, WIRED, Nov. 2000, available at http://www.wired.com/wired/archive/8.11/microsoft_pr.html. Likewise with the Whole Foods merger. See, e.g., Judith Levy, *FTC: Whole Foods–Wild Oats Merger Would Thwart Competition*, SEEKING ALPHA, June 20, 2007, <http://seekingalpha.com/article/38876-ftc-whole-foods-wild-oats-merger-would-thwart-competition>.

ing documents in which firms admit to price fixing. Instead, it refers to all manner of information regarding firms' costs, demand, strategic decision-making, and so forth. It seems rather implausible that all such information could be eliminated, kept fully hidden, or elaborately concocted to present a reasonably consistent picture that departs radically from the truth. And if such was attempted, the inevitable lapses—as well as the attempts at coordinating the charade—may well belie what was happening.

The firms' task is even more daunting because we are examining the possibility of coordinated behavior. For example, suppose that a group of firms significantly elevates price to a supracompetitive level. Consider how difficult it would be for all of them to fabricate consistent stories about a cost shock that had not in fact occurred. Each would need relevant departments to invent data, produce reports, and so forth, and this fabricated information would have to largely coincide across the firms yet appear to have been independently generated. All of this would need to happen quickly and be documented before the moment of the pricing decision. And no one at any of the firms could leave traces, or subsequently be induced to testify, about any of it. Finally, the fabricated cost shock would have to be consistent with external evidence. On the other hand, if there really was a cost shock, and accordingly the price increase was an ordinary response of competitors, there would likely be substantial corroborating material produced in the ordinary course of business by all of the firms, and it would be consistent with external indicators.

Sometimes covering tracks may be possible. When there is a simple, standardized product sold in a stable market, perhaps only a few individuals who have price-setting authority at each firm need to be involved in pricing decisions, and they may be able to operate privately, leaving no recorded clues. Such seems to have been the case in a number of prosecuted price-fixing cases, although some traces may ultimately have surfaced. But there, internal searches often focused on admissions of explicit, secret price-fixing communications between firms. It is quite another matter for such firms to hide all cost and demand information or, as just described, to concoct in parallel elaborate, phony information that appears to have been collected and disseminated throughout each firm.

In all, internal evidence is no panacea, but it seems to have substantial potential.¹⁵⁶ Moreover, as with the other bases for inference considered throughout this Part, different means can be employed together. Thus, if the question is whether a price increase can be explained by a common cost shock, one would look to external evidence, such as market data and information from

¹⁵⁶ It also has costs, both adjudication costs and ex ante costs implicit in what has just been discussed in that firms may undertake some efforts at subterfuge, which may entail direct costs as well as losses from reductions in internal operational efficiency.

suppliers, and also to internal evidence from the firms in the industry. Using both, a more reliable conclusion can be drawn.

2. *Interfirm Communications*

Interfirm communications—particularly explicit ones, often taken to be secret—are the focus of many commentators' views of what the regulation of oligopolistic price elevation is or should primarily be about. The question for present purposes, however, is how interfirm communications are relevant in making inferences about whether successful oligopolistic coordination has taken place. To a substantial extent, the answer is already incorporated in the preceding subsections.

First, interfirm communications contribute to our understanding of what firms have done and why. In that respect, they are much like some of the other types of internal evidence just considered. Thus, when there is evidence not just on the existence of interfirm communications but also on their content,¹⁵⁷ they may provide a strong basis for inference if they are sufficiently explicit and reliable,¹⁵⁸ especially when they refer to prior, successful price elevation—although the same is true of many other types of internal evidence. Likewise, the caveats concerning internal evidence are relevant, some with particular force. In addition to problems of authority, ambiguity, and conflicts, the concern that firms may hide their tracks is particularly great because interfirm contacts may, as noted, be confined to a few individuals whereas much other internal activity is more ubiquitous.

Note that public interfirm communications, ranging from advance price announcements to commentary on industry conditions, are not hidden and may be as revealing as secret discussions. Indeed, if public interchanges were *per se* legal, firms could simply move their meetings from hotel rooms to joint press conferences. Even when they are purportedly directed toward the world, sequential advance announcements arriving at a consensus price may indicate

¹⁵⁷ When all we know, for example, is that secret meetings took place, we may still be quite suspicious of firms' activities, which can be useful in making inferences for purposes of liability. However, as will be discussed in section IV.A, if we wish to levy sanctions that are based on the level of price elevation, it will be necessary to rely primarily on other evidence for that purpose, although our confidence that such other evidence is not entirely mistaken may be greater. *See also infra* section V.B (discussing the punishment of attempts to elevate prices).

¹⁵⁸ Interfirm communications pose an additional issue concerning reliability because they may be designed to mislead competitors. For example, a firm may try to convince others it will go along with price elevation because it intends to cheat from the outset, hoping to gain at others' expense. One would often need further evidence to know whether firms actually engaged in price elevation. Such other evidence would often be sufficient in itself. More to the point, interfirm communications may nicely complement such evidence if the latter does not, standing alone, resolve all doubt as to what transpired. Another, perhaps more farfetched, possibility is that firms would orchestrate discussions designed to mislead enforcers, perhaps by having a discussion about a supposed cost shock, as suggested in note 159.

what firms are doing. In this respect, they may supplement information from price moves themselves; for example, if an industry-wide price increase is due to a common cost shock and firms are behaving competitively, they will have less need to feel out their rivals before raising their prices.¹⁵⁹ Of course, to the extent that public communications are understood to provide a basis for inferring a violation, they would be discouraged, possibly being driven underground.

Second, as already mentioned in subsection 1.c, interfirm communications can serve as an important facilitating device and thus support an inference of oligopoly pricing. These might include public statements—advance price announcements, predictions about industry demand or costs, and open discussions of various matters at trade association meetings—as well as secret meetings. These communications may help firms reach a consensus, which can be challenging when there are multiple equilibria and there may be disagreement about which is best, and also when more complex understandings are required, perhaps because of product heterogeneity. In addition, direct interchange may help identify and coordinate the punishment of cheaters, although there are problems regarding the credibility of firms' statements in this context.¹⁶⁰

Relatedly, as a facilitating practice, some forms of interfirm communications are an attractive target because there is little social cost if firms are deterred from engaging in them. Secret meetings in hotel rooms come to mind. For other communications, such as public price announcements, trade association activities, and participation in standard setting or joint ventures, there are varying degrees of possible benefits that need to be considered.

Consequently, interfirm communications of all sorts can be important in making reliable inferences. It is also clear that their reliability may be enhanced, perhaps significantly, by combining them with the other evidence considered throughout this Part. Accordingly, evidence on interfirm communications should be viewed as a weapon in the detection arsenal but not as the be all and end all.

A related question is how probative is the lack of explicit interfirm communications. First, since firms attempt to keep them secret and may succeed in doing so, the failure to find such communications is hardly conclusive that

¹⁵⁹ Firms might still seek to learn from their competitors, for example, if the magnitude of the cost shock is somewhat uncertain, which complicates the inference process. Note that, if one gives sufficient credence to this qualification and wishes to avoid any interference with this interfirm learning process, one might then justify permitting explicit price discussions—or explicit discussions couched in terms of purported cost changes, which could cover for the same thing.

¹⁶⁰ For elaboration of these points, see Kaplow, *supra* note 3, § IV.B.2.

they did not occur. Second, successful oligopolistic coordination is sometimes possible without explicit, elaborate communications. Regarding both points, there is an important interaction with the conduciveness of conditions, discussed in subsection A.2. When conditions are most conducive, particularly when the number of firms is small, explicit interfirm communications may be less essential and also more difficult to detect.¹⁶¹ This suggests that, when the danger of oligopolistic price elevation is greatest, we should be less bothered by the lack of such evidence. Similarly, when firms' coordination problem is highly complex and likely to require extensive, explicit communications among large numbers of individuals, the absence of direct (or strong indirect) evidence that any such communications took place would make an inference of successful oligopolistic coordination notably weaker.

An additional point is that, in attempting to detect successful oligopolistic coordination, it generally makes little sense in cases where, let us suppose, no secret interfirm communications are detected to attempt to infer whether they have taken place through the use of circumstantial evidence. Since the reason we care about these communications is that they illuminate other evidence, nothing further is added by this process. Stated abstractly, if evidence E_1 , E_2 , and E_3 give rise to a probabilistic inference that internal communication C occurred, any implication from the set E_1 , E_2 , E_3 , and C about successful oligopolistic coordination can be no stronger than what one could have inferred from E_1 , E_2 , and E_3 directly.¹⁶²

C. LIABILITY ASSESSMENT

Few simple lessons can be drawn from the foregoing examination of the many types of evidence that may be used to support or negate inferences of successful oligopolistic coordination, particularly in light of the nature of the optimal decision-making framework discussed in section II.C, with its emphasis on ex ante effects on deterrence and the chilling of desirable behavior. This subsection nevertheless offers some observations on the relationship between avenues of proof and the social problem at hand.¹⁶³

First, it is useful to reflect on why the probability of oligopolistic price elevation is important. In some basic models of optimal sanctions, all that

¹⁶¹ Regarding the latter, see Hay & Kelley, *supra* note 59, at 20, 23–24 & n.15, and Robert T. Masson & Robert J. Reynolds, *Statistical Studies of Antitrust Enforcement: A Critique*, in AMERICAN STATISTICAL ASSOCIATION, 1977 PROCEEDINGS OF THE BUSINESS AND ECONOMICS STATISTICS SECTION, PART I, at 22, 25–26 (1978).

¹⁶² That is, even if direct evidence of C would strengthen the ultimate inference, if the only evidence of C is due to the inference from E_1 , E_2 , and E_3 , nothing is added.

¹⁶³ This subsection sets to the side questions of liability for the use of facilitating practices—that is, the condemnation of facilitating practices as such rather than making inferences from the use of facilitating practices about oligopolistic coordination.

matters is the expected harm.¹⁶⁴ Here, further decomposition is necessary, in part because the social consequences of errors may well be asymmetric. In particular, subsection II.C.2's discussion of chilling effects suggests that significant harm may result if highly competitive industries (especially with constant marginal cost) face a nontrivial prospect of sanctions for price elevation. Accordingly, it is important that competitive firms not be subject to even moderate expected sanctions,¹⁶⁵ which implies that the uncertainty in any estimate of price elevation is important.¹⁶⁶

Specifically, it matters how much of the distribution of possibilities involves no price elevation. Different forms of proof, viewed individually and in various combinations, may differ substantially in the likelihood of false positives. For example, some forms of proof may be highly probative of the magnitude of price elevation, yet admit a nontrivial risk of false positives. In contrast, other sorts of evidence—perhaps certain internal evidence or interfirm communications—may be highly probative of whether there is coordinated oligopoly pricing (greatly reducing the likelihood of a false positive) but give little indication of the extent of price elevation.

It is also worth considering further the importance of the conduciveness of conditions with regard to the avoidance of false positives so as to reduce chilling effects. Presumably, there are a large number of markets where conditions are quite unconducive to successful coordinated price elevation. If such cases are not screened out, they might give rise to many false positives—perhaps a small portion of all possible cases that could arise but a large absolute number. In addition to the adjudication costs that would be involved, there is the danger that the distribution of such cases would not be seen as random *ex ante*, but instead certain groups of firms in particular circumstances might have significant reason to fear that their actions would be mistaken for oligopolistic coordination, producing serious chilling effects.¹⁶⁷ Moreover, when

¹⁶⁴ See, e.g., Louis Kaplow & Steven Shavell, *On the Superiority of Corrective Taxes to Quantity Regulation*, 4 AM. L. & ECON. REV. 1 (2002) (showing that the existence of uncertainty about harm is best addressed by having actors pay the expected marginal harm).

¹⁶⁵ It may not be sufficient for such firms to treat the prospect of expected sanctions as a current marginal cost (which they would) because they would then need to raise price as a consequence, which itself could further raise expected sanctions.

¹⁶⁶ Observe that the nature of the cost of false positives is assumed throughout this subsection to concern cases of no actual price elevation. If, as section V.C explores, we may wish to exonerate price elevation attributable to the unilateral exercise of market power rather than to coordinated oligopolistic price elevation, the chilling costs from errors in assessing the existence of price elevation beyond that due to unilateral market power would be different (generally, much less, and possibly even benefits, as will be discussed there).

¹⁶⁷ It may not seem obvious that random chilling effects are innocuous whereas concentrated ones are costly. The core rationale that supports this view is based on the conjecture that chilling effects rise disproportionately in the expected sanction (they are nonlinear) at low levels. For example, if 1000 markets are each subject to an expected sanction of 1, firms in these markets

coordinated oligopolistic price elevation does occur in industries that are un-conducive, the elevations will, on average, tend to be smaller in magnitude and duration. Accordingly, insisting on moderately or perhaps highly conducive conditions seems sensible.¹⁶⁸ Recall from subsection III.A.2, however, that high conduciveness is not sufficient, particularly if deterrence is reasonably effective.

Other features of industries may bear on the risk of chilling effects that may be associated with one or another type of evidence.¹⁶⁹ For example, regarding demonstrations of price elevation that involve direct attempts to measure marginal cost, firms in some industries may each produce a single product with a simple cost structure whereas other industries may be inhabited by firms that produce multiple, related products with complicated technologies and substantial costs that may not obviously be fixed or variable in a relevant time frame. Accordingly, the optimal weight to give to evidence will depend on whether a specific sort of evidence in a particular type of industry is likely, viewed *ex ante*, to pose a problem of chilling effects or to enhance deterrence substantially. Regarding the latter, note that the importance of any type of evidence will depend on what other sorts of proof are likely to be present if coordinated price elevation is indeed taking place. If some type of evidence is unreliable and involves chilling costs but other, superior evidence is likely to be available under the circumstances—conditional on undesirable behavior actually having occurred—then the former evidence might optimally be given little weight. In contrast, if the problematic evidence is likely to be the only type of evidence in such an industry, it may need to be given greater emphasis.

A final point in considering the probability of oligopolistic coordination is that, as subsection II.C.1 explains, a high proportion of actual positives will be false in a well functioning system because successful deterrence greatly reduces true positives. This notion is not inconsistent with the need to avoid

might each, in anticipation, elevate price slightly, which would not materially raise the likelihood of detection and have little social cost (since marginal deadweight loss is zero when price elevation is zero). In contrast, if expected sanctions are concentrated—perhaps 10 markets face an expected sanction of 100—distortion might be great and some might become monopolized, as discussed in subsection II.C.2

¹⁶⁸ As mentioned at various points throughout, some views of more traditional approaches to liability, which emphasize explicit interfirm communications, may have the opposite implication regarding conduciveness. For the reasons just given in the text (and others, including the argument to follow on the magnitude of elevation), this approach serves to reduce rather than maximize social welfare. For substantial elaboration, see Kaplow, *supra* note 9.

¹⁶⁹ Additionally, the analysis in subsection II.C.2 suggests that the social costs attributable to a given prospect of error will depend on industry conditions. Notably, if marginal costs are rising, competitive firms might be able to avoid false positives (with little attendant social cost) by somewhat reducing price, but if marginal costs are constant, false positives can be more problematic.

chilling effects because this other phenomenon concerns the ratio of false to total positives, ex post, whereas chilling effects are primarily determined by the absolute level of false positives anticipated ex ante.

A second observation is that the analysis in section II.B indicates the importance of the magnitude of oligopolistic price elevation. Not only is harm rising with this magnitude, but efficiency costs generally rise nonlinearly. Marginal deadweight loss rises, beginning from zero, as price is elevated above marginal cost.¹⁷⁰ Waste due to excessive entry in the homogeneous goods case similarly is zero at the margin when price equals marginal cost and increasing nonlinearly thereafter. When product variety is valuable, it is plausible that gains are greatest at the margin for small elevations, with the marginal benefit falling thereafter.¹⁷¹

This observation suggests that it may be best in many settings to require evidence of a significant price elevation. Most simply, since there are substantial adjudication costs, it may not be worthwhile to pursue small cases. This basic claim may not, however, be true. One may not know that the elevation is small until most of the investigative costs are sunk. In addition, due to the prospect of deterrence, it may well be optimal to have a policy of pursuing cases with small elevations even if, ex post, the litigation costs exceed the harm in the case at hand.¹⁷²

A second and probably more important reason that it may often be appropriate to require proof of significant price elevation concerns the tradeoff of deterrence benefits and chilling costs. When price elevation is large, not only are deterrence benefits particularly great but it is also probably true that chilling costs are low because the likelihood of a false positive is small. In contrast, when price elevation is very small, not only are deterrence benefits insignificant but chilling costs are probably a larger concern because a nonexistent elevation has a nontrivial likelihood of being misperceived as a small elevation. As mentioned in the discussion of the probability of successful oligopolistic coordination, the prospect of false positives will depend on the

¹⁷⁰ In contrast, it was mentioned that marginal forgone consumer surplus is falling with price elevation, even though the total sacrifice in consumer surplus is rising.

¹⁷¹ An interesting implication of this point combined with the preceding one on chilling effects is that a mean-preserving increase in the variance of the distribution of possible price elevation has ambiguous welfare consequences. For positive values of price elevation, greater variance implies greater social loss. But greater variance also implies a greater likelihood of false positives (no price elevation) and thus of chilling effects as well. Thus, the shape of the distribution matters, an aspect of which forms the basis for the next point in the text.

¹⁷² This point is standard in the economics of law enforcement, which suggests further that the difficulty can be too little rather than too much enforcement if society relies on private enforcement motivated by the prospect of recovering damages or public enforcers who are rewarded in ways leading them to focus on high-stakes cases. See, e.g., Steven Shavell, *The Social Versus the Private Incentive to Bring Suit in a Costly Legal System*, 11 J. LEGAL STUD. 333 (1982).

type of evidence, some of which indicates the magnitude of price elevation (price wars, price discrimination, direct measurement of marginal cost) and some of which may not (the existence of interfirm communications, internal references to the existence of an industry understanding, the use of facilitating practices). But, for many types and combinations of evidence, it seems useful to require that a substantial oligopolistic price elevation be demonstrated.¹⁷³ This suggestion is reinforced when considering, as mentioned with regard to conduciveness, that there exist many market settings in which there is not in fact coordinated price elevation. Indeed, when deterrence is reasonably effective, such elevation is unlikely even when conditions are fairly conducive to successful coordination. Finally, if private lawsuits are permitted, the concern for meritless (or low-merit) litigation designed to extract settlements may further counsel in favor of setting a higher threshold.¹⁷⁴

Anticipating some of the analysis in section IV.A, it is interesting to consider briefly the extent to which the operation of sanctions may be a partial safety valve in cases with low price elevation. If penalties are limited to fines or damages that are themselves proportional to the magnitude of price elevation, then the prospect of sanctions being imposed when no price elevation is mistaken for small price elevation may not cause very serious chilling effects. In contrast, if there are significant sanctions for even small elevations, which seems to an extent to be the practice in some jurisdictions, then the problem would be much more serious. Clearly, the threshold for liability and level of sanctions need to be optimized together.

The foregoing point combined with the preceding analysis in this Part suggests that it is conceptually useful to view the problem of designing a legal regime as one of determining how evidence maps into sanctions. Ordinarily, laws contain elements of liability, each of which must be met by some standard of proof. Liability often exhibits an all-or-nothing character. If there is no liability, there are no sanctions. If there is liability, then we separately assess sanctions by some other set of criteria, perhaps with its own proof burdens. What results from this or any other regime can be described equivalently by a

¹⁷³ Regarding evidence indicating the magnitude of price elevation, it was explained in prior subsections that the risk of error varies greatly, being most significant when attempting to measure marginal cost directly in certain types of industries but less significant when using price wars or price discrimination to indicate successful oligopolistic price elevation. Even regarding the latter sorts of proof, however, larger elevations tend to involve less risk of false positives. For example, what appear to be price wars involving small price cuts may really be competitive reactions to changes in cost or demand that are too small to measure, and small price discriminations are more likely to be attributable to undetected cost differences in serving different groups.

¹⁷⁴ So-called strike suits may nevertheless arise. Even if they cannot be dismissed quickly under an elevated standard, such suits may be less credible if the liability standard is tougher to meet, in which case plaintiffs would be in a weaker position to extract significant settlements, which in turn would deter suits in these circumstances. *See infra* section V.A.

function that, for any set of evidence that might be adduced, simply assigns certain sanctions. In a standard regime, all combinations of evidence that yield no liability map to sanctions of zero. For sets giving rise to liability, one can combine as well the evidence used to assign penalties to determine what sanctions attach to the complete set of evidence.

More generally, sanctions could in principle be any function of the full set of conceivable evidence. For example, instead of absolute liability thresholds, one might assign higher sanctions not only when the extent of price elevation seems to be greater but also when the probability is higher. Small or uncertain elevations might be assigned low rather than zero sanctions. Given administrative costs, however, such may not be optimal. Sanctions are considered further in Part IV; for the present, the point is that both the probability and magnitude of detected oligopolistic price elevation are important in assigning sanctions, whatever their level may optimally be. And, as this section has emphasized, it is necessary to consider explicitly the significance of deterrence benefits and chilling costs when deciding, for example, how high a probability of price elevation or how great a magnitude should be required. Finally, it is important to keep in mind that the process of liability assessment will be anticipated by firms. Therefore, in addition to straightforward deterrence and chilling, there will be other adjustments in behavior—for example, as noted in subsection A.1.a., coordinating firms may raise prices gradually so as to render detection of oligopolistic elevation more difficult—which further complicates the problem of how best to combine various types of evidence in making judgments about liability.

IV. SANCTIONS

Detection must be combined with sanctions in order to reduce the extent of coordinated oligopolistic price elevation in the economy.¹⁷⁵ Section A explores fines and damages (mostly discussed interchangeably although important differences are noted), section B adds imprisonment, and section C discusses injunctions.

Although there is a substantial literature on the economics of law enforcement addressed to the choice among types of sanctions and their optimal magnitude,¹⁷⁶ work on competition policy has devoted little attention to these questions. Much existing analysis is incomplete or misleading as a consequence of this omission. For example, literature on rules of liability often is significantly guided by the implicit or explicit assumption that injunctions will

¹⁷⁵ Note that if the target was not all coordinated oligopolistic price elevation but only a subset, perhaps that which arises due to particular sorts of explicit interfirm communications, much of the analysis would remain applicable.

¹⁷⁶ For a survey, see Polinsky & Shavell, *supra* note 10.

be a central or the sole remedy when this may well not be optimal and also is not in accordance with existing practice. And some work on the magnitude of penalties finds those under consideration, usually government-imposed fines, to be too low but does not take into account that those sanctions may be in addition to others, such as private damages and imprisonment. Accordingly, it is valuable to sketch at the outset some of the considerations that bear on the choice among types of sanctions before considering each in greater depth.

Discussion elsewhere in this article takes fines or damages to be central. Such sanctions are widely employed for price fixing in many jurisdictions. Moreover, basic law enforcement theory suggests that their use is often desirable, for these sanctions tend not to be costly in themselves (in contrast to imprisonment and injunctions), they deter behavior (which injunctions alone do not), and they may be calibrated to the extent of harm caused and the likelihood of detection. Regarding the latter, the general prescription is that, ideally, expected sanctions should be set equal to expected external harm, so fines or damages should equal actual harm multiplied by the inverse of the likelihood of successful detection. Because oligopolistic firms attempt to keep their price coordination secret, this latter point is important in the present setting, and this is the rationale for employing treble damages in private lawsuits in the United States—which, note, are in addition to any fines and prison terms imposed by government enforcers.

Because the use of imprisonment is socially costly—resources are consumed in running the system and the loss to imprisoned individuals is not matched by any direct social offset (unlike fines, which are transfers)—it tends to be optimal not to use it unless monetary sanctions are insufficient. Imprisonment, which is used in price-fixing enforcement, especially in the United States, can nevertheless be important on account of the difficulty of detection, which may require for adequate deterrence a level of fines that exceeds firms' assets, and also because of firms' difficulty in controlling employees, some of whom may have incentives to violate the law even when such is against the interest of the firm as a whole.

Injunctions are notably different from the other two types of sanctions. Most importantly, they do not by themselves achieve deterrence. Furthermore, if fines and damages (and imprisonment) are to be employed, why is there a further need for an injunction? After all, when the other sanctions succeed at deterrence, whether to impose injunctions becomes moot. Also, injunctions are parasitic on the other sanctions because firms have little incentive to abide by injunctions unless they fear penalties from failing to do so. Injunctions do differ from other remedies in that, instead of attempting to induce compliance, they may implement it directly, such as by an agency regulating firms' prices going forward or restructuring an industry so as to make conditions no longer

conducive to successful oligopolistic coordination.¹⁷⁷ Such remedies, of course, are often quite costly in themselves.

A complete analysis of optimal enforcement must consider additional issues as well,¹⁷⁸ but they will be set aside for present purposes (except for some brief mentions, notably, in section V.A). Among them are whether private suits should be used instead of or in addition to public enforcement, how private enforcement should be operated (permission of class actions, allocation of attorneys' fees, determination of who should be permitted to sue), which firms should be liable at all, how obligations for fines or damages should be allocated among the firms (particularly if the defendants do not constitute the entire market or if some are judgment proof), the strategic use of leniency policy (considered briefly in subsection III.A.1.a), and liability for attempts (examined in section V.B). Finally, analysis of detection and sanctions should not, in principle, be compartmentalized, a point elaborated in section III.C.

A. FINES AND DAMAGES

The use of fines or damages—monetary sanctions—is generally desirable, as the introductory remarks suggest. The general principle is that fines or damages (which for the most part will not be distinguished in this section, with fines often used as a shorthand) should equal external harm times a probability multiplier, in the simplest case equal to one divided by the probability that sanctions will be imposed. For example, if firms contemplating price fixing anticipate that there is only a 50 percent chance of being caught and fined, fines should (at least somewhat) exceed twice any profits, for otherwise their activities will be profitable even taking into account expected sanctions.¹⁷⁹

The core deterrence logic is simple and familiar, but is worth emphasizing because it seems that many competition law commentators fail to appreciate it, as discussed in section C, despite Posner's clear presentation of the basic idea in his seminal article decades ago.¹⁸⁰ Whatever methods firms might use to achieve coordinated oligopolistic price elevation—and whether or not the particular means ultimately will be discovered—firms' decision-making calculus changes if they expect to be subject to sanctions. If they anticipate a large gain from successful oligopolistic coordination and no penalties, they

¹⁷⁷ Imprisonment also directly restricts the future behavior of imprisoned individuals. But if ex ante incentives are inadequate to deter, firms might replace such individuals, who, having the same incentives as those replaced, would likewise not be deterred.

¹⁷⁸ See Polinsky & Shavell, *supra* note 10.

¹⁷⁹ Throughout, the analysis simplifies by abstracting from risk aversion.

¹⁸⁰ See Posner, *Oligopoly*, *supra* note 7, at 1588–93.

will undertake great efforts to accomplish it. However, if there are fines, and if the expected level of fines exceeds expected profits from coordination, then they will voluntarily forgo such efforts and aim to ensure that their employees or other agents do not engage in this activity. The logic is no different for, say, the application of a corrective tax to firms that pollute.¹⁸¹ In imposing the tax, the government simply charges firms for their pollution. How pollution is reduced is up to the firms: they may change their production methods, seek to invent new technologies, or reduce output. The government does not need to know which is employed or in what combination. As long as pollution is detected and firms are fully charged for the harm attributable to it, they will control it to an optimal extent. With price fixing, the optimal degree is ordinarily zero, and if expected sanctions are equal to harm to others and thus exceed any profits, that is the level firms will choose.

Subsection 1 further explores particularly relevant aspects of the harm component of the basic formula for optimal fines. To simplify the exposition, it for the most part implicitly assumes that the probability of sanctions is one, so that expected sanctions simply equal the stated sanction. Subsection 2 elaborates on the probability multiplier. Finally, subsection 3 briefly discusses the law on fines in the United States and European Union, which, as will be seen, does not closely match what seems to be optimal. Much of the discussion sets to the side other sanctions, although they will occasionally be noted because of their significance and the fact that the theory of optimal sanctions encompasses the entirety of sanctions as well as having implications for particular components.

1. *Determination of Harm*

What counts as social harm is controversial in the present setting. As subsection II.B.1 explains, total social welfare is an appealing objective to maximize, but many would consider only consumer surplus. As an initial matter, this difference is not important for present purposes because the harm component of the standard optimal sanction prescription refers to all harm that is external to the firm, which is consumer surplus in any event. Thus, if group of firms elevates price, reaping 100 of profits and causing deadweight loss of 25, the harm component of the sanction is 125, which is the total loss in consumer surplus. The point is that firms will in general be induced to behave optimally when they bear the full social costs of their actions. The 100 in profits is directly matched by a loss of 100 to consumers, and sanctions must reflect the latter to achieve deterrence. Clearly, if the expected sanction was only 25 rather than 125, the firm would not be deterred.

¹⁸¹ See, e.g., Kaplow & Shapiro, *supra* note 13, at 1100.

It is also dangerous to set expected sanctions simply equal to firms' profits, or, equivalently, a measure of lost consumer surplus that omits deadweight loss. Firms that compare a profit of 100 to an expected sanction of 100 would be indifferent about price elevation. We might hope that they would refrain from the activity, perhaps because there are additional expenses of undertaking it and on account of adjudication costs not included when stating that the expected sanction is 100. The problem is that there will inevitably be errors in setting the sanction, so if the system aims at setting an expected sanction that just equals firms' expected profits, sometimes the actual sanction will be too high and other times too low.¹⁸² As long as firms do not anticipate such errors, they would still bear an expected sanction of 100. However, if they do anticipate them—for example, they may have knowledge of aspects of their industry that lead adjudicators to systematically underestimate harm, or to overestimate it—then some will be underdeterred, and the additional fines on other firms in different circumstances provides no consolation.¹⁸³ Accordingly, brinkmanship is not good policy. Expected sanctions need to exceed expected profits from harmful activities. It suffices to follow the general prescription that the target should be expected external harm. Alternatively, one could, for example, omit deadweight loss but otherwise augment oligopoly profits in determining this component of the formula for the optimal sanction.

The analysis to this point places a floor on the magnitude of the optimal expected fine. There should also be a ceiling that ideally is no higher than necessary to deter. (The foregoing point about measurement errors indicates that this suggestion is overly simple. In general, there will be a tradeoff; higher fines will usefully augment deterrence in some instances, presumably less and less so as fines rise to ever-higher levels.) A standard reason in the optimal law enforcement literature concerns overdeterrence, which is to say deterring even those acts with social benefits in excess of social costs. For price fixing, there may be few such acts—and, when there are, which subsection II.B.2 indicates is possible, there is no particular reason to believe that such would be the cases where further increments to already high fines would make the difference.

More importantly, as emphasized throughout, chilling costs are quite important. The prospect of false positives tends to discourage some desirable behavior. For a given probability of false positives, the problem will be worse when sanctions are higher. Accordingly, lower sanctions tend to be desirable, all else equal. More broadly, the analysis at the close of section III.C indicates

¹⁸² Even if a system makes no errors, firms' ex ante expectations may be subject to some error, which would produce much the same problem.

¹⁸³ See generally Louis Kaplow & Steven Shavell, *Accuracy in the Assessment of Damages*, 39 J.L. & ECON. 191, 194, 204–05 (1996); Louis Kaplow & Steven Shavell, *Private Versus Socially Optimal Provision of Ex Ante Legal Advice*, 8 J.L. ECON. & ORG. 306, 312–16 (1992).

that optimal sanctions should not be viewed independently of optimal liability rules. For example, it may be that sanctions should be greater when proof is more certain, in significant part because the chilling costs will be lower when false positives are less likely.

It is therefore important to attempt to measure harm reasonably well when determining the magnitude of fines (or other sanctions). Obviously, this task will often prove difficult for many of the same reasons that the accurate determination of liability is challenging. Indeed, some of the evidence is basically the same. When liability is established by pricing patterns or price elevation, as described in subsections III.A.1.a and III.A.1.b, a measure of harm is the basis for finding liability, that is, that some harm due to successful oligopolistic coordination has occurred.¹⁸⁴ This measure, in turn, can provide the benchmark for assessing sanctions.¹⁸⁵ Some internal evidence and interfirm communications, discussed in section III.B, may similarly indicate the magnitude of the overcharge. However, other evidence is more ambiguous. Price stickiness may indicate coordination without demonstrating the extent of price elevation. The existence of interfirm communications may likewise fail to reveal the size of overcharges, although they may if sufficient content is revealed. Similarly, the use of some facilitating practices may signal coordination, or perhaps attempted coordination, but not indicate the degree of success, if any. In the latter sorts of cases, it is necessary to inquire directly into the magnitude of price elevation for purposes of setting the level of sanctions. And once such evidence is examined, it makes sense to use this information to refine inferences regarding liability as well.¹⁸⁶

2. Probability Multiplier

When sanctions are not imposed with certainty, higher sanctions are necessary to achieve a given level of deterrence. A central difficulty in determining

¹⁸⁴ Price elevation for purposes of assessing sanctions is sometimes determined by comparing prices charged during the period of illegal price elevation to those before or after it. The “before” comparison corresponds to detection by attempting to identify a price jump at formation. The “after” comparison assumes that at some point, perhaps when an investigation commences, price will drop to the competitive level. This latter measure, however, creates perverse incentives because firms that fail to reduce their prices not only continue to earn supracompetitive profits but also reduce damages for past actions. See Joseph E. Harrington, Jr., *Post-Cartel Pricing During Litigation*, 52 J. INDUS. ECON. 517 (2004).

¹⁸⁵ See, e.g., Baker & Bresnahan, *supra* note 70, at 15 (methods of detecting supracompetitive pricing “often provide, as a byproduct, an estimate of the cost to buyers of the exercise of market power by sellers”).

¹⁸⁶ Relatedly, there is no reason to be reluctant to include such evidence when considering liability on the ground that the evidence is costly or potentially ambiguous, in light of the fact that it is necessary to use it in calibrating sanctions. If proceedings were bifurcated (*see infra* section V.A), so that liability is determined first, then there may be some cost savings in ignoring such evidence until there is an affirmative finding of liability.

the appropriate probability multiplier is that it is necessary to know what the probability is in a given enforcement regime. This probability, in turn, is the ratio of cases in which sanctions are (properly) imposed to the total number of violations, but since the violations are largely secret, it is quite difficult to know what the multiplier should be. The empirical evidence surveyed in section II.D indicates that there are many successful prosecutions of cases that seem fairly clearly to have involved significant price elevations, and they resulted in large sanctions. As long as those sanctions were notably greater than firms' profits—a point that is subject to dispute, as discussed in subsection 3—it would follow that there must be a significant number of undetected violations since not all prospective violators are deterred. With some laws, like those against murder and automobile theft, it may be fairly straightforward to measure, at least approximately, the number of unsolved cases. But individuals who pay elevated prices are often unaware of this fact, for what is difficult for enforcers or experts to determine is beyond the reach of individual consumers, although large buyers of intermediate goods may have some direct knowledge and an incentive to learn about and report violations. Nevertheless, there is some information on this question, which suggests that the probability is rather low.¹⁸⁷

In addition, the probability of sanctions is not exogenous. Instead it depends, as noted, on the substantive law and rules of proof, and also on enforcement effort. Regarding the latter, public enforcers are given budgets and set priorities, which could be modified. It is commonly suggested that current prosecutions usually follow from customer complaints (presumably the sorts of large customers just mentioned, especially given the intermediate-goods industries in which most prosecutions occur) and from informants. If these cover most actual instances of price elevation, they may suffice. However, if there indeed is significant undetected price elevation, then one option is to employ other enforcement methods that require greater initial effort, such as using econometric methods to identify plausible instances of price fixing, followed by more targeted investigations. Alternatively, it may be cheaper simply to raise sanctions, achieving the same increment to the expected sanction while economizing on enforcement effort, a standard prescription in the law enforcement literature. The main reasons to increase effort in the present setting probably are limits on sanctions (such as caused by the judgment-proof problem, which may be significant for single-product firms that substantially elevate price), problems of selective targeting (if firms in certain settings know they will not be prosecuted using current means, raising sanctions on firms in other settings does not help), and concern about chilling costs (greater

¹⁸⁷ See sources cited *supra* note 65 (suggesting that the probability may be in the 10–20 percent range).

effort, by yielding more cases with true positives that can be demonstrated, may enable an increase in proof burdens that reduces false positives and thus expected sanctions on firms that behave properly).

Private enforcement also deserves mention. Some jurisdictions provide for private suits. In this setting, enforcement effort is endogenous, determined by the expected payoffs from litigation compared to litigation costs. Higher damage awards,¹⁸⁸ laxer proof burdens, and other measures encourage litigation, which itself can be costly, although to the extent that it sufficiently augments deterrence, it may reduce costs because there would be fewer defendants whom it would make sense to sue. In practice, many (but not all¹⁸⁹) private suits for price fixing constitute follow-on litigation to successful government prosecutions, the effect of which is to augment government sanctions with private damage awards. Some further discussion of private litigation appears in section V.A.

As mentioned, sanctions may fail to be imposed both because of difficulty in initial detection and also because of difficulties of proof. Regarding the latter, raising sanctions to enhance deterrence is more problematic because, as noted, this policy also tends to increase chilling effects.¹⁹⁰ This point essentially restates the question of how optimally to determine both liability and sanctions when proof is uncertain to varying degrees.

¹⁸⁸ The fact that the level of damages directly affects private incentives to sue, but may not so directly affect government enforcement incentives (depending on how government enforcers are rewarded)—and, relatedly, that the resulting probability of sanctions differs in the two cases—constitutes one of the two central reasons that optimal damages may differ from optimal fines. The other is that victims ordinarily receive damages awards but not fine proceeds (although neither is inherent, as decoupling or nondisbursement of class action proceeds is possible and governments could distribute fine proceeds to victims if they wish). As a consequence, victims' incentives to avoid harm are influenced, and some have suggested that this may be an important adverse effect of private damages actions for price fixing. *See, e.g.*, Jonathan B. Baker, *Private Information and the Deterrent Effect of Antitrust Damage Remedies*, 4 J.L. ECON. & ORG. 385 (1988); David Besanko & Daniel F. Spulber, *Are Treble Damages Neutral? Sequential Equilibrium and Private Antitrust Enforcement*, 80 AM. ECON. REV. 870 (1990); William Breit & Kenneth G. Elzinga, *Antitrust Enforcement and Economic Efficiency: The Uneasy Case for Treble Damages*, 17 J.L. & ECON. 329 (1974); Stephen W. Salant, *Treble Damage Awards in Private Lawsuits for Price Fixing*, 95 J. POL. ECON. 1326 (1987).

¹⁸⁹ *See infra* note 197 (discussing facts about private antitrust enforcement in the United States, including the importance of private damage suits relative to government enforcement actions).

¹⁹⁰ Consider in this light Posner's interesting suggestion that, if anything, sanctions for more subtle forms of successful oligopolistic coordination should be higher, not lower (or zero, as many commentators advocate), because more subtle methods are harder to detect. *See* Posner, *Oligopoly*, *supra* note 7, at 1590–91. Perhaps the main reason this point is likely to be valid is that coordination achieved with the aid of highly explicit interfirm communications can be detected not only through such means but also through other observations, such as of price wars, price elevation, and the like, which can also be employed with respect to more subtle forms of coordination. However, if proof in more subtle cases tends to be more uncertain, there is a greater concern for chilling effects, which tends to favor moderated sanctions. It is unclear which factor is usually greater, and it is easy to imagine instances going in both directions.

3. *Law on Fines*

In the United States, private suits for damages measure harm as the overcharge—relative to a competition benchmark—times the quantity sold,¹⁹¹ which omits deadweight loss, and then award injured victims three times this amount, along with attorneys' fees.¹⁹² The multiplier, established a century ago, might be rationalized primarily in light of the difficulty of detection but also in part by the omission of deadweight loss, although the magnitude of other sanctions, notably government-imposed fines and imprisonment, are not reflected in such analysis.

One might have thought that fines would ordinarily be determined similarly, in accordance with the analysis in subsections 1 and 2. That is, the magnitude would be based on some fairly direct measure of harm—proportional to the magnitude of the overcharge times the units sold, perhaps with an addition for deadweight loss—and then multiplied by some factor reflecting the probability of sanctions. One could also imagine that there may be offsets allowed for other penalties so that sanctions as a whole met such a target.

In fact, the normal determination of fines in the United States and the European Union works rather differently. Although subject to reservations and some ambiguity, it appears that both law and practice base the fine on something closer to a fixed fraction of firms' revenues in the affected markets, with little or no adjustment to reflect the actual price elevation.¹⁹³ Although moti-

¹⁹¹ See, e.g., 2A PHILLIP E. AREEDA, HERBERT HOVENKAMP, ROGER D. BLAIR & CHRISTINE PIETTE DURRANCE, *ANTITRUST LAW* ¶¶ 392, 395 (3d ed. 2007); Kaplow, *supra* note 3, § III.C.4.

¹⁹² On the less developed use of private damages in the European Union, see Sergio Sorinas, *Remedies and Fines*, in 3 *EU COMPETITION LAW: CARTEL LAW: RESTRICTIVE AGREEMENTS AND PRACTICES BETWEEN COMPETITORS* 477, 584–86 (Mario Siragusa & Cesare Rizza eds., 2007) (noting that private damages are not allowed in Community courts but are left to member states, and that they have been rather limited, with the law highly varied and underdeveloped). For a proposal to allow actions for single damages, see COMMISSION OF THE EUROPEAN COMMUNITIES, *WHITE PAPER ON DAMAGES ACTIONS FOR BREACH OF EC ANTITRUST RULES* (2008).

¹⁹³ Fines for violations of U.S. competition law are governed by the Sentencing Guidelines. UNITED STATES SENTENCING COMMISSION, *2008 FEDERAL SENTENCING GUIDELINES MANUAL*, pt. R (Nov. 2008); see ABA SECTION OF ANTITRUST LAW, *SENTENCING GUIDELINES IN ANTI-TRUST: A PRACTITIONER'S HANDBOOK* (1999). Harm is measured by taking 20 percent of the volume of commerce involved. Use of the 20 percent factor, making no adjustment for higher or lower overcharge increments, seems to be widespread. See *id.* at 43. EU law is unclear regarding the extent to which the magnitude of the overcharge (however defined) is relevant in setting fines, either under the 2006 Fining Guidelines or in practice. See Commission Guidelines on the Method of Setting Fines Imposed Pursuant to Art. 23(2)(a) of Reg. No. 1/2003, 2006 O.J. (C 210) 2; *id.*, § 1.B.21 (general rule that “the proportion of the value of sales taken into account will be set at a level of up to 30% of the value of sales”); Sorinas, *supra* note 192; see also *id.* at 534 (“However, the 2006 Fining Guidelines make clear that cartel-type agreements . . . will ‘as a matter of policy, be heavily fined’. . . . Therefore, in cartel cases the percentage level can be expected to be equal or close to 30%, and the other factors mentioned in the 2006 Fining Guidelines will probably come into play and lead to a lower percentage only in exceptional circumstances.”).

vated on grounds of simplicity, this approach seems surprising in light of the huge variation in overcharges across cases.¹⁹⁴ Accordingly, setting to the side other penalties (which may be substantial in the United States because of private suits, but less so elsewhere), deterrence is likely to be highly inadequate when large overcharges occur since fines will be less than firms' profits, even ignoring any probability discount. This point may help to explain why, as section II.D presents, there are numerous prosecuted cases, many of which involve substantial price elevation.

The analysis of subsection 1 indicates that this approach is also highly problematic in cases involving small overcharges. The concern is not so much with overdeterrence per se as with chilling effects. The excessive application of fines is unnecessary to deter true violations involving small overcharges, and some such cases may involve false positives because, as discussed in section III.C and elsewhere, cases not involving actual violations are much more likely to present evidence of small price elevations than of large ones. Indeed, it was previously mentioned that low sanctions for low overcharges provide a partial safety valve against chilling costs by reducing the magnitude of expected sanctions in settings in which errors are more likely. These considerations suggest that, if anything, it might be best to moderate sanctions—relative to the best estimate of harm and an appropriate multiplier—when overcharges are low rather than imposing relatively high sanctions in such cases.

Another question of interest is how the current, standard level of fines relates to typical overcharges. Some argue that fines are too low on average, a position that is consistent with the aforementioned evidence of a high rate of detected and prosecuted violations.¹⁹⁵ As a consequence, collusion may indeed be profitable in some cases.¹⁹⁶ Keep in mind, however, that analyses limited to fines alone can be misleading if other sanctions, notably damages in private suits and imprisonment, are substantial.¹⁹⁷

¹⁹⁴ See, e.g., sources cited *supra* note 59.

¹⁹⁵ See CONNOR, GLOBAL PRICE FIXING, *supra* note 59; Connor & Lande, *supra* note 59. Previously, some commentators had expressed concerns about U.S. fines being overly severe. See, e.g., *id.* at 526–27 & n.79; Mark A. Cohen & David T. Scheffman, *The Antitrust Sentencing Guideline: Is the Punishment Worth the Costs?*, 27 AM. CRIM. L. REV. 331, 352–56 (1989). (It seems apparent from Connor and Lande's examination of prior surveys, *supra*, at 532 tbl. 1, that Cohen and Scheffman examined an unusually small number of cartels and found mean and median overcharges that are well below those reported in other surveys.)

¹⁹⁶ See, e.g., John M. Connor, *Global Cartels Redux: The Amino Acid Lysine Antitrust Litigation (1996)*, in THE ANTITRUST REVOLUTION: ECONOMICS, COMPETITION, AND POLICY 252 (John E. Kwoka, Jr. & Lawrence J. White eds., 2004) (Connor was plaintiffs' expert); see also Lawrence J. White, *Lysine and Price Fixing: How Long? How Severe?*, 18 REV. INDUS. ORG. 23 (2001) (White was defendants' expert).

¹⁹⁷ See, e.g., Robert H. Lande & Joshua P. Davis, *Benefits from Private Antitrust Enforcement: An Analysis of Forty Cases*, 42 U.S.F. L. REV. 879, 880 (2008) (finding "that almost half of the

B. IMPRISONMENT

As mentioned in the introduction to this Part, nonmonetary sanctions, notably imprisonment, tend to be inferior to fines or damages because of their greater social cost, and thus they are optimally used primarily when monetary sanctions are insufficient. When an individual or firm pays a fine or damages award, the funds are transferred, to the government or plaintiffs (often victims), so the sanction is not socially costly in itself.¹⁹⁸ With imprisonment, the utility loss suffered by those incarcerated—the prospect of which provides the deterrent effect—is a social loss not otherwise offset. In addition, it is socially costly to operate prisons. Despite these substantial disadvantages, imprisonment is increasingly used in the United States for price-fixing violations.¹⁹⁹

Two key rationales for the use of nonmonetary sanctions are relevant in the present context. The first is to make up for sanction insufficiency, notably, when defendants may be judgment proof. When fines or damages are limited to firms' profits, there would not often be a problem. But, as explained in section A, monetary payments should exceed profits and, moreover, they need to reflect a probability multiplier in light of the difficulty of detection, with treble damages being the norm in private suits in the United States. Accordingly, it seems plausible that firms that significantly elevate prices for years may not, when ultimately detected, have sufficient assets to pay the full fine or damages award, particularly for firms elevating price on their only product or products.

underlying violations were first uncovered by private attorneys, not government enforcers, and that litigation in many other cases had a mixed public/private origin," and concluding that "private litigation probably does more to deter antitrust violations than all the fines and incarceration imposed as a result of criminal enforcement by the DOJ"; *id.* at 893–95 (reporting that total payouts in just the cases in their sample exceeded total fines in all cases by more than four to one, and that limiting their sample of forty large cases just to those in which there were both private recoveries and criminal fines, the former were larger than the total of the latter for all cases, not just those in their sample); *id.* at 897, 909 tbl. 5 (of the approximately \$18 billion in total private recoveries, about \$8 billion were not from follow-on actions and an additional \$4 billion were from cases of mixed origin).

¹⁹⁸ This statement abstracts from the administrative costs of the system, including extracting the payments—the former of which are common for imprisonment and the latter of which are often negligible relative to the costs of prison—and also from risk aversion.

¹⁹⁹ *See, e.g.*, Scott D. Hammond, Recent Developments, Trends, and Milestones in the Antitrust Division's Criminal Enforcement Program, Presented at ABA Section of Antitrust Law Spring Meeting (Mar. 26, 2008), <http://www.usdoj.gov/atr/public/speeches/232716.htm>. Total imprisonment rose from an average aggregate of nine years of imprisonment per fiscal year in the 1990s to a record high of eighty-six aggregate years in 2007. *Id.* at 4. Average sentence length was eight months in the 1990s and has (roughly) trended upwards, reaching a record of thirty-one months in 2007. *Id.* at 6. *See also* Connor, *supra* note 65, at 91–92 tbl. 1. Outside the United States, however, imprisonment is ordinarily not authorized and is rarely used, and fines against individuals are also quite limited. *See* DIRECTORATE FOR FINANCIAL, FISCAL AND ENTERPRISE AFFAIRS, COMPETITION COMMITTEE, OECD, REPORT ON THE NATURE AND IMPACT OF HARD CORE CARTELS AND SANCTIONS AGAINST CARTELS UNDER NATIONAL COMPETITION LAWS 10–12 (2002). However, such policies are increasingly being revisited.

The second rationale, which is especially relevant to sanctions levied on particular individuals, relates to the agency problem within a firm.²⁰⁰ Most obviously, the threat of imprisonment (or fines on individuals) can usefully introduce a divergence between a firm's and its agents' incentives. In cases in which the firm is not adequately deterred, the prospect of imprisonment may nevertheless deter key employees, making it difficult for firms that wish to violate the law to do so.

Also important is the reverse situation, where it is the agent rather than the firm that may be the one inadequately deterred. For example, managers or sales personnel may receive compensation, including bonuses, and have promotion prospects tied to measures of performance. In such cases, they may have an incentive to engage in price fixing even when it is against the firm's interest. The reason for this possible difference in perspectives is that, even if expected sanctions are adequate to deter the firm, making it affirmatively in its interest to quash improper behavior, the firm may not be able to induce its agents to comply. They may calculate that there are large personal gains from violations, with low chances of detection. Moreover, if they are detected, such as when the firm is successfully prosecuted, the firm may not be able to heavily sanction the individuals. Usually, the most it can do is fire them, which may cost them future earnings as well as unvested deferred compensation. But these costs may not always be sufficient to deter the relevant individuals. In such cases, the firm would like to be able to sanction its misbehaving employees more severely, but it cannot. Thus, the prospect of government-imposed sanctions may provide a useful supplement.²⁰¹

To use a sanction of imprisonment (or fines levied directly on individuals), particular culpable individuals must be identified.²⁰² This need raises two challenges: determination of who in principle should be subject to imprisonment and detection of any such individuals in a particular case. These issues are

²⁰⁰ In the absence of agency problems and the judgment-proof problem, it tends not to matter in theory whether sanctions are levied on an entity or its agents, who can rearrange any such allocations by contract (such as indemnity provisions), unless those are outlawed. *See generally* Reinier H. Kraakman, *Corporate Liability Strategies and the Costs of Legal Controls*, 93 *YALE L.J.* 857 (1984).

²⁰¹ *See* A. Mitchell Polinsky & Steven Shavell, *Should Employees Be Subject to Fines and Imprisonment Given the Existence of Corporate Liability?*, 13 *INT'L REV. L. & ECON.* 239 (1993). As noted, the logic in the text justifies individual sanctions but not necessarily imprisonment; its justification will, again, typically be the judgment-proof problem, in this case with regard to individuals rather than firms.

²⁰² There is also the possibility of criminal liability imposed on the firm, but since imprisonment of a firm is not coherent and fines (whether or not deemed criminal) are already addressed, it is unclear how much independent significance remains. *See generally* V.S. Khanna, *Corporate Criminal Liability: What Purpose Does It Serve?*, 109 *HARV. L. REV.* 1477 (1996). A criminal conviction may, due to other rules, have independent consequences; for example, a firm convicted of bid rigging might be barred from future bidding, although such a prohibition could also follow as a result of civil liability.

generally present when attempting to sanction specific individuals in a legal entity for actions as agents in serving that entity. Accordingly, similar problems arise when firms illegally dispose of toxic waste, bribe government officials, violate safety regulations, and so forth.²⁰³

Because the cost of sanctions is greater and because the nature of chilling effects may differ when it is particular agents who are faced with sanctions—and their interests can diverge from firms' interests—different proof burdens may be optimal.²⁰⁴ And, of course, significantly higher proof burdens are ordinarily employed. This and other dimensions regarding the use of imprisonment warrant further study in the present context, and with regard to sanctions applied to agents more generally.

C. INJUNCTIONS²⁰⁵

As suggested in the introduction to this Part, injunctions are not obviously important in the present setting. Moreover, it is not apparent that they are heavily used.²⁰⁶ Nevertheless, many commentators seem to be fixated on their centrality—often implicitly, in arguing against a price-fixing rule on the

²⁰³ Even if price-fixing liability is confined to a narrow, precisely defined set of interfirm communications, there will be important questions of individual liability concerning individuals who are not, for example, caught on tape in the hotel room discussing prices, but who instead are their superiors or others engaged in price setting or strategizing who have knowledge of such behavior.

²⁰⁴ A particular concern is that, if individuals face a significant risk of heavy sanctions on account of false positives, chilling effects could be large. Even if the firm could absorb a certain level of expected sanctions imposed by mistake, more risk-averse individuals may not be able to do so. For this reason, individuals are often indemnified by firms, but some risk may remain (depending on the conditions of indemnification) and, if imprisonment is to be employed, financial indemnification may not be an entirely effective substitute (unless it is quite large).

²⁰⁵ Discussion is confined to orders to cease price fixing and related behavior. Structural relief is discussed briefly in note 208.

²⁰⁶ For example, exploration of the website and publications on workload and enforcement of the U.S. Department of Justice Antitrust Division reveals substantial information on fines and imprisonment, but not (that this author could readily locate) on injunctions. Examination of competition law treatises in the United States and other research does not reveal significant attention devoted to injunctions with regard to price fixing. It appears that injunctions are used with regard to explicit, open practices, such as when an organization requires price fixing or employs particular facilitating devices. For example, in civil antitrust actions against the airlines industry (*Airline Tariff*, for sending coded messages through the electronic price-posting service (United States v. Airline Tariff Publ'g Co., et al.)), NASDAQ market makers (for activity designed to maintain supracompetitive bid-ask spreads (United States v. Alex Brown & Sons, et al.)), and a firm engaged in FCC license auctions (for using price digits to send messages (United States v. Mercury PCS II, L.L.C.)), defendants were enjoined from the specific challenged practices (and also in general terms from agreeing to fix prices). For documents on the pertinent cases, see U.S. Dep't of Justice, Antitrust Division, Antitrust Case Filings, <http://www.justice.gov/atr/cases.html> (last visited Apr. 7, 2011). Most U.S. Department of Justice cases that involve price fixing are criminal, and injunctions do not appear to be employed.

ground that it would be difficult to fashion an injunction commanding compliance.²⁰⁷ Accordingly, it is important to examine injunctions further.

Injunctions do not deter violations. If the only legal consequence of liability for theft, murder, or tax evasion was being subject to an order not to repeat the violation in the future, crime might become rampant. Oligopolists would hardly be encouraged to forgo large profits by the prospect that, if they were caught, they might have to desist from continuing to earn supracompetitive rewards in the future. Instead, deterrence is accomplished through the use of fines and damages, along with imprisonment, if necessary. Once they are in place, the supplemental role of injunctions is unclear. Injunctions themselves could also be punitive, but injunctions are not ordinarily designed in this way, and such an approach tends to be inefficient.²⁰⁸

Ordinarily, injunctions are used to help ensure prospective compliance. With regard to coordinated oligopolistic pricing, it is unclear why injunctions are important for this purpose either.²⁰⁹ Why not simply rely on fines and damages, and imprisonment if necessary, in order to deter future violations? The basic logic is the same as that regarding ordinary deterrence: if the expected cost of violation exceeds firms' expected profits, they will be induced to comply. Indeed, even if there is an injunction against future violations, it too must be enforced. Presumably, firms will be led to comply with the injunction under the same conditions: when the expected sanction for violation is greater than the expected gain therefrom.

The foregoing raises the question of what, if anything, injunctions add to the preexisting legal command. Perhaps sanctions were previously insufficient to accomplish deterrence, and the penalties for violating an injunction are greater. But if this is so, one must ask why sanctions are not higher in the first

²⁰⁷ See, e.g., KAYSER & TURNER, *supra* note 104, at 143–44 (“By the very fact that we are dealing with practices—that is, conduct—the appropriate remedy is always the injunctive remedy: cessation of the practice.”); Turner, *supra* note 5, at 669; Jonathan B. Baker, *Identifying Horizontal Price Fixing in the Electronic Marketplace*, 65 ANTITRUST L.J. 41, 47 (1996).

²⁰⁸ Sanctions tend to be more efficient and effective when designed directly as such. The threat that firms might, for example, be forced by an injunction to deliver their goods through circuitous routes in the future would accomplish deterrence, but it would be better to fine them an amount equal to the cost of complying with the injunction and not require them actually to waste the resources. Some injunctions may impose costs indirectly, such as structural relief in the competition law setting. But, for the reason just given, this result tends only to make sense as a side effect. For example, if oligopolists were forced to deconcentrate, they would bear costs (in addition to forgone supracompetitive profits) if they lost scale economies, but that would be an undesirable rather than beneficial consequence of the relief, which would need to make sense in spite of this effect.

²⁰⁹ More generally, in contexts in which behavior generates negative externalities, fines or damages tend to be more efficient than injunctions because the government needs to know only about (or estimate) the level of harm (which information is required for any remedy) and not also about compliance costs. See generally Louis Kaplow & Steven Shavell, *Property Rules Versus Liability Rules: An Economic Analysis*, 109 HARV. L. REV. 713 (1996).

instance. Also, if there is some good reason, sanctions could be raised for repeat offenders—which is often done in the criminal law—without any separate need for an injunction.²¹⁰ Another aspect of the problem is that coordinated pricing may be undeterred when it is particularly difficult to detect. It is again unclear how an injunction would help; the industry could be subject to greater scrutiny in the future, but this could be done by enforcers and private plaintiffs in any event. Perhaps instead, firms might be subject to ongoing internal monitoring, making future furtive behavior more difficult. Note that in such cases the key difference made by the injunction would not be in the acts it prohibits or the sanctions it sets, but rather in the information it generates. Yet another possibility is that a government agency might set firms' prices or directly control other aspects of their behavior; that is, direct government regulation might be substituted for decentralized marketplace decision-making that operates under the threat of competition law sanctions.

Supposing that more ordinary injunctions were to be employed, consider what they might require. In other realms, injunctions and government regulations sometimes dictate specifics of behavior and in other settings focus primarily on results. Thus, a factory might be required to adopt a particular pollution-control technology or it might be ordered to meet a stated emissions target in any way it can. If the problem is waste discharges that contaminate groundwater, particular disposal techniques may be commanded or firms might be ordered to cease certain types of discharge, which they might choose to accomplish by changing technology, improving maintenance and monitoring, or halting production in a vulnerable location. With respect to coordinated oligopolistic price elevation, firms could be commanded what prices to charge (that is, something akin to public utility regulation) or to cease coordinating their behavior.

This latter possibility has caused great concern among commentators. The objection is that the failure to behave in a coordinated fashion, when conditions are conducive to success, would be somehow unnatural or irrational and hence futile to require—the asserted implication being that liability must be limited to the commission of specified acts, such as particular forms of interfirm communication.²¹¹ This complaint, however, is quite puzzling since

²¹⁰ Regulatory commands can be useful when firms may be judgment proof, but such commands must be enforced directly, such as through ongoing monitoring, discussed in the text to follow, rather than relying solely on probabilistic, after-the-fact detection.

²¹¹ See, e.g., AREEDA & HOVENKAMP, *supra* note 2, at 150 (“In the preceding hypothetical, however, we cannot reasonably expect firm *L* to refrain from increasing its prices when it feels that the market would accept them, or the others to refrain from following.”); *id.* at 231 (cannot employ “a legal rule [that] tells the oligopolist to close its eyes to the immediate and direct market impact of its own output choices, as well as to the subsequent market impact of its rivals’ probable responses to its own output decision”); *id.* at 232 (cannot employ a rule that “tells each firm to ignore the profit-maximizing signals emitted by the market”); MAHER M. DABBAH, EC

the response is both obvious and has been offered before. The simple point is that what is natural and rational depends on whether sanctions are imposed. It may be natural and rational for a hungry shopper to steal an apple, for youth with public artistic impulses to create graffiti, or for a spendthrift manager to embezzle funds—that is, if such were legal. However, if these activities are illegal and subject to heavy sanctions, engaging in them becomes irrational.

AND UK COMPETITION LAW 268 (2004) (cannot make illegal firms' behavior that constitutes profit maximization); EINER ELHAUGE & DAMIEN GERADIN, GLOBAL ANTITRUST LAW AND ECONOMICS 801–02 (2007) (referring to “the problem that firms [in] oligopolistic markets cannot avoid knowing their prices are interdependent when they set their prices, so that it would be hard to define any prohibition in a way that tells firms how to behave”); *id.* at 835 (“If so, how could one define the offense in a way that oligopolists could avoid behaving illegally? Is it practicable to ask them to ignore the reality of their price interdependence when making their pricing decisions?”); HERBERT HOVENKAMP, THE ANTITRUST ENTERPRISE 128 (2005) (objecting to the condemnation of interdependent behavior standing alone because “implicit in condemnation of any practice under the antitrust laws is that the defendant was obliged to behave in some other way than it did”); WHITE HOUSE TASK FORCE ON ANTITRUST POLICY 5 (1968), *reprinted in* Antitrust & Trade Reg. Rep. (BNA) No. 411, spec. supp. II (May 27, 1969) (“antitrust law . . . cannot order the several firms to ignore each other's existence”); SCHERER & ROSS, *supra* note 58, at 342 (“How should oligopolists change their behavior so as to avoid breaking the law? Must they begin ignoring their interdependence in pricing decisions, when to do so would be irrational?”); SIGRID STROUX, US AND EC OLIGOPOLY CONTROL 114 (2004) (“Imposing competitors to disregard their rival[s'] behaviour would obviously be nonsense, as it would require them to behave irrationally.”); IVO VAN BAELE & JEAN-FRANCOIS BELLIS, COMPETITION LAW OF THE EUROPEAN COMMUNITY 51 (4th ed. 2005) (“[S]ometimes it is only rational commercial behaviour which makes competitors align their conduct. In such a case undertakings should not be punished for doing what makes sense commercially.”); Edward H. Chamberlin, *Duopoly: Value Where Sellers Are Few*, 44 Q.J. ECON. 63, 65 (1929) (“Each is forced by the situation itself to take into account the policy of his rival in determining his own”); Giorgio Monti, *The Scope of Collective Dominance Under Articles 82 EC*, 38 COMMON MKT. L. REV. 131, 145 (2001) (arguing that remedies are infeasible as long as there exists a rational economic explanation for the oligopolists' behavior); Turner, *supra* note 5, at 666 (“Particularly is this so when the behavior involved, setting the ‘profit-maximizing’ price in light of all market facts, is not only legally acceptable but vitally necessary to make competitive markets function as they are supposed to function.”). Interestingly, most of these references postdate Posner's work (in one case, co-authored) that is cited in the note to the following paragraph, yet they seem to overlook its basic point with regard to the deterrent effect of sanctions affecting what firms would find rational to do.

Even ignoring that sanctions change firms' incentives, the sometimes-expressed view that interdependence is inevitable that is reflected in the aforementioned views—*see also, e.g.*, KAYSER & TURNER, *supra* note 104, at 27 (suggesting that recognized interdependence is “extremely likely” when concentration is even moderately high)—is not empirically well grounded. Evidence on successful prosecutions, discussed in section II.D, reveals the use of explicit and sometimes highly elaborate direct communications even in very concentrated industries, and industrial organization research for decades on the structure-conduct-performance paradigm, discussed in subsection III.A.2, suggests that successful coordination often fails even in highly concentrated industries. Moreover, most commentators who have offered an opinion on the subject (overlapping in part with those cited just above in this note) assert that coordination is actually quite difficult in the absence of explicit communications. *See* sources cited in Kaplow, *supra* note 3, at 762–63 n.192. Accordingly, both the empirical basis and apparent consensus behind the premise for this common argument seem on reflection to be lacking. Indeed, for those who believe that successful coordinated price elevation through pure recognized interdependence is challenging or impossible in most settings, it is hard to understand the basis for the concern in the text, which presumes that coordination is so attractive and straightforward that it is virtually irresistible.

Oligopolistic price coordination is no different. If there are no sanctions, a firm may find it attractive to follow a leader's price increase and to avoid undercutting the industry's supracompetitive price because of the allure of sustained oligopoly profits. But if such acts are associated with sufficiently high penalties, then the firm would find it irrational to follow the leader's price increase and profitable to undercut elevated prices. Therefore, the question of whether a command to refrain from coordinated oligopolistic pricing—whether issued through a particular injunction or by the law more generally—will succeed depends on the adequacy of expected sanctions. This adequacy, in turn, depends on detection and on the magnitude of sanctions.²¹² It does not depend on firms' managers possessing some mystical ability to engage in a form of reasoning unknown to humankind.

V. ADDITIONAL CONSIDERATIONS

A. INSTITUTIONAL ISSUES

The framework for decision-making presented in section II.C largely abstracts from the administration of competition policy. Enforcement institutions vary widely across jurisdictions and sometimes within them—regarding the latter, especially in the United States, which has public and private suits, state and national enforcement, and two distinct systems at the national level. These differences are potentially important in themselves, and they also may influence what rule is optimal—and vice versa, which is to say, if certain rules are notably better in principle, it may accordingly be desirable to make institutional choices that best facilitate implementation of those rules. This section is confined largely to identifying issues rather than resolving them, both because there exist independent literatures on institutional design and because the control of successful oligopolistic coordination is but one of the tasks of a competition regime. It is useful to organize thinking around three features of the administrative apparatus: enforcement effort, error rates, and cost.

Enforcement effort is particularly important in the present setting because of the difficulty of detection and resulting concerns about underdeterrence, as

²¹² See POSNER, ANTITRUST, *supra* note 7, at 97–98 (“Tacit collusion is not an unconscious state. . . . The threat of a damages judgment for supracompetitive pricing will influence their pricing decisions; what would be irrational would be for the oligopolists to ignore such a threat.”); RICHARD A. POSNER & FRANK H. EASTERBROOK, ANTITRUST: CASES, ECONOMIC NOTES AND OTHER MATERIALS 333 (2d ed.1981) (“Oligopolistic interdependence is not an unconscious state. If tacit collusion can be detected at all in an industry, and can be proved at trial, there should be no special difficulty in proving that the same practices have continued despite an injunction against them. The defendants can alter their behavior even if the behavior was once in their interest. Sanctions for disobedience change the incentives. The question is: Can tacit collusion be proved in the first place?”); Posner, *Oligopoly*, *supra* note 7, at 1592 n.80 (“All I am arguing is that a deliberate restriction of output by competitors is conduct that rational men can avoid—and will avoid if it is made sufficiently costly to them to engage in it.”).

suggested by the evidence reviewed in section II.D and discussed further in subsection IV.A.3. As mentioned, higher sanctions are a substitute for enforcement effort, but there are important limits. In the United States, where enforcement is generally thought to be the strongest, there are private suits for treble damages as well as public enforcement, potentially by numerous government entities, with the possibility of significant fines and imprisonment. Allowing private damages suits motivates victims who may have access to information not readily available to the government; such information could be conveyed to officials, but the incentive to develop the information may be less without the prospect of compensation.²¹³ Private suits are also useful because they provide a check against laxity or industry capture of government enforcers. Government also may have advantages, such as those due to economies of scale and scope, a potential willingness to investigate when private parties may not be in a position to capture enough reward to motivate them, and the ability to employ certain enforcement instruments (for example, wiretaps and the use of informants) and sanctions (notably, imprisonment).

Each approach also has costs, an important one being the possibility of overzealous enforcement. Private parties have incentives to bring unmeritorious cases if there is a sufficient probability of factfinder error or, in some circumstances, if adjudication costs are sufficiently high, in both cases in an attempt to extract significant settlements.²¹⁴ Likewise, government enforcers, eager to advance their careers or augment their agencies' budgets through high-visibility prosecutions, can also be overly aggressive.²¹⁵ For these reasons, the best choice among possible enforcers and the optimal way to incentivize them are hardly obvious.²¹⁶

Error rates and costs are determined both by who brings enforcement actions (and how they are motivated) and by the system of adjudication. In most countries (and in the United States for actions brought by the Federal Trade Commission), expert agencies are employed, typically with appeals brought in generalist courts. Direct enforcement in generalist courts is another possibil-

²¹³ See *supra* note 197 (presenting facts about private enforcement in the United States).

²¹⁴ The main limit concerns the credibility of going forward, which may not be present if prospects of success are low and the plaintiffs themselves would bear significant litigation costs. See, e.g., Kathryn E. Spier, *Litigation*, in 1 HANDBOOK OF LAW AND ECONOMICS 259, 305–07 (A. Mitchell Polinsky & Steven Shavell eds., 2007).

²¹⁵ Overenforcement can also arise due to protectionism.

²¹⁶ See, e.g., Gary S. Becker & George J. Stigler, *Law Enforcement, Malfeasance, and Compensation of Enforcers*, 3 J. LEGAL STUD. 1 (1974); William M. Landes & Richard A. Posner, *The Private Enforcement of Law*, 4 J. LEGAL STUD. 1 (1975); A. Mitchell Polinsky, *Public Versus Private Enforcement of Fines*, 9 J. LEGAL STUD. 105 (1980); Steven Shavell, *The Optimal Structure of Law Enforcement*, 36 J.L. & ECON. 255 (1993). An additional issue is the need to coordinate sanctions, which is rendered more complex when sanctions may result from actions of different enforcers operating through different systems. For example, in the United States, treble damages are not reduced if defendants have already paid substantial fines to the government.

ity, employed in the United States for private plaintiffs and for the Antitrust Division of the Department of Justice—with the further idiosyncratic feature that juries are often used as factfinders.²¹⁷ It is worth noting that none of these match-ups is inevitable: one could, for example, use expert agencies to adjudicate private suits, or have government enforcement actions adjudicated by the same entity as the initiator, a distinct agency, a specialized court, or a generalist court.

The virtues of expert adjudication of one sort or another are obvious. In most areas of competition law, and certainly including oligopolistic coordination, complex economic issues are involved, and the mass of relevant information that must be assessed in a given case is often vast. There is frequent but not widespread use of specialized adjudicators, but often not with the relevant expertise. In particular proceedings, decision-makers may be guided by parties' (partisan) experts or by adjudicator-appointed (purportedly neutral) experts.²¹⁸ Use of the latter, in whole or in part, seems highly advantageous in light of the potential for bias by paid experts,²¹⁹ yet such use seems atypical,

²¹⁷ U.S. courts' distrust of juries, whether on account of a pure inability to understand the complex subject matter or also due to a fear of anti-defendant bias, helps to explain the tendency in recent decades to raise procedural hurdles for plaintiffs and, relatedly, to substitute judges' decisions for juries' by resolving more cases through legal motions decided by the court. Similar concerns may also help to explain Supreme Court reversals of jury verdicts in fully tried cases. *See, e.g., Brooke Grp. Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209 (1993); *Spectrum Sports v. McQuillan*, 506 U.S. 447 (1993). One of the difficulties posed by this approach is that the judges are themselves generalists, often with negligible relevant experience, and trial judges in particular may be biased in favor of defendants regarding dispositive motions because of the significant effect on their workload caused by even a single, significant antitrust case. Or they may err in favor of plaintiffs, hoping that cases will settle, thereby avoiding appellate review.

²¹⁸ For example, in the United States, Federal Rule of Evidence 706 authorizes court-appointed experts.

²¹⁹ Consider, for example, F.M. Scherer's criticism of Posner's more direct, economically oriented approach toward oligopoly pricing.

Granted, with *enough* strong evidence of this type, a competent economist could in good conscience testify that even though no meetings in smoke-filled rooms were proved, it was quite unlikely prices could have been set as observed had there been no tacit restrictive understanding. The trouble is, economic analysis is an elastic instrument and, I am sorry to report, some economists' consciences are also elastic, so one can find economists who with apparent conviction will explain away any pattern of behavior, however bizarre, as the consequence of special but highly competitive industry circumstances. Sometimes they may even be right, if there exists any absolute measure of "right" in such complex matters. Every tacit collusion case under Posner's scheme would be a "big case," drawing teams of economists to ply the courts with their expert but conflicting opinions. In the end, the decision would turn significantly upon whose experts were more credible. It would not, I fear, be a system highly likely to yield either truth or justice, especially when private respondents pay \$1,000 per day for "credibility" (including extensive preparation) while the Government is limited to \$150 or (in exceptional cases) \$250.

Scherer, *supra* note 8, at 983 (note that the daily rates, updated to the present, translate more closely to hourly rates).

for reasons that are not apparent. When a single agency serves as investigator, prosecutor, and adjudicator, their experts have a mixed character, perhaps fairly neutral at the outset when attempting to identify whether an investigation should be pursued (although maybe not if the political stakes are high), but possibly partisan once in adjudication.

The nature of the tribunal as well as the selection of experts should have a direct and, one would suppose, substantial impact on the quality of outcomes: greater and more neutral expertise should enhance deterrence and reduce chilling effects because both false negatives and false positives should be fewer. The main potential tradeoff concerns cost, for higher expertise is often more expensive. However, the opposite may be true in the present context. A battle of purely partisan experts before a decision-maker with little or no expertise may well, in addition to being less accurate, be more costly as each party invests substantially in attempts to sway the tribunal and presentations are drawn out because even basic concepts must be explained rather than taken as part of presumed knowledge. Moreover, the greater unpredictability that may arise and resulting higher costs have the further undesirable consequence of facilitating meritless challenges, which itself worsens the problem of chilling effects.

It is useful to examine the particular problem of controlling oligopoly pricing with attention to different enforcement stages. For competition agencies, it is possible that different methods might be better suited at different points in the process. For initial identification of cases to investigate more thoroughly, agencies might, in addition to relying on buyer complaints and informants, employ economists who would study industry data on pricing, market shares, and the like, as well as features bearing on the ease of coordination, to identify where violations are most likely to be present—and this is so independently of what ultimately must be proved in an adjudication to establish liability.²²⁰

Another stage-related challenge involves the formal screening of cases by an adjudicator. Because of the aforementioned concerns about overzealous enforcement, with concomitant litigation costs and chilling effects, it is desirable for meritless cases to be eliminated before undertaking long, expensive proceedings—involving both pre-trial work, including investigation or discovery and preparation of expert reports, and trials. This consideration explains the increasing emphasis in the United States on granting motions to

²²⁰ See, e.g., POSNER, *ANTITRUST*, *supra* note 7, at 55, 64, 69 (emphasizing that, even if the law is interpreted narrowly, enforcement agencies should use broader economic criteria to decide where to focus their investigations); Rosa Abrantes-Metz & Patrick Bajari, *Screens for Conspiracies and Their Multiple Applications*, *ANTITRUST*, Fall 2009, at 66; Harrington, *supra* note 70. For skepticism about the prospects for economic screening by agencies, see Scherer, *supra* note 8, at 982–83.

dismiss at the outset of case filings in court and motions for summary judgment before proceeding to trial.²²¹ Unfortunately, this task is far easier said than done precisely because of the elaborate and complex nature of the evidence and analysis involved. Because so much of the total expense of adjudication is borne before an actual trial commences (investigation, discovery, experts), the greatest benefits are realized if meritless cases can be dismissed up front. The problem is that, at the outset, there is very little basis for ascertaining whether cases of the sort under examination here have merit.²²²

Because of these inherent difficulties with early dispositions, it is appealing to consider some more explicit sequencing (i.e., bifurcation) of litigation than is ordinarily undertaken. For example, if it is desired, on one hand, not to dismiss nearly every case, and, on the other hand, not to allow highly costly proceedings when cases may well be meritless, a compromise might be struck under which there is some preliminary, partial factfinding, perhaps based primarily on publicly available information. For example, parties (or the tribunal) might engage the services of a limited number of experts to offer initial opinions on the likelihood of successful oligopolistic coordination and on the conduciveness of industry conditions. If the initial showing is sufficiently weak, the case would be dismissed, but if it is reasonably strong, the case would be permitted to proceed. It is speculative how reliable an assessment could be made based on limited evidence, and in designing such a system, it would have to be decided just how much or little effort—use of experts, access to defendants' pricing data, and so forth—would be permitted. However, it might be possible to eliminate a significant portion of meritless cases at well under the full cost of complete proceedings. Moreover, if such is indeed possible, meritless cases would much less attractive to initiate because their threat value in extracting significant settlements would be reduced.

For competition regimes that rely on specialized agencies, it would seem that such procedural reform may well be possible. Indeed, it already exists to an extent, since agencies can drop an investigation at any point if they think it lacks merit. It may still be useful, however, to permit defendants—parties

²²¹ The most pertinent cases are *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544 (2007), on motions to dismiss, and *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574 (1986), on summary judgment.

²²² Even if the use of explicit communications is a prerequisite to liability, as long as proof by circumstantial evidence is allowed, an enforcer (private party or government agent) would be in a position to allege that, taking into account all manner of evidence concerning whether the market in question is conducive to successful oligopolistic coordination and whether such successful coordination is taking place, such is likely to have been employed. The U.S. Supreme Court in *Twombly* acted as if it had some way to assess such potentially complex matters at the motion-to-dismiss stage, but what that might be does not seem possible to extract from the opinions in that case. An interesting possibility would be to disallow circumstantial evidence, which would largely limit liability to cases in which an informant could be obtained at the outset. See Kaplow, *supra* note 9, § 4.1.

under investigation—to instigate such a preliminary determination, particularly in light of the potential problem of overzealous government enforcers. For proceedings in court, it may require more substantial reform to enable such an approach, although aspects might be implemented through more subtle means.²²³

It remains to consider the disposal of cases after all investigation, discovery, and preparation of expert reports, but before trial. This possibility is especially important in the absence of the just-described, more preliminary proceeding. However, it too is more difficult than is usually imagined. The problem is that, unless a party's case has negligible evidentiary support, it seems difficult for an adjudicator to find that a trial is unnecessary without in essence conducting the trial to determine whether the evidence is sufficiently strong. In the United States, courts often grant defendants' pre-trial motions for summary judgment in antitrust cases, and a reading of the opinions suggests that they in essence do resolve substantial factual disputes that, in principle, they are supposed to avoid (by denying the motion and allowing trial to proceed).²²⁴ Thus, in one respect, they do behave in the manner just described. One limitation on achievable cost savings is that such decisions come only after many costs are sunk. Also, because judges are not supposed to be behaving in this manner, they cannot organize proceedings in a way that best enables them to make the requisite decisions.²²⁵

In sum, issues of institutional design are quite important for competition regulation because of the scope and complexity involved. Costs are high and the risk of false positives producing chilling effects and false negatives undermining deterrence are great. Moreover, there is a tension between enforcement intensity, which is necessary to address the deterrence problem, and high administrative costs, including chilling effects that arise when well-behaved firms can effectively be threatened with protracted litigation. As mentioned at

²²³ In the United States, Federal Rule of Civil Procedure Rule 42(b) allows for separate trials of different issues, and Rule 52(c) for partial findings as a trial proceeds (when a judge is the factfinder). Both, however, contemplate already being at the trial stage. It may still be possible, however, through deft use of pretrial conferences and by requiring plaintiffs to make more substantial showings as a prerequisite to obtaining fuller discovery from objecting defendants, for judges to manage cases and provide periodic indications of their leanings in response to information that is presented. It also seems possible that the use of court-appointed experts and special masters could help perform similar functions.

²²⁴ Consider, for example, the cases on antitrust conspiracies discussed in Kaplow, *supra* note 3, at 743 n.151.

²²⁵ Sometimes courts also dismiss cases at the outset when factual disputes exist, essentially inferring from the weakness of a plaintiff's complaint that it seems unlikely that its case is substantial. To the extent that this practice becomes more common after *Twombly*, perhaps plaintiffs will begin to offer expert reports and other evidence at this early stage, even though such proffers are supposed to be unnecessary. At that point, the motion to dismiss would itself, to an extent, become somewhat like the compromise procedure discussed just above in the text.

the outset of this section, these issues are relevant to the choice of legal rule. Also, if one legal rule is in principle substantially superior to another, that preference may well bear on how the supporting institutions and procedures ought to be designed.²²⁶

B. LIABILITY FOR ATTEMPTS

In some instances, the available evidence may establish that firms attempted to coordinate their behavior to elevate price but not show that they succeeded. Proof regarding success may be ambiguous even though the attempt is fairly clear, or coordination may be demonstrated to have failed. Many sorts of evidence, including much of that discussed in subsections III.A.1 and III.B.1, bear directly on the occurrence of successful oligopolistic coordination. However, certain evidence, including some involving the existence of interfirm communications addressed in subsection III.B.2, may relate directly to firms' attempts.²²⁷ For example, evidence of a secret meeting at which it was decided that the firms would jointly elevate price does not establish that they succeeded (and there may be evidence that they did not).²²⁸ Such cases raise the question of whether it is optimal to punish attempts.²²⁹

At first, it may seem that punishing attempts is unnecessary and possibly undesirable, and sometimes this view is correct. Such punishment is costly, which cost may be unproductive in light of the fact that no harm occurred. Furthermore, there may be mistakes in identifying attempts, the prospect of

²²⁶ To restate one aspect of this point in the U.S. context, there is, to a degree, a choice between employing highly simplified but necessarily poor rules in order to adapt to the use of lay juries (and, arguably, generalist judges) or instead crafting more sensible rules that may require restructuring how enforcement takes place. Also of great importance in the United States, the allowance of private treble damages actions, even if helpful for deterrence, adds greatly to potential total administrative costs as well as to concerns about chilling effects. It may make more sense to limit or abolish private suits, making up the slack with substantially higher fines (*see supra* subsection IV.A.3) than to adopt a legal rule that renders permissible much of the most socially damaging oligopoly pricing out of the concern that a more open-ended approach would induce excessive private litigation.

²²⁷ Also related is the treatment of facilitating practices, addressed in subsection III.A.1.c, particularly if they are themselves to be subject to sanctions—that is, independent of their providing a basis for inference regarding successful coordination. In a sense, facilitating practices can be seen as attempts or as evidence that an attempt is contemplated.

²²⁸ However, when there is evidence of ongoing meetings—to discuss cheating, price and market share adjustments, and the like—it may be clear that successful oligopolistic coordination has characterized the past and present, in which case the meetings would constitute evidence of success.

²²⁹ *See generally* Steven Shavell, *Deterrence and the Punishment of Attempts*, 19 J. LEGAL STUD. 435 (1990). Interestingly, Posner criticizes the commonly advanced view that the law should focus on explicit communications (which are a means to an end and evidence thereof often proving an attempt but not success) rather than on the undesirable outcome (oligopolistic price elevation) as one that punishes attempts while excusing success itself. *See* POSNER, *ANTI-TRUST*, *supra* note 7, at 53–55 (referring to “the law’s emphasis on the attempt to fix prices rather than on economic effects”).

which could chill desirable behavior. Indeed, if there is no evidence of success or if there is affirmative evidence of failure, perhaps we should be skeptical that there really was an attempt. When masked individuals enter a bank with guns drawn, but ultimately need to flee without any money in hand, we have little doubt about what they were attempting. However, when attempts are inferred from ambiguous internal documents or from meetings whose existence seems likely but whose content is unknown, the prospect of error may be great. Of course, sometimes there will be fairly clear evidence of an attempt, perhaps a tape of the secret price-fixing session.

The benefits of penalizing attempts—which are widely punished in the criminal law—typically concern deterrence. An alternative to punishing attempts would be to raise sanctions for success. However, when success is difficult to prove, as it often is for price fixing, and when there are limits to sanctions that can be imposed, punishment of attempts helpfully augments deterrence. It may also reduce costs: if some attempts cases are easy to establish, enforcers might be able to forgo other more difficult and expensive cases while still maintaining deterrence. In addition, when attempts are clear and sufficient for liability, costly inquiries into success may be avoided—although evidence of success may nevertheless be required to set sanctions.

Another important consideration involves chilling costs. All else equal, it is optimal to raise penalties for more clearly undesirable acts while reducing sanctions for more ambiguous ones. If sometimes attempts are clear and they are subject to significant penalties, lower sanctions can be applied in more ambiguous cases of success, or some might be excused as a consequence of an elevated burden of proof made possible (while maintaining deterrence) by the punishment of attempts. Taken together, these reasons suggest that punishment of attempts at oligopolistic price coordination probably makes a good deal of sense, particularly when the evidence is strong.

An important caveat, previously suggested by the discussion of the relevance of the conduciveness of conditions, is that imposing sanctions on attempts that fail (or where success is unknown but seems unlikely) may not much enhance deterrence in realms where it is most socially valuable, while still imposing chilling costs. The point is that deterrence is not very important in industries where success is highly unlikely. Few firms would attempt oligopolistic price elevation in any event, and even those that do so may be almost certain to fail. Moreover, given that few will try, it may be more likely that findings of apparent attempts involve false positives. The strength of this reservation depends on the quality of the evidence bearing both on attempts and on failure. Thus, when conditions are conducive and success is merely hard to prove, punishing attempts is probably valuable. Moreover, even when the foundation is weaker, there may be some deterrence benefits, for firms uncertain about success will be better deterred. In addition, deterrence may suffer

from the prospect that firms that make undesirable attempts might be exonerated when success cannot be established and they might convince an adjudicator that success was unlikely or absent. After all, the firms themselves may have a better view about the *ex ante* likelihood of success, and their very attempt may attest to their belief that it was at least plausible that they would succeed. Put another way, firms' attempts, when they are subject to the risk of sanctions, themselves provide evidence of the plausibility of success, assuming of course that the evidence on the attempts themselves is substantial.

Another question concerns sanctions for attempts. Ordinarily, penalties should be set to reflect the harm caused, as discussed in section IV.A. For failed attempts, harm is zero, and for attempts of ambiguous outcome, harm is uncertain. Sanctions for attempts, accordingly, will need to reflect prospective harm. More precisely, one should consider the preceding analysis of the benefits of punishing attempts in augmenting deterrence while reducing chilling costs and set penalties, for both attempts and in cases involving demonstrated success, to maximize social welfare in light of these considerations. Although a mere truism, it does guide thinking: we know which factors make punishing attempts more valuable, and these same factors tend to favor stronger sanctions as well.

C. UNILATERAL MARKET POWER

The analysis of oligopoly thus far focuses on distinguishing two categories of behavior: purely independent behavior—taken to be synonymous with competition, where firms take prices as given—and interdependent behavior—where firms take rivals' reactions into account and thereby are able to elevate price.²³⁰ This section examines a further distinction within the independent-behavior category: between competitive behavior—where rivals equate marginal cost to a given price—and interactions in which firms, although taking rivals' behavior as given, nevertheless elevate price through what is referred to as the unilateral exercise of market power. In the latter case, firms find it profitable to sell at a price above rather than equal to marginal cost, without assuming that other firms will respond by raising their prices, and to maintain price in excess of marginal cost rather than slightly undercutting others' prices, even if they do not take into account any retaliation by rivals. This characterization merely restates the point that the exercise of unilateral market power entails price elevation without the sort of coordination (interdependence) described in section II.A.²³¹

²³⁰ See Kaplow, *supra* note 3, §§ I.B, II.A.

²³¹ The use of the term unilateral market power and its contrast to coordinated oligopolistic behavior (interdependence) on one hand and competition on the other hand is conventional in economics but is often confused in competition law commentary (including some written by economists), where the term "independence" is sometimes used to encompass interdependence—

This article has largely proceeded as if such unilateral market power did not exist. When it might, a number of questions arise. First, what treatment of unilateral price elevation maximizes social welfare? This topic is the analogue to that addressed in section II.B, where it is implicitly assumed that oligopolistic price elevation is due to coordination. Second, how can we distinguish price elevation attributable to the exercise of unilateral market power from that due to successful oligopolistic coordination? The answer to this question is important if the response to the first question implies that different treatment is optimal in the two cases. As will be discussed, many of the methods of detection considered in sections III.A and III.B specifically identify interdependence whereas others identify price elevation without necessarily suggesting what mode of interaction produced it. Third, in situations in which it is difficult to distinguish the two causes of price elevation, what approach to liability makes sense in light of the social costs arising from this possibility of misclassification? These errors are different from errors of, for example, mistakenly treating competitive behavior, with price equal to marginal cost, as if there was oligopolistic price elevation. Accordingly, ex ante behavioral effects tend to differ as well, and sometimes the effect of anticipating mistaken liability in this setting is to improve rather than reduce social welfare.

These questions have received virtually no attention in the competition policy literature on price fixing and related matters,²³² although there is an extensive industrial organization literature on unilateral market power and consideration of such power plays a significant role in analyzing which horizontal mergers should be challenged.²³³ Accordingly, the analysis in this section is preliminary. It considers industries with multiple leading firms for cases with homogeneous goods and with differentiated products, the latter corresponding to the case often referred to as monopolistic competition.²³⁴

that is, both behavior in which rivals' behavior is taken as given and strategic behavior that takes into account rivals' anticipated reactions. *See id.* § I.B.5.

²³² Richard Markovits—in a series of papers that, to a significant extent, are a critique of Turner, *supra* note 5, and Posner, *Oligopoly*, *supra* note 7—emphasizes that price can exceed marginal cost without coordination. *See, e.g.*, Richard S. Markovits, *Oligopolistic Pricing Suits, the Sherman Act, and Economic Welfare: Part I: Oligopolistic Price and Oligopolistic Pricing: Their Conventional and Operational Definition*, 26 STAN. L. REV. 493 (1974); Richard S. Markovits, *Oligopolistic Pricing Suits, the Sherman Act, and Economic Welfare: Part III: Proving (Illegal) Oligopolistic Pricing: A Description of the Necessary Evidence and a Critique of the Received Wisdom About Its Character and Cost*, 27 STAN. L. REV. 307 (1975). Posner's otherwise critical reply accepts this point. *See* Richard A. Posner, *Oligopolistic Pricing Suits, the Sherman Act, and Economic Welfare: A Reply to Professor Markovits*, 28 STAN. L. REV. 903, 911–13 (1976). But neither author substantially illuminates the questions considered here.

²³³ *See* U.S. Dep't of Justice & Fed. Trade Comm'n, Horizontal Merger Guidelines §§ 6, 7 (2010) (distinguishing between unilateral and coordinated effects); European Comm'n, Guidelines on the Assessment of Horizontal Mergers Under the Council Regulation on the Control of Concentrations Between Undertakings, 2004 O.J. (C 31) 5, at ¶¶ 24–28.

²³⁴ Unilateral exercise of market power by a single, dominant firm (a monopoly) is examined briefly in subsection II.B.2.b.

Suppose that a market is inhabited by five symmetric firms, each having a 20 percent market share. The homogeneous goods case is most relevant for the analysis of interdependent oligopolistic price elevation because coordination is more difficult when products are differentiated. In analyzing this case, economists distinguish between whether firms are understood to compete on price or on the quantity of goods sold.²³⁵ In the former case, if firms have constant marginal cost, the result will be competitive, which is to say that there can be no unilateral exercise of market power. If prices are at the competitive level, any firm that raises price loses all of its sales to rivals, which sell the identical good and maintain their lower prices. And if prices were ever elevated, a firm that cut its price slightly would take the entire market, greatly increasing its profit at the others' expense. This result should not be surprising; after all, it is believed that many markets with at least a few firms do behave quite competitively, and that coordination would be required if the firms are to do better for themselves. There are further complications when marginal costs are rising (or there are capacity constraints) that will not be examined further here.²³⁶

Accordingly, economists generally focus on quantity competition, drawing on Cournot's seminal discussion.²³⁷ In the most plausible formulation of this story, a two-stage process is envisioned. Firms choose their quantities in an initial period, after which these quantities are all sold in a subsequent period, in which price is determined to clear the market. When each firm chooses its quantity, it is assumed to take other firms' quantity choices as given. This formulation implies some price elevation. If a firm considers reducing its quantity somewhat, this deviation would decrease the total quantity supplied (other firms, recall, are assumed to keep their quantities constant), which in turn implies a higher market-clearing price. If initially the firm was behaving competitively, so that price equaled marginal cost, this higher price will increase its profits. All firms are assumed to behave likewise, resulting in a supracompetitive equilibrium price. (This elevated price, however, is below the monopoly price.)

In this model, the equilibrium price is elevated, but the exercise of market power is unilateral. If the firms coordinated their behavior, price would be

²³⁵ See FRIEDMAN, *GAME THEORY*, *supra* note 13, at 54–57; TIROLE, *supra* note 13, ch. 5; VIVES, *supra* note 13, chs. 4, 5; SHAPIRO, *supra* note 13, at 333–56.

²³⁶ Formal solutions using mixed strategies are presented in Partha Dasgupta & Eric Maskin, *The Existence of Equilibrium in Discontinuous Economic Games, II: Applications*, 53 *REV. ECON. STUD.* 27 (1986), and Eric Maskin, *The Existence of Equilibrium with Price-Setting Firms*, 76 *AM. ECON. REV. (PAPERS & PROCEEDINGS)* 382 (1986). However, some analysts find the mixed-strategy solution inapt in terms of understanding actual strategic interactions among firms. See, e.g., James W. Friedman, *On the Strategic Importance of Prices Versus Quantities*, 19 *RAND J. ECON.* 607, 608, 616 n.13 (1988); Shapiro, *supra* note 13, at 346 & n.30.

²³⁷ See, e.g., sources cited *supra* note 235.

elevated further, perhaps to the monopoly level, but even with no coordination (unilateral action, taking other firms' quantities as given) there is some price elevation. Although this model is commonly used, there are questions about its realism. Specifically, once firms are in the second stage, why do they not raise their quantity, reducing their price somewhat, to increase their profits? The strongest answer is that there are some settings—and therefore, the Cournot model should be viewed as apt only in those cases—in which such second-stage competition cannot take place. Specifically, these involve capacity constraints, which make price cuts unattractive because each firm is unable to sell more, so the price cut can only reduce its profits.²³⁸ If firms do have excess capacity, then sustained supracompetitive prices would be attributable to interdependence.²³⁹

Viewing each of the above cases, with competition in prices and in quantities, it is not clear how often there will be significant unilateral exercise of market power with homogeneous goods. Supposing, however, that such price elevation does occur, let us return to our main questions. Initially, we can inquire into whether such price elevation is socially valuable. In the cases posed, it would seem that it is not. Static, allocative efficiency suffers. Regarding dynamic efficiency, we can consider whether such supracompetitive pricing is necessary to induce entry. As explored in subsection II.B.2.a, with homogeneous goods entry tends to be excessive when prices are elevated, with the degree of excess rising with the extent of elevation. The optimal amount of entry ordinarily arises when price equals marginal cost. The main exception is when fixed costs make it difficult to sustain even one or two firms in an industry, in which case, as noted, some price elevation may be socially desirable.

The prospect of supracompetitive pricing might also encourage other sorts of investment that perhaps reduce production costs or improve product quality or service, as discussed in subsection II.B.2.b. If one of our five firms makes such investments, however, it will tend to be rewarded without regard to the sort of price elevation under discussion. For example, a firm with lower costs can reduce price below that of its higher-cost competitors, gain market share, and earn positive margins on such additional sales as well as on sales it previously enjoyed. To be sure, it might profit even more with further elevated

²³⁸ See, e.g., TIROLE, *supra* note 13, at 218 (“More generally, what we mean by quantity competition is really a *choice of scale that determines the firm's cost functions and thus determines the conditions of price competition*. This choice of scale can be a capacity choice, but more general investment decisions are also allowable.”).

²³⁹ Firms also might coordinate in the first stage on capacity reduction, the feasibility of which may be eased by the fact that capacity expansions are often difficult to keep secret and take time to implement, although some firms may also try to preempt others, undermining coordination.

prices—or it may not, if price elevation serves to protect its rivals and thus limit the firm's ability to increase market share.

In sum, it seems that the exercise of unilateral market power by multiple firms in homogeneous goods industries tends to be detrimental (in contrast to the conclusion with dominant firms presented in subsection II.B.2.b). One might therefore wish to condemn such supracompetitive pricing along with that due to successful oligopolistic coordination. Suppose, however, that for one reason or another, such behavior does not give rise to liability. (It generally does not under existing competition law, although horizontal mergers that enhance such market power are condemned.)

Can we identify cases in which the unilateral exercise of market power is taking place? Often, it seems we could. For example, the Cournot story supposes that firms are selling at capacity. We can also consider how interdependence itself may be identified—as distinct from the unilateral exercise of market power—the subject of Part III.²⁴⁰ Finally, when there is nontrivial uncertainty about whether price elevation is unilateral or due to successful oligopolistic coordination, the cost of mistakenly concluding that the latter occurred is not great. Indeed, as just mentioned, unilateral price elevation with homogeneous goods tends to involve both allocative and dynamic inefficiency, so false positives (deeming to be interdependent price elevation that is really unilateral) may actually be beneficial rather than socially costly.

Now suppose instead that the five firms in our previous example, each with a market share of 20 percent, produce differentiated rather than homogeneous products. Analysis of firms' interactions in this case differs because it is no longer true that pricing slightly below rivals' prices captures the entire market.²⁴¹ Instead, as one firm reduces its price, taking rivals' prices as given, consumers only gradually substitute toward that firm's product, away from those of rivals. It follows that some unilateral market power exists in this setting, which is to say that firms will elevate their prices above marginal cost even when they take rivals' prices to be constant. To see this, suppose that all firms initially price at marginal cost. A firm that increases its price slightly does lose some sales to rivals, but it also earns more on those customers it retains. Since it was previously assumed to be pricing at marginal cost, the lost sales at first entail no sacrifice in profit, so at least some price increase will be profitable. Eventually, however, further elevation will become unprof-

²⁴⁰ Note, however, that direct measurement or other evidence of marginal cost indicating that it is below price does not distinguish interdependence from unilateral price elevation. But if in such cases one insists that price be shown to be substantially in excess of marginal cost, for the reasons adduced in section III.C, then one may have de facto exempted much unilateral market power in industries in which there are more than a few firms since unilateral market power in such settings is likely to be modest.

²⁴¹ See sources cited *supra* note 36.

itable. The situation is similar for the other firms, and their price increases in turn tend to make it profitable for the initial firm to increase its price further, and so on. At the resulting equilibrium, all firms charge prices above marginal cost (but, as in the homogeneous goods case, still below the monopoly price).

Our first question is whether such price elevation is desirable. As usual, there is allocative inefficiency due to the fact that price exceeds marginal cost. An important dynamic effect is to encourage entry. In the homogeneous-goods case, this effect is undesirable, but, as explained in subsection II.B.2.a, with product differentiation there is the benefit of increased product variety. When price equals marginal cost, variety tends to be too low. In the equilibrium involving firms' unilateral exercise of market power, variety could be either too high or too low. (In particular settings it may be possible to tell which.) The previous analysis explained why variety tends to be too little at the competitive price, with the benefits of variety falling and the degree of inefficiency rising as price is elevated further and further above marginal cost. Thus, it is probably unwise to adopt a broad rule condemning unilateral price elevation in this context, and it is ambiguous in the general case whether some reduction in price (say, through a chilling effect) would be desirable or detrimental.

Second, consider whether it is possible to distinguish the unilateral exercise of market power from elevations that may be due to successful oligopolistic coordination. Many means of detection identify interdependence per se and thus do distinguish the cases. Other sources of evidence (such as on marginal cost²⁴² or internal documents) may or may not be helpful, depending on the particulars.²⁴³

One of the main criteria helping to distinguish the sources of price elevation in this situation is the assumption of differentiated products itself. When differentiation is low, unilateral market power will tend to be negligible from this source. Moreover, interdependence may well be feasible. When differentiation is high, unilateral market power will be more substantial on this account, but interdependence tends to be infeasible. In other words, this

²⁴² It is again worth recalling that if, as section III.C suggests, liability based primarily on evidence of marginal cost being below price would only be assigned in cases of significant elevation, unilateral market power would be de facto exempted when product differentiation was not too large (which is precisely the cases that would be targeted, as noted in the text to follow). In addition, it was already explained that the benefit of price elevation in reducing any product variety deficit is likely to be small (or reversed, a cost of excess variety) when the price elevation is large.

²⁴³ For econometric studies that distinguish unilateral from coordinated exercise of market power, see Jonathan B. Baker & Timothy F. Bresnahan, *Estimating the Residual Demand Curve Facing a Single Firm*, 6 INT'L J. INDUS. ORG. 283 (1988), and Aviv Nevo, *Measuring Market Power in the Ready-to-Eat Cereal Market*, 69 ECONOMETRICA 307 (2001).

particular dimension of conduciveness will be particularly probative of what is taking place.

Finally, supposing that we try to distinguish the two sources of price elevation and might err, particularly by producing false positives, we should consider the social costs entailed because these will bear on how we should set proof burdens. As explained, some reduction in price when the true source of elevation is the exercise of unilateral market power tends to have ambiguous welfare effects in this case. Accordingly, it may then be important to attempt to further distinguish among settings involving product differentiation to see whether a positive or negative effect is more plausible. In any event, observe that the typical social cost of this sort of false positive does seem to be less (if indeed it is a net social cost) than that of errors in settings in which the market is fully competitive, with no actual price elevation taking place.

VI. CONCLUSION

Determination of an optimal regime for coordinated oligopolistic price elevation requires attention to the social objective, the problem of detection, and the design of sanctions. The purpose of the regime is to deter such price elevation, both because of the direct, static costs of supracompetitive prices and because of dynamic effects, notably, inducing excessive entry in homogeneous goods industries, which are those in which substantial coordinated price elevation is most likely. Because of difficulties with detection, achieving effective deterrence is not easy, and empirical evidence suggests that deterrence is inadequate even in the United States, which is regarded to have the toughest enforcement in terms of both overall (public and private) effort and the level of sanctions.

The appropriate framework needs to be attentive not only to errors of omission (false negatives, where violators are not sanctioned) but also to mistakes of commission (false positives, where innocent firms are sanctioned or subject to significant costs imposed by investigations and litigation). The nature of this tradeoff is considered explicitly, with special attention to the nature of chilling costs, that is, the undesirable *ex ante* effects on behavior attributable to the prospect of false positives. The likely magnitude of such costs varies by the nature of the industry and the type of proof.

Successful oligopolistic coordination may be detected through a number of means, ideally considered in combination. Its presence or absence may be demonstrated by market-based evidence, notably pricing patterns and indications of price elevation, and also by the use of facilitating practices. The degree to which industry conditions are conducive to collusion is also highly relevant. In a sense, conduciveness is a necessary condition, but not a sufficient one—among other reasons, because of the possibility that deterrence is

effective. Interfirm communications, the emphasis of much prior analysis of the subject, are also probative but hardly necessary, both as a logical matter and because proof may be lacking even when they are present. All bases for assessment may be established, reinforced, or negated by evidence internal to firms, including many sorts that do not speak directly to coordination. For example, if the main alternative explanation for a sudden price increase is a cost shock, firms' internal records may well indicate the likelihood and magnitude of such an event.

Because detection is the central challenge in addressing coordinated oligopolistic price elevation, this subject receives the most extensive attention, with a particular focus on the reliability of various types of proof and the risks of chilling effects. For example, demonstration of price elevation by comparison of price to marginal cost is more precarious than some other methods due to the danger of underestimating marginal cost; hence, it may be appropriate to require stronger evidence or proof of a greater magnitude of price elevation when employing this technique, as well as corroborating evidence.

The analysis of sanctions focuses on fines and private damage awards because monetary sanctions are heavily used and tend to be the most efficient sanction to the extent feasible—subject, notably, to the limit of firms' assets. As in most areas of law enforcement, sanctions should reflect harm. Large price elevations, which are disproportionately harmful, need to be adequately deterred, which requires high sanctions because firms' profits are greater. Cases that apparently exhibit small price elevations are more likely to involve false positives and thus generate chilling costs, so moderating expected sanctions is valuable, and less social cost from forgone deterrence results. Government fines in the European Union and the United States do not reflect this prescription very well, although private damages do. Furthermore, because the detection rate is low, multipliers need to be large, raising the possibility that the optimal magnitude of sanctions may be significantly higher than the level presently employed.

Imprisonment, used increasingly and chiefly in the United States, provides a helpful supplement in achieving deterrence due to the possible inadequacy of fines and damages. Furthermore, because of agency problems in firms, sanctions, including imprisonment, applied to individual violators may be valuable. Injunctions do not seem to be widely used and do not appear to be important in principle because they fail to generate deterrence and must be enforced through other sanctions in any event. Nevertheless, many writers seem to be preoccupied with injunctions and believe that rules of liability need to be crafted, and possibly significantly constrained, by the feasibility of injunctions of a particular type, tantamount to command and control regulation.

Some additional considerations are also noted. Particularly important are institutional issues, for the efficacy of any competition regime depends strongly on the cost and quality of the method of implementation, particularly in light of the economic complexity involved. Liability for attempts is seen to be valuable in light of the difficulty of achieving adequate deterrence. A pervasive complicating factor is the possibility that price elevation in an oligopolistic industry is due to firms' exercise of unilateral market power. Coordinated elevation is most likely in homogeneous goods industries, so that case is particularly relevant; it appears that unilateral market power is less likely and that it may well be possible to identify when its exercise is plausible. Moreover, the potential chilling costs from misclassification may not be that great—and may even be negative (that is, benefits) because moderating prices can be desirable in such industries.

This investigation is preliminary in a number of respects. First, there has been limited prior analysis of much of the relevant territory, especially from the current perspective. Second, a number of key choices, at least as a matter of degree, raise empirical questions. Although relevant evidence is examined, on many dimensions quite little is known, particularly with regard to chilling effects. Third, the interplay with enforcement institutions needs further attention. It may be, for example, that the optimal regime for a system with exclusive government enforcement using an expert agency is quite different from that for one allowing private suits adjudicated by generalist judges and lay juries. Despite these significant reservations, it does appear that wholesale rethinking is in order.

Most commentators—legal analysts and economists, in the European Union and the United States—believe that current law is best described as, and should be, a communications-based prohibition, limiting liability to cases in which the use of particular forms of communication have been employed. The predecessor to this article²⁴⁴ explores in detail how consensus characterizations of existing practice are difficult to state operationally, differ greatly from what many advancing them seem to contemplate, and conflict with much of what actually seems to take place.²⁴⁵ More relevant for policy purposes, this article offers a systematic, ground-up analysis of the social problem that turns out to differ even in basic subject matter coverage (not just in possible bottom lines) from most prior work on the topic.²⁴⁶ Careful examination of the social

²⁴⁴ See Kaplow, *supra* note 3.

²⁴⁵ See *supra* note 4.

²⁴⁶ The leading three arguments offered in favor of the traditional view can be related explicitly to the analysis in this article. One argument asserts a difficulty in attacking purely interdependent behavior because such would involve commanding firms to behave irrationally. See sources cited *supra* note 211. This criticism is mistaken because it omits considerations of deterrence; that is, applying heavy sanctions to certain choices will change what firms find rational to do. See *supra*

objective, the central challenge of detection, and the matter of appropriate sanctions does not generate the commonly favored rule, indicating that it indeed is not optimal. Furthermore, a sequel explicitly compares the traditionally supported prohibition to the economic approach developed here.²⁴⁷ It finds that the former requires more complex and uncertain economic inquiries to implement and tends to focus liability where deterrence benefits are low and chilling costs high, providing exoneration in sets of cases in which social harm is great and chilling costs more modest.

This body of work constitutes a strong critique of existing modes of thought, even though it does not advocate for a particular legal rule for the reasons just mentioned. It seeks to provide a sound framework for analysis and to advance our understanding on many subsidiary questions to the extent possible. It also indicates areas where further empirical research could valuably illuminate the subject.

As a final note, it should be mentioned that this article's criticism should not be seen as primarily directed at courts—and it is only secondarily a criticism of enforcement agencies—even though the analysis suggests that current approaches are misguided. It is principally the responsibility of academics and other analysts to assess policy, at which point it is their role and that of government agencies and advocates to bring the teachings to bear in implementation. This article seeks to contribute to, and in certain respects restart, a long-dormant policy debate on perhaps the most important aspect of competition regulation. It is hoped that the resulting efforts will increasingly guide the design of legal regimes.

section IV.C. Another objection is that making price elevation by oligopolists illegal is inconsistent with the legality of price elevation by monopolists. *See* sources cited *supra* note 42. This point ignores the purpose of separate, more stringent prohibitions on group behavior and, moreover, if accepted, would imply that classic cartels should be legal. *See supra* subsection II.B.2.b. Third, it is argued that remedies, particularly injunctive relief, directed at price elevation are problematic because they amount to price regulation. *See, e.g.,* Turner, *supra* note 5, at 669 (referring to a “purely public-utility interpretation of the Sherman Act”); AREEDA & HOVENKAMP, *supra* note 2, at 206 (noting “the absence of a practical remedy other than judicial price control”); *id.* at 232–33 (suggesting that the required injunction would be “equivalent to . . . compelling marginal-cost pricing” which “puts the antitrust tribunal directly in the price-control business”); ELHAUGE & GERADIN, *supra* note 211, at 835 (“Could courts figure out whether they had done so other than by asking whether the prices were reasonable?”); Baker, *supra* note 207, at 47 (“the only remedy is judicial price regulation—a complete non-starter”); Scherer, *supra* note 8, at 984. This claim is misconceived because, as mentioned, effective control may be accomplished through penalties that achieve deterrence rather than by relying on directive legal commands. *See supra* section IV.C. However, this argument is suggestive of an important concern with competition policy regarding price fixing that is undeveloped in the existing literature, namely, that the detection of violations can be quite difficult, raising the problem of false positives, the prospect of which chills desirable behavior. As explained especially in Parts II and III, these concerns are indeed central in devising price-fixing rules, but the analysis does not show that they imply the desirability of the conventional approach.

²⁴⁷ *See* Kaplow, *supra* note 9.

