
ARTICLE
WHY (EVER) DEFINE MARKETS?

Louis Kaplow

TABLE OF CONTENTS

I.	INTRODUCTION.....	438
II.	THE MARKET DEFINITION / MARKET SHARE PARADIGM.....	443
	<i>A. Preliminaries</i>	443
	<i>B. Analysis</i>	447
	1. Dominant Firm, Homogeneous Good, Small Rivals	448
	2. Market Redefinition	453
III.	STANDARD REFERENCE MARKET	459
	<i>A. Motivation</i>	459
	<i>B. Concept of a Standard Reference Market</i>	462
IV.	DEFINING THE RELEVANT MARKET	465
	<i>A. Choosing the Best Market Definition</i>	467
	1. Best as Most Accurate Measure of Market Power	468
	2. Best as Yielding the Right Legal Answer	473
	<i>B. What to Do?</i>	474
V.	ELASTICITY VERSUS CROSS-ELASTICITY	480
	<i>A. Relationship Between Elasticity and Cross-Elasticity</i>	482
	<i>B. Implications</i>	484
	1. Summation	484
	2. Revenue Weights	486
	3. Evidence Availability	490
	<i>C. Are Cross-Elasticities Ever Useful?</i>	491
	<i>D. Reflections</i>	493
VI.	ADDITIONAL TOPICS.....	495
	<i>A. Rivals' Elasticity of Supply</i>	496
	<i>B. Market Power and Condemnation</i>	498
	<i>C. Is Market Definition Ever Needed?</i>	502
	<i>D. The Allure of Market Definition</i>	506
	<i>E. The Law on Market Definition</i>	508
VII.	CONCLUSION.....	515

WHY (EVER) DEFINE MARKETS?

Louis Kaplow^{*}

Competition law is dominated by the market definition / market share paradigm, under which a relevant market is defined and pertinent market shares therein are examined in order to make inferences about market power. This Article advances the immodest claim that the market definition process is incoherent as a matter of basic economic principles and hence should be abandoned entirely. This conclusion rests on four arguments. First, meaningful inferences of market power in redefined markets cannot be made. Second, the paradigm relies on an unarticulated notion of a standard reference market whose necessity and prior omission signal a serious gap. Third and most important, determining what market definition is best is impossible without first formulating a best estimate of market power, rendering further analysis pointless and possibly leading to erroneous outcomes. Finally, the need to define markets engenders a mistaken focus on cross-elasticities of demand for particular substitutes rather than on the market elasticity of demand, which further reduces the quality of resulting market power inferences. Although the inquiry is conceptual, brief remarks on legal doctrine suggest that creating conformity may not be unduly difficult.

I. INTRODUCTION

In competition law cases, the most frequently decisive factor is the degree of existing market power or the extent to which it is enhanced as a consequence of the practice under scrutiny, whether an alleged exclusionary act or a horizontal merger. The importance of market power is well established under U.S. antitrust law, where it is a significant consideration under nearly all but a subset of per se doctrines (notably, on price fixing), and it plays a roughly similar role in competition re-

^{*} Finn M.W. Caspersen and Household International Professor of Law and Economics, Harvard Law School, and Research Associate, National Bureau of Economic Research. I am grateful to participants in workshops at Berkeley, Harvard, Stanford, the Department of Justice Antitrust Division, and the Symposium on Antitrust Economics and Competition Policy hosted by Northwestern University's Searle Center on Law, Regulation, and Economic Growth for helpful discussions and comments; Simon Franzini, Daniel Jacobson, Summer Kim, Gopi Panchapakesan, Ravi Ramchandani, and Brian Samuelson for research assistance; and Harvard Law School's John M. Olin Center for Law, Economics, and Business for financial support. This project has its origins in a seminar taught in 1979, a *Harvard Law Review* article published in 1982, the revision of the text and preparation of the teacher's manual on market power and market definition for Antitrust Analysis (with Phillip Areeda) in 1988, and an unpublished lecture presented to antitrust lawyers at the National Association of Attorneys General in 1990. Some hints also appear in my 2007 Handbook of Law and Economics antitrust chapter written with Carl Shapiro. I have benefited greatly from prior interchanges with coauthors, but neither they nor others who have offered comments should be held responsible for any views expressed herein (some of which conflict in part with what I have suggested previously).

gimes of the European Union and other jurisdictions.¹ Market power, in turn, is most often assessed under the market definition / market share paradigm, making market definition the most litigated issue in the field.²

The first step under the market definition / market share paradigm is to define a so-called relevant market. This market definition process involves choosing from among candidate markets that which most accurately depicts the extent of market power. The method can usefully be described as one that starts from the homogeneous goods market that includes the firm's or firms' product and then considers whether to redefine the market by broadening it to include substitute products (or regions, for geographic market definition). Next, one measures the firm's market share in that market or the collective share of a group of firms, such as when evaluating a horizontal merger, joint venture, or trade association activity. Finally, one infers from this share the degree of market power and assesses it against the pertinent legal standard. Although the market definition / market share paradigm is not the only way to evaluate market power — notably, there exist more direct techniques, some technical and others more qualitative — the process is normally employed and thought by some to be mandatory. In any

¹ See, e.g., 1 ABA SECTION OF ANTITRUST LAW, ANTITRUST LAW DEVELOPMENTS 550–54 (Jonathan M. Jacobson et al. eds., 6th ed. 2007) [hereinafter ANTITRUST LAW DEVELOPMENTS]; ABA SECTION OF ANTITRUST LAW, MARKET POWER HANDBOOK 5–9, ch. 2 (2005) [hereinafter MARKET POWER HANDBOOK]; *DG Competition Discussion Paper on the Application of Article 82 of the Treaty to Exclusionary Abuses* ¶ 23 (Dec. 2005) [hereinafter *DG Competition Article 82*], available at <http://ec.europa.eu/competition/antitrust/art82/discpaper2005.pdf>.

² See, e.g., *Eastman Kodak Co. v. Image Technical Servs., Inc.*, 504 U.S. 451, 469 n.15 (1992) (“Because market power is often inferred from market share, market definition generally determines the result of the case.”); ANTITRUST LAW DEVELOPMENTS, *supra* note 1, at 549 (“Defining a relevant market is often a critical issue, and sometimes the critical issue, in an antitrust case.”); 2B PHILLIP E. AREEDA, HERBERT HOVENKAMP & JOHN L. SOLOW, ANTITRUST LAW 135 (3d ed. 2007) (“In resolving market or ‘monopoly’ power issues, the courts have typically relied heavily on market definition and on the defendant firm’s share of the market thus defined.”); Jonathan B. Baker, *Market Definition: An Analytical Overview*, 74 ANTITRUST L.J. 129, 129 (2007) (“Throughout the history of U.S. antitrust litigation, the outcome of more cases has surely turned on market definition than on any other substantive issue.”); Robert Pitofsky, *New Definitions of Relevant Market and the Assault on Antitrust*, 90 COLUM. L. REV. 1805, 1807 (1990) (“Knowledgeable antitrust practitioners have long known that the most important single issue in most enforcement actions — because so much depends on it — is market definition.”). For the European Union, see, for example, Commission Notice on the Definition of Relevant Market for Purposes of Community Competition Law, 1997 O.J. (C 372) ¶ 4 [hereinafter Commission Notice on Market Definition], which notes that “[t]he definition of the relevant market in both its product and its geographic dimensions often has a decisive influence on the assessment of a competition case.” See also Thomas E. Kauper, *The Problem of Market Definition Under EC Competition Law*, 20 FORDHAM INT’L L.J. 1682 (1997) (comparing the approaches in the European Union to those in the United States, more often emphasizing similarities than differences, particularly with regard to monopolization and mergers).

event, other approaches are usually seen as subservient, providing information to guide market definition rather than to replace it.

The thesis of this Article is that the market definition process should be abandoned. The central, conceptual argument is that there does not exist any coherent way to choose a relevant market without first formulating one's best assessment of market power, whereas the entire rationale for the market definition process is to enable an inference about market power. Why ever define markets when the only sensible way to do so presumes an answer to the very question that the method is designed to address? A market definition conclusion can never contain more or better information about market power than that used to define the market in the first place. Even worse, the inferences drawn from market shares in relevant markets generally contain less information and accordingly can generate erroneous legal conclusions — unless one adopts a purely results-oriented market definition stratagem under which one first determines the right legal answer and then announces a market definition that ratifies it. Additional, largely unavoidable difficulties are identified with the economic logic underlying market redefinition. Because virtually all of the argument reveals inherent problems in the very conception of the market definition / market share paradigm, it follows that the conclusions here do not depend on one's assessment of the quality of various means of measuring market power either in general or in particular cases³ and that they are independent of the legal application at hand.

Prior criticism of the market definition / market share paradigm is extensive. Defects have been identified by courts, enforcement agencies, and both legal and economic commentators. No one believes that the market definition process is flawless or that market power inferences drawn from market shares are uniformly reliable, or even nearly so. Some favor alternatives in particular settings, either to complement or to replace the market redefinition approach. But prior work, individually and collectively, does not constitute a wholesale assault on the core logic of the methodology along the lines offered here.⁴ This characterization is reflected in the 2009–2010 U.S. Horizontal Merger

³ For example, they do not depend on whether a particular econometric technique is sound or can readily be applied in light of data limitations, or on what sorts of internal documents are found, or on how savvy purchasers are in predicting their responses to hypothetical price increases. This Article takes no position on how often adjudicators or agencies should employ more sophisticated econometric techniques, rely on informal methods of assessing the quality of substitutes, and so forth. It addresses only the use to which all such information should be put.

⁴ Leading prior critiques of market definition are discussed in notes 78, 79, and 80 (at which point they can more meaningfully be contrasted with the present analysis, especially that in Part IV), and past treatments of cross-elasticities (which suggest but do not develop some of the points in Part V) are presented in note 83. Some previous mentions by U.S. courts of market definition's limitations are noted in section VI.E.

Guidelines revision process in the questions posed for commentary, the comments submitted, and the final revised Guidelines.⁵

Part II begins by articulating the basic economic logic that underlies the market definition / market share paradigm. Market power is defined, and a standard formula relating market share to market power is presented. This formula, it is emphasized, properly applies only in a homogeneous goods market. Accordingly, the analysis focuses on how it may be extended to redefined markets, notably, when one or more substitutes are added to the original, narrow market. It turns out that the only way to make economically meaningful inferences about market power from market share in this redefined market — something that courts and commentators purport to do routinely — is to undo the market redefinition. Therefore, it is hard to make sense of existing practice, and the only route to correct the deficiency is to stay with the narrow, homogeneous goods market in all cases, that is, never to redefine markets. This approach, moreover, is entirely sound in that valid inferences about market power can be made in this market, a claim that is true regardless of how high or low a firm's (or firms') market share is or how many good substitutes of whatever strength are available.

Part III introduces the notion of a standard reference market. It explains that market power inferences based on market shares — which are necessary to speak intelligibly about market definition and to reach conclusions about market power once the relevant market is selected — presume the existence of some common benchmark. For example, a statement that a market share in a narrow market is misleadingly high or that a firm's lower market share in a broader market better depicts the situation presumes that those shares mean something. In order to operationalize such statements and help make intelligible past discussion of the subject, one may construct a standard reference market. However, reflection on the essential need for the concept and its absence in prior analysis reveals another vexatious hole in the paradigm.

Part IV advances the central criticism that it is impossible to choose the relevant market coherently unless one already has one's best estimate of market power in hand, in which case the entire exercise is pointless, or worse. Two criteria for market definition are considered, the most natural one (in terms of rationalizing current prac-

⁵ See *Horizontal Merger Guidelines Review Project*, FED. TRADE COMM'N, <http://www.ftc.gov/bc/workshops/hmg/index.shtml> (last visited Oct. 30, 2010) (links for the initial questions and comments); Public Comments on Revision to the Horizontal Merger Guidelines, FED. TRADE COMM'N, <http://www.ftc.gov/os/comments/hmgrevisedguides/index.shtm> (last visited Oct. 30, 2010); U.S. DEP'T OF JUSTICE & FED. TRADE COMM'N, HORIZONTAL MERGER GUIDELINES (2010) [hereinafter U.S. MERGER GUIDELINES].

tice) being that the best market is that which yields the most accurate inference about market power. Reflection on this standard reveals the problem. To determine whether market definition *A* involves less error than does market definition *B*, one needs to have a view about the magnitude of these errors. Now, these errors consist of the deviation between the market power inference derived under one or the other market definition and one's best estimate of actual market power. Therefore, assessing the errors, which is necessary to choose the relevant market, presupposes that one already has formulated one's best estimate.⁶ Hence, defining markets is useless.

Even worse, using the best market to make inferences, as just explained, involves error — not just error in some absolute sense, but compared to the best estimate with which one began. As a consequence, the resulting inference is generally worse than where one started, which implies that legal conclusions will be inferior to those one would have reached if the market definition process had been eschewed. Part IV examines another criterion that avoids this latter problem of excess error but also makes particularly stark the senseless, question-begging character of the market definition process.

Part V explores an additional mistake induced by the market definition / market share paradigm: the focus on cross-elasticities rather than on the market elasticity of demand. The market definition process asks us to determine whether the additional good in market *B* is a strong enough substitute for the good in market *A* that the two markets should be combined. This question, in turn, directs attention to the cross-elasticity between the two goods — the (percentage) rate at which changes in the price of the good in market *A* cause consumers to switch to the product in market *B*. However, the definition of market power in Part II and the standard formula relating market power to market share make it apparent that the market elasticity of demand, rather than particular cross-elasticities, is what is directly relevant to the ultimate question of market power. The analysis in Part V explains two respects in which this divergence produces misleading conclusions. The Part does note that cross-elasticities are sometimes quite helpful — notably, in evaluating horizontal mergers between firms producing differentiated products — but, under the proper analysis, the pertinent cross-elasticities are not used to define markets.

Part VI considers a number of additional topics. The discussion throughout, which often focuses on product market definition for concreteness, is linked to rivals' elasticity of supply. Matters pertaining to

⁶ As Part IV makes clear, the argument does not depend at all on the quality of this estimate of market power or the manner in which it was derived (for example, whether through econometric analysis or the conjectures of customers).

the definition of market power that are often conflated with rules for legal condemnation are clarified. The point that market definition is never needed is explored with respect to standard applications involving monopolization and horizontal mergers. The allure of market definition, despite its shortcomings, is examined. And finally, the law on market definition is related to the present critique. Although this is a conceptual paper on what makes sense as a matter of competition policy, a number of considerations are presented that suggest (but do not purport to establish) that existing legal doctrine may not offer a significant constraint on abandonment of the market definition / market share paradigm.

II. THE MARKET DEFINITION / MARKET SHARE PARADIGM

The essential reason for defining a market and examining a firm's market share in that market is to make inferences about market power. Section A elaborates this point, defines market power, and indicates why it is often thought useful to employ the market definition / market share paradigm to illuminate market power rather than relying on other means of inference. Section B explains the economic logic underlying the relationship between market share in a given market and market power and discusses how this logic depends on the manner in which a market is defined. In redefined markets, the inference process is seen to break down.

A. Preliminaries

It is uncontroversial that the core rationale for defining or redefining markets is to enable inferences about market power.⁷ Sometimes it is the prevailing level of market power that is relevant, as is said to be typical in monopolization cases under Sherman Act section 2 or EU Article 102 (formerly 82). The extent of market power may be thought directly relevant to assessing anticompetitive effects or may be seen as a useful screen — that is, a threshold inquiry prerequisite to further

⁷ See, e.g., AREEDA, HOVENKAMP & SOLOW, *supra* note 2, at 232 (“Finding the relevant market and its structure is typically not a goal in itself but a mechanism for considering the plausibility of antitrust claims that the defendants’ business conduct will create, enlarge, or prolong market power.”); Phillip Areeda, *Market Definition and Horizontal Restraints*, 52 ANTITRUST L.J. 553, 565 (1983); Baker, *supra* note 2, at 130; Commission Notice on Market Definition, *supra* note 2, ¶ 2 (“The main purpose of market definition is to identify in a systematic way the competitive constraints that the undertakings involved face. . . . It is from this perspective that the market definition makes it possible *inter alia* to calculate market shares that would convey meaningful information regarding market power for the purposes of assessing dominance or for the purposes of applying Article [101 (formerly 81)].”); see also Gregory J. Werden, *The History of Antitrust Market Delineation*, 76 MARQ. L. REV. 123, 189 (1992) (suggesting that the idea that “market delineation must be based on the underlying concern with market power was spreading rapidly in the early 1980s but certainly was not universal”).

consideration, to reduce the frequency with which single-firm practices need to be scrutinized and to mitigate the extent of false positives. Sometimes, the emphasis is on changes in market power, often a focus in assessments of horizontal mergers.⁸ (Observe that market power changes may be determined by comparing levels of market power before versus after, or with versus without an act or practice; accordingly, the language of levels will generally be used for convenience.⁹) This Article will not directly examine the purposes and proper role of market power requirements in competition regimes.¹⁰

If the point of the market definition process is to enable inferences about market power, it is important to identify what we mean by market power. The most frequent definition, and perhaps the easiest to make explicit, is the power to profitably¹¹ elevate price¹² above the competitive level, which in the case of perfect competition would be a price equal to firms' marginal costs (which would be equated through

⁸ See, e.g., U.S. MERGER GUIDELINES, *supra* note 5; Guidelines on the Assessment of Horizontal Mergers Under the Council Regulation on the Control of Concentrations Between Undertakings, 2004 O.J. (C 31) 5.

⁹ As discussed in section V.C, some particular methods do focus on changes in market power without being directly concerned with levels, such as in the use of diversion ratios in the analysis of the unilateral effects of horizontal mergers in industries with product differentiation. See, e.g., Joseph Farrell & Carl Shapiro, *Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition*, B.E. J. THEORETICAL ECON., Jan. 2010, art. 9, 1. Even then, however, levels are implicitly employed because premerger markups must be ascertained to determine how much firms' profits suffer as sales fall due to postmerger price increases.

¹⁰ A priori, it seems plausible that the justification for market power inquiries would be relevant to whether market definition is a useful tool. However, since there is substantial agreement that the primary (if not sole) purpose of defining markets is to make inferences about market power, itself understood in a particular way, it turns out that further exploration of the underlying purposes of the market power inquiry is largely immaterial to the present investigation. In part, this simplification is possible because the criticisms developed here are conceptual and general rather than pragmatic and particular.

¹¹ More precisely, the inquiry is ordinarily about the extent to which the profit-maximizing price exceeds the competitive level. (In the examination of horizontal mergers, particularly with the use of critical loss analysis, see *infra* section V.C, it is often asked instead whether a given price increase, usually five percent, following the U.S. Merger Guidelines, would be profitable, or at least break-even, rather than profit-reducing. See U.S. MERGER GUIDELINES, *supra* note 5, § 4.1.2. Typically, if the premerger price was set at a maximizing level and a postulated price increase is break-even, the profit-maximizing price increase would be less than the postulated amount, sometimes approximated as half as much.)

¹² Regarding the focus on price, see Louis Kaplow & Carl Shapiro, *Antitrust*, in 2 HANDBOOK OF LAW AND ECONOMICS 1073, 1080 (A. Mitchell Polinsky & Steven Shavell eds., 2007), which notes that "although anticompetitive harm can come in the form of reduced product quality, retarded innovation, or reduced product variety, our discussion will follow much of the economics literature and most antitrust analysis in focusing on consumer harm that comes in the form of higher prices. This limitation is not as serious as may first appear because higher prices can serve as a loose proxy for other forms of harm to consumers."

the competitive process).¹³ Some variations and qualifications are noted in the margin,¹⁴ and some others are explored in section VI.B.

This concept of market power is usually expressed using the Lerner index (L), also known as the price-cost margin¹⁵:

¹³ See, e.g., U.S. MERGER GUIDELINES, *supra* note 5, § 1 (“A merger enhances market power if it is likely to encourage one or more firms to raise price, reduce output, diminish innovation, or otherwise harm customers as a result of diminished competitive constraints or incentives. . . . For simplicity of exposition, these Guidelines generally discuss the analysis in terms of . . . price effects.”); AREEDA, HOVENKAMP & SOLOW, *supra* note 2, ¶ 501, at 109; MARKET POWER HANDBOOK, *supra* note 1, at 5 (“The economic concept of market power is central to the legal analysis of most antitrust cases.”); Dennis W. Carlton, *Market Definition: Use and Abuse*, 3 COMPETITION POL’Y INT’L 3, 5 (2007); William M. Landes & Richard A. Posner, *Market Power in Antitrust Cases*, 94 HARV. L. REV. 937, 937, 939 (1981); Gregory J. Werden, *Demand Elasticities in Antitrust Analysis*, 66 ANTITRUST L.J. 363, 373–74 (1998) (tracing the U.S. Supreme Court’s acceptance of this economic definition of market power beginning in the 1960s); *id.* at 373 (“The courts of appeals have widely used the economic definition of market power.”); *DG Competition Article 82*, *supra* note 1, ¶ 24 (“Market power is the power to influence market prices, output, innovation, the variety or quality of goods and services, or other parameters of competition on the market for a significant period of time. . . . An undertaking that is capable of substantially increasing prices above the competitive level for a significant period of time holds substantial market power and possesses the requisite ability to act to an appreciable extent independently of competitors, customers and consumers.”); VALENTINE KORAH, AN INTRODUCTORY GUIDE TO EC COMPETITION LAW AND PRACTICE 119 (9th ed. 2007) (“In *Continental Can*, the ECJ implicitly accepted the Commission’s definition of a dominant position based on the economists’ concept of power over price . . .”).

¹⁴ A commonly cited definition is that “[m]onopoly power is the power to control prices or exclude competition.” *United States v. E.I. du Pont de Nemours & Co. (Cellophane)*, 351 U.S. 377, 391 (1956). Two differences may be noted briefly. Most obviously, this formulation adds the alternative of being able to exclude competition. Some have rationalized this phrasing as indicating that a monopolist (the focus of inquiry in *Cellophane*) might use its power either to raise price or instead, perhaps in a predatory manner, to exclude competition. Others have interpreted “or” as “and,” emphasizing that the ability to raise prices over a significant period of time requires somehow keeping competitors at bay. See, e.g., Richard Schmalensee, *Another Look at Market Power*, 95 HARV. L. REV. 1789, 1795 (1982). The other, more subtle difference — echoed more sharply in some other versions as well — is the reference to the ability to “control” price (others have referred to a firm having discretion or, in the words of Case 322/81, *Michelin v. Comm’n*, 1983 E.C.R. 3461 § 30, at 3503, in the European Union, the ability “to behave to an appreciable extent independently of its competitors and customers”). Most of these alternatives are vague and potentially misleading. A firm with little market power can still control its price, although significant elevations would be unprofitable, and a firm with great market power cannot simply ignore competitors and customers and, moreover, if it wishes to maximize profits, will feel compelled to select a single, particular price, just as would its low-power counterpart. See, e.g., PHILLIP AREEDA, LOUIS KAPLOW & AARON EDLIN, *ANTITRUST ANALYSIS* 485–86 (6th ed. 2004); Kaplow & Shapiro, *supra* note 12, at 1098; see also *DG Competition Article 82*, *supra* note 1, ¶ 23 (in elaborating that dominance requires the ability to behave independently of competitors and customers, states: “The notion of independence, which is the special feature of dominance, is related to the level of competitive constraint facing the undertaking(s) in question. For dominance to exist the undertaking(s) concerned must not be subject to effective competitive constraints. In other words, it thus must have substantial market power.” (footnote omitted)). These and other differences in definitions of market power are ignored here.

¹⁵ See, e.g., AREEDA, HOVENKAMP & SOLOW, *supra* note 2, ¶ 503b, at 118–19; Kaplow & Shapiro, *supra* note 12, at 1080; Landes & Posner, *supra* note 13, at 939, 941.

$$(1) \quad L = \frac{P - MC}{P}.$$

In this formula, P denotes price and MC is marginal cost.¹⁶ The Lerner index indicates the portion of the price that is in excess of what is required to cover marginal cost.¹⁷ If price equals marginal cost, as with perfect competition, the numerator equals zero and hence L is zero: that is, no market power. Suppose that price is 10% above marginal cost (e.g., $P = 110$, $MC = 100$); then $L = 10/110 \approx 0.09$, or 9%. That is, approximately 9% of the price of 110 constitutes the markup — the profit on the marginal unit — above the marginal cost of 100 of producing that unit. Price 50% above marginal cost (e.g., $P = 150$, $MC = 100$) yields an L of approximately 33%, and a price 100% above (double) marginal cost (e.g., $P = 200$, $MC = 100$) yields an L of 50%. Thus, a higher value of the Lerner index indicates greater market power.¹⁸

Given this understanding of market power, the role of the market definition / market share paradigm is, on its face, obscure. Market shares, whether in a properly defined relevant market or in any other, do not appear in the definition of market power. Instead, one only sees price and marginal cost. It would seem that, if one wished to know the level of market power, one would, therefore, examine price and marginal cost. In order to get to market definition (much less for it to be the standard technique), it seems that three propositions must be true: (1) direct determination of market power is frequently difficult; (2) market definition is somehow helpful in determining market power indirectly; and (3) the market definition / market share paradigm is often the best alternative among all the other possible techniques or at a minimum is usually a helpful complement. Section B takes up the second statement, and the remainder of the Article is concerned with the third. Here, we briefly examine the first.

¹⁶ Specifically, MC is the level of marginal cost at the current quantity of output. If marginal cost is rising with quantity and price is elevated, it follows that, if price were reduced from a supracompetitive level to the competitive level, quantity would be higher and thus marginal cost would be higher, so L would overstate the markup relative to the competitive price.

¹⁷ One could instead examine, for example, the fraction P/MC (which equals $1/(1-L)$), in which case perfect competition would be associated with an index value of 1, price 10% in excess of marginal cost would result in a value of 1.1, and so forth. Or one could consider $(P-MC)/MC$ (which equals $L/(1-L)$), which in turn is equivalent to $(P/MC)-1$, so perfect competition would (as with the Lerner index) have an index value of 0, price 10% in excess of marginal cost would have a value of 0.1, and so forth. Although these alternatives may seem a bit more straightforward, all are substantially equivalent (since they are all monotonic transformations of L), the use of the Lerner index is conventional in economic analysis of the subject, and as will be seen (notably, in expressions (2) and (4)), L has simpler properties that render its use convenient.

¹⁸ In the limit, $L = 100\%$ implies an infinite price or a marginal cost of zero, so for finite prices and strictly positive marginal cost, we have $L < 100\%$.

The assumption that direct determination of market power is frequently difficult is well accepted,¹⁹ although some may view it to be overstated.²⁰ Simple application of the Lerner index formula requires measuring price and marginal cost. Price is usually straightforward, although difficulties may arise, for example, in the presence of bundled pricing (whether multiple products are sold together or various services, including delivery and credit, are supplied jointly) or when different customers are charged different prices. Marginal cost is far more challenging, both because of difficulties in identifying which costs are variable (which depends on the relevant time period) and due to the presence of common costs that may be hard to allocate among the multiple products a firm produces. The problem of measuring marginal cost is usually regarded to be serious or at least problematic with sufficient frequency that other means of inferring market power are potentially attractive.

B. Analysis

This section considers the manner in which the market definition / market share paradigm allows inferences about market power. It begins by examining a simple setting in which a dominant firm sells a homogeneous good, subject to the possibility that small rivals are able to meet some of consumers' demand for the product and that consumers will to an extent respond to higher prices by switching to other products. In this setting, the relevance of the dominant firm's market share will be explained. The remainder of the section will take up the question of market definition, focusing in particular on redefinition that entails broadening the market beyond the homogeneous goods market originally examined.²¹ For concreteness and conve-

¹⁹ See, e.g., ANTITRUST LAW DEVELOPMENTS, *supra* note 1, at 229.

²⁰ Economists' favored methods of predicting unilateral price effects of horizontal mergers of differentiated products suppose that premerger margins can be determined. Joseph Farrell and Carl Shapiro (economists who have spent time in antitrust enforcement agencies) argue that, although "gross margins are hard to measure using public data . . . 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." Farrell & Shapiro, *supra* note 9, at 18. Others are more skeptical. See, e.g., MASSIMO MOTTA, COMPETITION POLICY 116 (2004); Baker, *supra* note 2, at 142–43; Werden, *supra* note 13, at 394.

²¹ Since the initial setting involves perfectly homogeneous goods, the question whether a narrower market might be contemplated does not arise, at least as the market definition problem is ordinarily understood. In a world in which all products are differentiated, one might suppose that the market is originally taken to include only a single product, the one produced by the firm under scrutiny. For mergers, of course, at least two firms' products would need to be considered. For further discussion of differentiated products mergers, see section V.C.

nience, the discussion will focus on issues involved with product market definition.²²

1. *Dominant Firm, Homogeneous Good, Small Rivals.* — Consider how a single firm acting unilaterally to maximize profits chooses its price. If it is a perfect competitor, it has no interesting choice to make: if it charges even slightly above the prevailing market price, it will make no sales whatsoever. (Keep in mind that perfect competition supposes that competitors are offering an identical product; this assumption of homogeneous goods will be retained in this subsection.) Furthermore, the firm has no reason to charge less than the prevailing price, for it is assumed that it can sell all of its output at the market price. Accordingly, the firm will indeed sell at the market price. Moreover, it will select a level of output at which its marginal cost equals price, as is familiar.

Now suppose instead that competition is imperfect — specifically, that raising the price somewhat above the market price will not result in the firm losing all of its sales. The firm will face a simple tradeoff as it increases price: On one hand, each unit of price increase will earn it an additional unit of profit on all of the output that it continues to sell. On the other hand, each unit of price increase will result in some loss of sales, which will reduce profit by $P-MC$ per sale that is lost. (The reason is that, for an incremental lost sale, the firm forgoes P in revenue and no longer incurs MC in cost.) The profit-maximizing price is that at which these forces are in balance.

²² It is common to view the problems of product and geographic market definition as similar, and in certain respects they are. The major difference is that, in some settings involving geographic market definition issues, consumers' preferences may have a particular structure that may be highly relevant to the analysis. For example, if the products are otherwise identical, shipping costs for different customers are the same, and preferences over shipping time do not differ across consumers, everyone will value the substitute identically; in particular, all will prefer it to the local product if and only if its price, adjusted for transportation costs and delay, is lower. For some purposes, it may be useful to treat geographically remote supply having such characteristics similarly to present rivals and to potential entrants, although aspects of the analysis differ (for example, remote suppliers' response to a local price change will depend on the demand elasticity in the remote market as well as on remote suppliers' cost functions; for further related discussion, see notes 38 and 85 on partial versus general equilibrium analysis). Details will not be pursued here. When consumers' preferences for the product from the remote market vary more broadly — for example, the product may itself have different characteristics or a different brand image, or preferences over delivery time may vary — then geographically remote suppliers are more like local suppliers of differentiated (or simply different) products, so the standard analysis of product markets is more closely applicable. This differentiated products perspective on geographic market definition is emphasized in Louis Kaplow, *The Accuracy of Traditional Market Power Analysis and a Direct Adjustment Alternative*, 95 HARV. L. REV. 1817, 1835–43 (1982). Subsequently, product differentiation has become a central consideration in the theory of international trade, particularly to explain two-way trade flows between countries at similar stages of economic development. See, e.g., Paul Krugman, *The Increasing Returns Revolution in Trade and Geography*, 99 AM. ECON. REV. 561, 562–65 (2009) (revised version of 2008 Nobel lecture).

It is straightforward to derive the following characterization of the firm's pricing decision²³:

$$(2) \quad L = \frac{P - MC}{P} = \frac{1}{|\varepsilon_f|}.$$

The term ε_f denotes the *firm's* elasticity of demand. (This concept is to be contrasted with the market elasticity of demand, to be mentioned shortly.) It is defined as the percentage change in the firm's sales as the firm increases its price. Specifically, $\varepsilon_f = (dQ/dP)(P/Q)$, where Q denotes the quantity of the firm's output. The absolute value is taken because the elasticity, as defined here, is a negative number (a price increase causes sales to fall). The rough intuition behind expression (2) is straightforward: the greater the rate at which sales fall as price rises (and thus the larger the denominator on the right side, making the overall value of the right side smaller), the less the firm will find it profitable to elevate its price P above MC . The specifics in the derivation essentially track the logic of the preceding paragraph.

It is worth elucidating ε_f , the firm's elasticity of demand. This elasticity indicates the rate at which the firm loses sales *to all sources* as it increases its price. The firm might lose some sales to rivals, who are induced to produce more as a consequence of the price increase. It might lose sales to other firms that produce close substitutes. And it may lose sales to distant substitutes — indeed, products most would not think of as substitutes. (If the price of pizza rises, a consumer may purchase more of other Italian foods, other foods generally, or other unrelated goods, whether video games, clothing, or a nicer automobile; for elaboration, see Part V.) Note that the firm, in making its pricing decision, does not care per se where consumers who cease to buy its product go instead. All it cares about is the rate at which they reduce their purchases of its own product as the price rises. If this rate is high, the firm will not wish to elevate price very much above marginal cost. If the rate is low, it will profit by raising prices substantially.

Suppose next that our dominant firm faces a number of small rivals, sometimes referred to as a competitive fringe. Each of these firms is taken to produce the same product as the dominant firm, which is to say that we continue to assume homogeneous products.²⁴ Furthermore, in saying that each of these firms is “small,” it is meant

²³ See, e.g., Kaplow & Shapiro, *supra* note 12, at 1080. Beginning with the expression for profits (price times output, minus cost, itself as a function of output; that is, $PQ - C(Q)$, where C is the cost function), one takes the derivative with respect to price (keeping in mind that output is implicitly a function of price, as reflected by the firm's demand curve), rearranges terms, and makes a substitution using the definition of the firm's elasticity of demand.

²⁴ See, e.g., *id.* at 1081; Landes & Posner, *supra* note 13, at 944.

that they are assumed to behave as price-taking competitors.²⁵ Accordingly, if the dominant firm sets the price P , each of the rivals is assumed to expand output to the point at which its marginal cost just equals P . The dominant firm's demand in this model is often described as a residual demand; it sells only to those who continue to buy the product (rather than some other product) and are unable to obtain the product from firms in the competitive fringe.²⁶

This elaboration does not change the prior analysis so much as make more concrete what lies beneath the firm's elasticity of demand. Recall that, in addition to losing sales to rivals, price increases reduce total industry sales of the product. The firm's elasticity of demand reflects both of these factors, and some further derivations establish the following relationship²⁷:

$$(3) \quad |\epsilon_f| = \frac{|\epsilon_d| + (1-S)\epsilon_r}{S}.$$

This formula for the (absolute value of the) firm's elasticity of demand introduces three new terms: $\epsilon_d = (dX/dP)(P/X)$, the *market* elasticity of demand (X denotes total market demand, whereas Q , introduced previously, denotes what the dominant firm sells); $\epsilon_r = (dY/dP)(P/Y)$, the rivals' elasticity of *supply* (Y denotes rivals' total supply); and S , the dominant firm's market share (and hence $1-S$ indicates rivals' aggregate share).²⁸

²⁵ If they are not all small and hence do not all behave as price takers, further analysis is required. Some of the pertinent considerations are mentioned in section VI.C's discussion of unilateral and coordinated effects in horizontal mergers. See also Janusz A. Ordover, Alan O. Sykes & Robert D. Willig, *Herfindahl Concentration, Rivalry, and Mergers*, 95 HARV. L. REV. 1857 (1982) (extending the analysis to the case in which multiple firms in the homogeneous goods market are not small).

²⁶ This approach derives from Karl Forchheimer, *Theoretisches zum Unvollständigen Monopole*, in 32 JAHRBUCH FÜR GESETZGEBUNG, VERWALTUNG UND VOLKSWIRTSCHAFT IM DEUTSCHEN REICH 1 (Gustav Schmoller ed., 1908). See Gavin C. Reid, *Forchheimer on Partial Monopoly*, 11 HIST. POL. ECON. 303 (1979).

²⁷ See, e.g., Kaplow & Shapiro, *supra* note 12, at 1081-82. This derivation begins by expressing the dominant firm's demand as the total industry demand for the product minus that portion supplied by the competitive fringe. When one takes the derivative of that expression with respect to price, rearranges terms, and uses the definitions of the pertinent elasticities and of market share, the expression in the text results. A version of this derivation first appeared in George J. Stigler, *Notes on the Theory of Duopoly*, 48 J. POL. ECON. 521, 523-24 (1940). Stigler was also the first to use this formula to emphasize the point that market share alone tells only part of the story — one that, in isolation, can be misleading. See *id.* at 524.

²⁸ The term ϵ_d appears between absolute value bars because, like the firm elasticity, the market elasticity of demand as defined in the text is negative: as price rises, the quantity demanded in the market as a whole falls. By contrast, the term ϵ_r is positive to begin with: as price rises, the quantity supplied to the market by rivals increases. On another note, it should be understood that all of these terms are endogenous, in that they depend on the price. As price rises, both elasticities may change (it is sometimes supposed that the elasticity of demand tends to be higher and of supply perhaps lower as prices rise, although neither need be true), and the dominant firm's share

This formula, on reflection, is quite intuitive. First, as the firm raises its price and thus the market price for the product, consumers buy less of that product in total. The market elasticity of demand (sometimes referred to simply as the elasticity of demand) indicates the rate at which this switching occurs. That is, the market elasticity of demand, ε_d , indicates the percentage fall in total sales of the product as the market price is increased. Clearly, a larger (in magnitude) market elasticity of demand implies a larger firm elasticity of demand.

Why is this market elasticity divided by S on the right side of the expression? The reason is that changes in the market demand must be proportioned by the quantity supplied by the firm to translate into the firm elasticity. To see this, suppose that the dominant firm's share is 50%, so $S = 0.5$. Now, if the market elasticity of demand is 2.0 (and, for simplicity, suppose for the moment that the rivals' supply elasticity is zero, so we can ignore the other term), then the firm's elasticity is 4.0. If price increases by one percentage point, market demand falls by two percentage points. But this fall is hypothesized to come entirely from the dominant firm (we are taking rivals' supply to be constant). Since the firm's output is only half the industry total (recall that its market share is 50%), the two percentage-point fall in industry demand translates into a four percentage-point fall in its own demand.²⁹

Second, consider the rivals' elasticity of supply, ε_s . This elasticity is the percentage change in rivals' supply as the price increases. (One can think of this term as referring primarily to expansion by existing rivals, but in a broader interpretation it could be understood as including supply substitution and entry (see section VI.A).) A high elasticity indicates that a modest price increase will induce rivals to expand their output significantly. The underlying reasoning is that the rivals are assumed to act as price-takers. When price rises, they will increase their output to (re)equilibrate their marginal cost to the price. Now, if their marginal costs rise only gradually with output, they will increase sales substantially, yielding a high supply elasticity. Similarly, if their marginal costs rise steeply (or, in the extreme case, if they are capacity constrained), they would find it profitable to raise their supply little (if at all).

A higher rivals' supply elasticity implies a higher firm elasticity of demand, as the formula indicates. Note that the rivals' supply elasti-

will in general change (as price increases, rivals supply more and the dominant firm less, so its share will be lower at higher prices). Accordingly, each of these terms refers to levels at the prevailing price.

²⁹ If the firm's initial quantity is 50 and that of the industry is 100, then the industry quantity falls by 2, from 100 to 98, and the firm's falls by 2, from 50 to 48. A 2-point fall on a base of 50 is 4%, and the elasticities, recall, are defined in percentage terms.

ty is weighted in the numerator by $1-S$, the rivals' aggregate market share. This reflects the fact that the impact of a given *percentage* increase in sales by the rivals on our dominant firm is proportional to the rivals' share.³⁰ Simply put, a given percentage increase, say 2%, will have double the impact in taking sales from the dominant firm if the base of that 2% increase is twice as high. Finally, as with the market elasticity of demand, this overall impact is divided by S (in the denominator of the above expression) for precisely the same reason.

It is useful to employ this decomposition of the firm's elasticity of demand to restate the expression for the dominant firm's market power, as follows:

$$(4) \quad L = \frac{P - MC}{P} = \frac{1}{|\epsilon_f|} = \frac{S}{|\epsilon_d| + (1 - S)\epsilon_r}.$$

Expression (4) combines equations (2) and (3) — in the case of the latter, inverting both sides. This expression informs us, as we intuitively appreciated in any case, that the dominant firm's market power is greater the lower is the magnitude of the market elasticity of demand, the lower is the magnitude of rival firms' elasticity of supply, and the higher is the firm's market share.

Because market share is our focus, it is useful to expand on the intuition why a greater market share implies more market power. Essentially, the reasoning elaborating expression (3) is now being recast in somewhat different terms. Market share augments market power, all else equal, for two sets of reasons. One is that the higher the dominant firm's share, the lower is its rivals' aggregate share, $1-S$, and thus the smaller is the impact of a given elasticity of rivals' supply, as just explained. Second, the higher the dominant firm's market share, the lower is the percentage by which it must reduce its own output on account of a given fall in demand and rise in rivals' supply due to a price increase; also, the larger is the base (retained sales) on which it profits from charging the higher price.

In all, as has long been known, and in recent decades has been more rigorously understood in competition policy circles, we can identify the basic economic logic indicating why market shares are relevant to inferences of market power. There remains, however, an important question: in what market are we measuring the dominant firm's market share? The answer is that, because the present analysis is explicitly predicated on the assumption of homogeneous goods, the market in

³⁰ If rivals sell 50 of 100 units, a 2% increase on the base of 50 is 1 unit, which is only 1% of the 100 units in the market.

which the share is measured for application of the formulas (3) and (4) is the narrow, homogeneous goods market.

A key point, which will be emphasized in much of the subsequent discussion as well, is that this formula yields the correct measure of market power without any need to engage in further analysis of market definition, regardless of how many substitutes may exist and how close some of the substitutes may be. If, for example, there are many good substitutes, then the market elasticity of demand, ε_d , will tend to be high, and expression (4) accordingly indicates that market power will tend to be low. If there are few good substitutes, then the market elasticity will tend to be low, and expression (4) indicates that market power will tend to be high. Part V will elaborate and qualify aspects of these statements, but remaining undisturbed will be the crucial lesson that formula (4) is correct, and, if the market elasticity of demand is properly understood, the market power inference on that account will be valid.

2. *Market Redefinition.* — At this point, there appears to be no need ever to engage in analysis of market redefinition. This conclusion not only is correct based on what has been considered thus far, but also will remain so as the analysis proceeds. Suppose, nevertheless, that one insists on allowing for market redefinition, as is commonly done. In particular, suppose that one contemplates redefining the market by expanding it to include a good substitute. After such redefinition, the market would no longer be a homogeneous goods market, so strictly speaking the foregoing analysis is inapplicable. This subsection considers how it might be extended to cover this case — the possibility of which analysts, enforcement agencies, and courts largely take for granted. As will be seen, the homogeneous goods assumption proves to be extremely important, and in ways that are not well appreciated.

Before proceeding with the analysis, however, it is useful to reflect on why we might be engaging in market redefinition, particularly in light of the point that formula (4) is correct — regardless of whether there are, say, a number of very close substitutes that are outside the market. For redefinition to make any sense, it must be that there is some aspect of the above formula that is difficult to measure (approximate) very well and, furthermore, that redefining the market to include substitutes will somehow make it easier to apply the formula correctly in the broadened market. As will be seen, the analysis does not bear out this hope.

To preview the analysis itself, observe that we will broaden the market by adding in the market for a substitute product. We put to the side the question of how good a substitute should be in order for it

to be included;³¹ rather, taking as given that the appropriate criterion for inclusion is met, we will consider how one should proceed to analyze the firm's market share in this broadened market.³² We need some method, some formula, that translates market share and other features into our measure of market power. We have a formula, expression (4), that accomplishes this feat; however, as emphasized in subsection 1, our formula is valid only for a market with homogeneous products, which we no longer have. Accordingly, we seek a new formula — some adjustment to expression (4) — that will enable valid market power inferences from the firm's market share in the expanded market.³³ And, consistent with our motivation for market redefinition, it will need to be the case that application of the new formula is somehow easier or more accurate than use of expression (4) for the homogeneous goods market.

If we combine some substitute-product market (in which, we will assume, our dominant firm does not produce), the firm's market share will fall from S to a level that will be denoted by \hat{S} . This new, lower value would be computed by dividing the firm's sales not by total sales in the original homogeneous goods market, as before, but by the sum

³¹ The reader may imagine that we have chosen the closest substitute (setting aside just how that is to be defined) and that the chosen substitute is quite close. None of the analysis to follow depends on how this choice is made or whether it is correct by conventional criteria.

³² If it turned out that broadening markets did make sense, the manner in which one analyzed market shares in the broadened market would presumably bear on the criterion for when the market should be broadened. The analysis in Parts IV and V will cast doubt on whether there could be a coherent principle in any event.

³³ Some analysts have discussed how, when markets are combined, the firm's lower share will be offset, for example, by there being in some sense a lower demand elasticity in the broadened market, with the suggestion that market definition would not matter as long as the proper adjustments were made. *See, e.g.*, Kaplow & Shapiro, *supra* note 12, at 1091; Landes & Posner, *supra* note 13, at 962. But the actual manner in which the values of other terms in the formula need to be adjusted in order to make correct inferences has not been examined. Richard Schmalensee offers the closest exception in an appendix; he presents a formula for the market demand elasticity in the combined market that is a function of the elasticities in the two separate markets and the (weighted average of the two) cross-elasticities of demand between the two products. *See* Schmalensee, *supra* note 14, at 1815–16. This formula, however, is addressed to the question of how much market power a hypothetical monopolist of both markets would have, in particular, if it increased the price in both markets by the same proportion (an inquiry like that often used to apply the U.S. MERGER GUIDELINES, *supra* note 5, § 4.1.1; see note 99 for further discussion). This question differs importantly from that of how to determine the market power of the dominant firm in the initial market once the additional market has been combined. Moreover, the new formula requires knowledge not only of the market demand elasticity in the initial market — which, combined with the firm's share in that market (and setting aside rivals' supply response, as Schmalensee does), would be sufficient to determine market power — but also of additional pieces of information: the market demand elasticity in the other market and the (weighted average) cross-elasticities between the two. Therefore, it cannot be that combining markets makes it easier to determine market power.

of this total and the total sales in the market we just combined.³⁴ If we wish to infer the firm's market power in this broader market, we might expect to employ the following formula:

$$(5) \quad L = \frac{P - MC}{P} = \frac{1}{|\epsilon_f|} = \frac{\hat{S}}{|\hat{\epsilon}_d| + (1 - \hat{S})\hat{\epsilon}_r}.$$

Expression (5) is the same as expression (4) except that, on the right side, each of our key terms takes a modified value (denoted with a “^”). Presumably, those who have supposed that one can begin with formula (4) and adjust each of the terms to make inferences in the broader market have something like this expression in mind.

If we limited our attention to the share falling from S to \hat{S} , we would be making a big mistake. Rather, we have to develop an interpretation of this lower share \hat{S} in the new market, which in turn requires figuring out what the two new elasticity terms represent. We know a priori that they must be lower. After all, expression (4) is valid, and the left side of expression (5) is the same as that in expression (4), so it must be that the right sides of the two expressions have precisely the same value. Moreover, the lower share, \hat{S} , in formula (5) reduces its value. (The intuition was presented in subsection 1. From the expression itself, the lower \hat{S} in the numerator reduces the value of the right side, and the lower \hat{S} in the denominator, since it appears as $1 - \hat{S}$, results in a higher value of the denominator as a whole, which further reduces the value of the right side.) Hence, one or both of the elasticity terms in the denominator necessarily have a lower magnitude.

But what are the lower values for these modified elasticities? One could choose figures that yield the right answer. (For example, we could keep the rivals' supply elasticity constant and figure out how much we need to raise the demand elasticity to restore the original, correct value to the right side of expression (5).) But this procedure is entirely unsatisfactory. First, it would involve a different adjustment in different cases — in which event there is no formula, that is, no all-purpose tool that can be employed broadly. Second, it patently assumes that we already know the answer, in which case there is no point in redefining the market in the first place. Recall that the motivation for market redefinition is that it is somehow easier or more accurate to use information in the combined market to make inferences.

Accordingly, we need to figure out how each of the elasticity terms needs to be modified to produce a valid rendering of formula (5). It

³⁴ Even this simple adjustment can involve complications, notably, whether to measure shares by units sold, revenue, or some other metric.

will suffice to determine what alteration of each would be necessary if the value of the other elasticity were zero; it will be clear that combining the two results will work for any values of the initial two elasticities. Begin with the market elasticity of demand.

$$(6) \quad \hat{\varepsilon}_d = \frac{\hat{S}}{S} \varepsilon_d.$$

As we knew would have to be the case, the value of this modified market elasticity of demand is lower than that of the initial elasticity. Indeed, it is lower by a factor that just offsets the proportionate fall of the numerator in moving from expression (4) to expression (5). For the rivals' supply elasticity, we have:

$$(7) \quad \hat{\varepsilon}_r = \frac{(1-S)\hat{S}}{(1-\hat{S})S} \varepsilon_r.$$

This modification is a bit more convoluted. The \hat{S}/S factor, just like that in expression (6), offsets the fall in the numerator. The $(1-S)/(1-\hat{S})$ factor offsets the increase in the weighting term in the denominator. It should be apparent that, if each of the two elasticities is defined in this manner, the value of expression (5) will equal that of expression (4) in all cases, so the market power inference from formula (5) will be correct.³⁵

Again, we have a stark problem. Why would one ever want to combine markets? In order to interpret the lower share in the new, broader market, we need to make two further adjustments, each of which involves undoing the effect of broadening the market in the first place. Taken together, all of the adjustments cancel; the effect is a complete wash.

Observe that the information we need to apply formula (5) — to determine market power in the expanded market — includes every bit of information that we needed to apply formula (4) — for market power in the original, narrow, homogeneous goods market. Hence, there is no respect in which it could possibly be easier to use formula (5) than formula (4).³⁶

Furthermore, aside from the market share \hat{S} itself,³⁷ no actual information pertaining to the new market is used. We do not use any information about the market elasticity of demand from the other market or about the supply elasticity in that market. Perhaps the for-

³⁵ It should likewise be clear that no other general formula for adjustment would work.

³⁶ The only difference is that, to apply formula (5), we need also to compute \hat{S} . As it turns out, however, this difference also is immaterial. After all, if any error is made, it will wash out once the adjustments in expressions (6) and (7) are completed, assuming that we use the same (possibly erroneous) value of \hat{S} in doing so.

³⁷ *But see supra* note 36.

mer omission is most surprising.³⁸ Might there be some other way to express the market elasticity of demand in the combined market, perhaps using information about the market elasticity in the second market or the cross-elasticity of demand between the two markets? The answer is no, or at least not in any meaningful sense, a conclusion that is further documented in the margin.³⁹

³⁸ Regarding the latter omission, one might suppose that the magnitude of the overall supply elasticity in the market being combined would matter. Moreover, this dependence may be thought to exist despite the fact that additional supply of the second product does not in itself reduce demand for the dominant firm's product as is the case when rivals producing goods identical to those of the dominant firm expand their supply. Suppose, for example, that the supply elasticity in the second market is very low. Then, a rise in the price for the product in the first market would not induce as much switching to the product in the second market because, as more purchasers switched, the price of the good in the second market would be bid up as a consequence, reducing the relative benefit of substitution. Implicitly, any such effect could be seen as reflected in the market demand curve in the first market and thus in the market demand elasticity in that market. (It is conventional, however, to define this curve and state this elasticity under the assumption that other prices are fixed; this approach is termed partial equilibrium analysis. See, e.g., DAVID M. KREPS, *A COURSE IN MICROECONOMIC THEORY* 279–80 (1990). Similar assumptions underlie formula (8) in section V.A for the market elasticity of demand as a weighted sum of other products' cross-elasticities. See *infra* note 85.)

³⁹ As examined in Part V — see especially the discussion of expression (8) in section A — the market elasticity of demand in a homogeneous goods market can be decomposed into a weighted sum of cross-elasticities with all other products. To be sure, the cross-elasticity with the market we are combining would be one of those terms. But that leaves the rest, which taken together correspond neither to that original cross-elasticity nor to the market demand elasticity in the second market. Stated formally, we have:

$$|\hat{\epsilon}_d| = \frac{\hat{S}}{S} |\epsilon_d| = \frac{\hat{S}}{S} |\tilde{\epsilon}_d| + \frac{R_f - \hat{S}R_d}{R_f} \epsilon_{21}.$$

In this expression, R_f is the revenue of the dominant firm; R_d is the total revenue in the original, narrow, homogeneous goods market; and ϵ_{21} is the cross-elasticity, specifically, the increase in the quantity purchased of the good in the second market as the price in the first market is increased. Finally,

$$|\tilde{\epsilon}_d| = 1 + \sum_{i=3}^N \frac{R_i}{R_d} \epsilon_{i1}.$$

Here, R_i is the total revenue in each of the other markets (besides the two on which we are focusing, which have been denoted markets 1 and 2), and ϵ_{i1} is the cross-elasticity for each of the other goods. It is apparent that these expressions neither illuminate the basic ideas nor provide a simpler means of assessing market power than that provided in formula (4), a subject further discussed in Part V. One could perform yet further convolutions: expressing the market elasticity of de-

In all, it is difficult to understand how the market redefinition process could facilitate making sensible inferences about market power. This observation is not meant to suggest that direct application of formula (4) for the narrow, homogeneous goods market is usually easy. Rather, the point is that moving to expression (5) does not help. To get the right answer, one essentially has to undo the market redefinition. This almost entirely unrecognized point casts the market definition / market share paradigm in a new light.⁴⁰

On reflection, perhaps this negative conclusion should not be surprising. In the field of industrial organization economics, which devotes substantial attention to matters of market power and competition policy more generally, the concept of market redefinition does not really exist.⁴¹ The fact that those economists who focus the most on

mand in the combined market in terms of the market elasticity in the original market, as in expression (6); further decomposing that elasticity into a weighted sum of cross-elasticities, as in expression (8) in section V.A; and pulling out the particular cross-elasticity for the second market, but then adding it back in, in order to return to the original elasticity. But, again, this task would be pointless. (If one was interested not in our dominant firm's market power, but in the ability of a hypothetical monopolist in the combined market to proportionately raise prices in the two markets, then there does exist a way to combine information about the two market elasticities and the cross-elasticities between the two products to illuminate this question. See *supra* note 33; see also *infra* note 99 (discussing the Merger Guidelines' hypothetical monopolist test).)

⁴⁰ Regarding the lack of prior recognition, most sophisticated observers are aware of the need to interpret market shares carefully in redefined markets — and in original markets — but have not appreciated that if one wishes to undertake such interpretations in a logical fashion, one is essentially forced to return to the original, narrow, homogeneous goods market. Perhaps the sharpest criticism that is in similar spirit to that presented in the text is offered in Schmalensee, *supra* note 14, at 1797. He does not mention that the method cannot be used even in the case of homogeneous products once one engages in market redefinition, and his more detailed elaboration, see *id.* at 1797–98, does not feature the core problem with market redefinition identified in the text, although his discussion does indicate some of what can go wrong when one does not fully undo the market redefinition. See also Kaplow, *supra* note 22, at 1830 n.54 (discussing a numerical example based on the *Cellophane* case in which making inadequate adjustment in the broader market may substantially understate market power relative to the correct result derived using the formula in the narrower, homogeneous goods market).

⁴¹ See, e.g., Franklin M. Fisher, *Economic Analysis and "Bright-Line" Tests*, 4 J. COMPETITION L. & ECON. 129, 132 (2008) [hereinafter Fisher, *Economic Analysis*] ("What, then, does economic analysis have to say about market definition? In one sense, the answer is 'Nothing at all.' The question of what is 'the' relevant market never arises in economics outside of antitrust. Moreover, . . . it is not a question that has a precise, well-defined answer."); Franklin M. Fisher, *Horizontal Mergers: Triage and Treatment*, J. ECON. PERSP., Fall 1987, at 23, 27 ("Market definition is an artificial construction created by antitrust litigation. For any other purpose of economic analysis, the binary question of whether particular firms or products are 'in' or 'out' of a given market is a meaningless one. Even in antitrust cases, that question is not a useful one if substantive results turn on the answer."); Kaplow & Shapiro, *supra* note 12, at 1091 ("Although we have discussed the economic analysis of market power at some length, the concept of market definition has not yet appeared directly."); *id.* at 1170 ("[T]he common approach of defining markets and looking to market shares . . . does not play a central role in economists' analysis of market power."); Adriaan ten Kate & Gunnar Niels, *The Relevant Market: A Concept Still in Search of a Definition*, 5 J. COMPETITION L. & ECON. 297, 298 (2009) ("[I]t is hard to find a satisfactory description or definition [of what a relevant market is] in textbooks on microeconomics or industrial

the pertinent issues have never found it helpful (or possible) to develop a way of making market power inferences in combined, nonhomogeneous goods markets using as one key input a firm's market share in the combined market is a powerful signal that there is no meaningful or useful way to do so. One might add that many economists in this field, including some leading ones, have held positions in enforcement agencies or consulted extensively on particular cases, so their exposure to the issue and motivation to find a solution have both been significant.⁴² Instead, some have devoted substantial and successful efforts to developing a number of more direct means of inferring market power that do not rely on market redefinition.⁴³

III. STANDARD REFERENCE MARKET

Some notion of a standard reference market is a necessary, central element in most discussions of market definition and statements about the implications of various market shares. Nevertheless, this concept has remained almost entirely inchoate in cases and commentary alike. Section A explains why some such benchmark is indeed implicit in the market definition / market share paradigm. Section B expands on its content and in the process reveals a serious gap both in current discourse and in applications of the paradigm.

organization, if the concept is mentioned at all in such readings.”); *see also* Gregory J. Werden, *Market Delineation and the Justice Department's Merger Guidelines*, 1983 DUKE L.J. 514, 516 (noting that “[e]conomic theory does not require the delineation of markets, and most economists would not know how to begin to delineate one,” but going on to state that the U.S. Merger “Guidelines employ the concept of an ‘antitrust market,’ a market delineated for the sole purpose of antitrust analysis”).

⁴² This observation does, however, raise the question of why they have not previously debunked the practice rather than merely offering reservations. Interestingly, before the market definition / market share paradigm became prominent, economists had expressed hostility toward the approach. *See* Werden, *supra* note 7, at 126; *see also* Edward H. Chamberlin, *Product Heterogeneity and Public Policy*, AM. ECON. REV., May 1950, at 85, 86–87 (“‘Industry’ or ‘commodity’ boundaries are a snare and a delusion — in the highest degree arbitrarily drawn, and, wherever drawn, establishing at once wholly false implications both as to competition of substitutes within their limits, which supposedly stops at their borders, and as to the possibility of ruling on the presence or absence of oligopolistic forces by the simple device of counting the number of producers included.”).

⁴³ *See, e.g.*, MICHAEL D. WHINSTON, LECTURES ON ANTITRUST ECONOMICS 100–14 (2006) (surveying literature on the estimation of market power in mergers); Jonathan B. Baker & Timothy F. Bresnahan, *Empirical Methods of Identifying and Measuring Market Power*, 61 ANTITRUST L.J. 3 (1992) (surveying techniques); Jonathan B. Baker & Timothy F. Bresnahan, *Estimating the Residual Demand Curve Facing a Single Firm*, 6 INT'L J. INDUS. ORG. 283 (1988) [hereinafter Baker & Bresnahan, *Residual Demand*]; Cory S. Capps, David Dranove, Shane Greenstein & Mark Satterthwaite, *Antitrust Policy and Hospital Mergers: Recommendations for a New Approach*, 47 ANTITRUST BULL. 677 (2002); Kaplow & Shapiro, *supra* note 12, at 1178–80 (surveying literature on methods for mergers); sources cited *infra* note 107 (on critical loss and related methods of assessing competitive effects of differentiated products mergers).

A. Motivation

For present purposes, set aside the criticism in Part II that casts into serious doubt why one would ever wish to redefine markets and whether it is possible to make sensible inferences from market share to market power in redefined markets (short of reversing the redefinition). Return to the simple notion that market shares tell us something about market power; in particular, higher shares ordinarily imply greater power and lower shares less power. But greater or less than what? And by how much?

Consider the following rosy scenario. Imagine that all properly defined (a.k.a. relevant) markets were identical in the following respect: a given share *S* in any such market corresponds to the same degree of market power as that same share *S* indicates in any other such market. If this were so, then whenever anyone referred to a market share in a proper market, they could just as easily be making an explicit statement about the magnitude of extant market power. Anyone privy to the translation table would understand exactly how much market power was conveyed by any specified market share.

Such a depiction, of course, is not merely an oversimplification but a fantasy. There is substantial variation in the market power conveyed by a given market share even if we confine attention to markets that are defined correctly according to some criterion. Formula (4) in Part II makes clear that a dominant firm's market power depends not only on its market share but also on the market elasticity of demand and rivals' elasticity of supply. If, say, the elasticity of demand is very high, it may be that the market is defined too narrowly, and if it is very low, too broadly. But even after such cases are eliminated (by redefining the market to produce a correct one), substantial variation is possible.

This simple point, that market shares viewed in a vacuum can be highly misleading and therefore must be interpreted in context, has been part of mainstream thought for decades.⁴⁴ William Landes and Richard Posner's important 1981 article, which includes various numerical examples applying formula (4), illustrates the phenomenon. For example, in one of the columns in their first table, as the hypothesized market elasticity of demand rises from 1.0 to 1.5 to 2.0, the market share necessary for a dominant firm to impose a given level of deadweight loss increases from 49% to 61% to 70%.⁴⁵ Put another way, a 49% share in the first market conveys the same market power

⁴⁴ See, e.g., Kaplow & Shapiro, *supra* note 12, at 1187-88; Pitofsky, *supra* note 2, at 1810-11; Schmalensee, *supra* note 14, at 1800-01. As section VI.E notes, courts have long recognized the point.

⁴⁵ Landes & Posner, *supra* note 13, at 955 tbl.1 (column for a market size of \$200 million).

as a 70% share in the third market; hence, a given share conveys significantly different levels of market power in the two different markets. In their second table, they consider two possible values for the market elasticity of demand and for rivals' elasticity of supply. When both are at their high levels, the market share required for a 20% price elevation is 61%, but when both are at their low levels, a share of 23% suffices.⁴⁶

One response to this situation — the one advanced here — is to forgo the market redefinition approach entirely. But courts and commentators (including Landes and Posner⁴⁷) do not adopt that stance. Accordingly, it is necessary to confront how one can make meaningful statements about the market power implications of any given market share in light of the large variation in the market share / market power relationship. Consider, for example, the following sorts of proclamations, which are undoubtedly familiar to the reader:

1. “The defendant’s market share of 80% is quite high, which usually means that it has a great deal of market power.”
2. “The defendant’s market share of 80% is indeed high; however, due to the aforementioned factors particular to this market, that share conveys much less market power than meets the eye.”
3. “Indeed, the defendant’s market share of 55% is lower than we might usually require, but in this market it conveys as much power as we would ordinarily associate with a much greater market share.”
4. “The plaintiff’s proposed market definition is too narrow. The defendant’s market share of 80% in that market conveys much more market power than exists in this case. Accordingly, we accept the defendant’s broader market definition, in which its share is only 30%, which we believe more accurately depicts its actual market power.”
5. “Although the defendant’s market share in the narrow market that the plaintiff favors is 80% and the share in the next broader market that the defendant advocates is 30%, the

⁴⁶ *Id.* at 958 tbl.2; *see id.* at 947–48 (“Thus the degree of market power in this example for a firm with an 8% market share is identical to that of a firm with an 80% share in a different market, because of offsetting differences in demand elasticities.”).

⁴⁷ *See infra* note 78.

truth is not near either of these extremes and instead lies somewhere in between.”

In reflecting on these statements, consider in particular the following phrases: (1) “usually means,” (2) “less . . . than meets the eye,” (3) “as we would ordinarily associate with,” (4) “conveys much more market power,” and (5) “the truth . . . lies somewhere in between.” Each communicates that the various market share figures under discussion typically correspond to some particular degree of market power — in the mind of the speaker and, the speaker presumably believes, in the mind of the audience for the remarks. For purposes of market definition and redefinition, statements like (4) are especially important. To reject a market as too narrow, as in this example, is to state that the market share in that market conveys a misleadingly high impression of extant market power, which presupposes that there is some level of market power associated with such a share (and, furthermore, that this level of market power is higher than exists). Likewise, to reject a market as too broad is to state that the market share in that market conveys a misleadingly low market power assessment, which similarly presupposes that it conveys some assessment.

Discussions of market shares therefore operate on the implicit but essential assumption that each share level is somehow (at least approximately) associated in all of our minds with some benchmark amount of market power. Otherwise, these statements would be meaningless. However, the preceding analysis in this section indicates that any given market share does not imply a unique level of market power, or anything close to it; moreover, it observes that this fact is broadly understood. How can these two propositions be reconciled?

B. Concept of a Standard Reference Market

This apparent contradiction can be avoided if discussions within the market definition / market share paradigm presume the existence of some normal or standard — albeit hypothetical — reference market that can be invoked as a benchmark. That is, in each of the five statements in section A’s illustration, it is implicitly assumed that there exists what I will call a standard reference market that provides the basis for comparison. When we see a given market share, it “usually means” something; this is the amount of market power that “meets the eye” or that we “ordinarily associate” with it. And if a share conveys “more” market power, it is more than that standard reference level. If the truth lies “somewhere in between” a market share of 80% and one of 30%, it must be that our impression of market power is associated with some intermediate market share in some such standard reference market, and that impression in turn is below the level of power similarly associated with 80% and above that with 30%. On occasion, dis-

cussions of the subject are more explicit, using language like “ordinarily associated with,” “typical,” “average,” or “benchmark” levels,⁴⁸ but usually references are almost entirely implicit (such as in saying that power is “greater” but not stating “than what”).⁴⁹

What is this standard reference market? In a sense, it does not much matter, as long as it provides a common benchmark for everyone.⁵⁰ Still, because it must be shared if communication employing market shares is to enable common understanding, the standard reference market that everyone implicitly must have in mind needs to be specified. But it never has been. Moreover, one suspects that most who speak in the language of market shares and market power have not really reflected much on the matter.⁵¹ As a thought experiment, consider administering a survey with one of the following two questions: “How much market power typically exists (specifically, how much will a dominant firm’s price exceed marginal cost) when a dominant firm has a market share of 70% in a properly defined market?”

⁴⁸ E.g., Kaplow, *supra* note 22, at 1820, 1827 & n.45, 1847; Schmalensee, *supra* note 14, at 1789, 1813–14; AREEDA, KAPLOW & EDLIN, *supra* note 14, at 492, 495; Kaplow & Shapiro, *supra* note 12, at 1188 & n.171 (also providing the most explicit prior reference to this idea). This concept has also played a central role in criticism of Landes and Posner’s claim that market power is systematically overstated on account of the frequent omission of direct consideration of some or all of the four factors that they identify as bearing on the market power implications of a given market share. See Kaplow, *supra* note 22, at 1819–21.

⁴⁹ To be clear, this practice is not confined to courts or to casual commentators but is virtually ubiquitous, even being employed by some of the most sophisticated analysts. For example, Franklin Fisher states: “A market will thus be well defined if and only if *the share measurements it leads to provide some reasonable index of the true power of the alleged firm . . .*” Franklin M. Fisher, *Diagnosing Monopoly*, Q. REV. ECON. & BUS., Summer 1979, at 7, 13 (emphasis added).

⁵⁰ There are some limitations. For example, suppose that we imagine the standard reference market to be one with a market elasticity of demand of 2.0 and a rivals’ supply elasticity of 1.0. In that market (and in any with a demand elasticity above 1.0) there is a finite upper bound to the Lerner index. (Here, if the market share S is 100%, it is clear from formula (4) that $L = 0.5$, so that the profit-maximizing price would be double the marginal cost.) For any such upper bound in the benchmark market, it is possible that there would be an actual market exhibiting even greater market power (simply suppose a 100% market share and a lower demand elasticity than assumed in the benchmark). A solution would be to take the market demand elasticity in the standard reference market to be 1.0 and use some positive supply elasticity. Then a share of zero would convey zero market power (the numerator in expression (4) would be zero and the denominator positive) and a 100% share would convey infinite market power (the numerator and denominator would both be one, and $L = 1$ implies an infinite price, as long as marginal cost is positive).

⁵¹ This gap has been mentioned. See, e.g., AREEDA, KAPLOW & EDLIN, *supra* note 14, at 492 (“Once we have a firm’s market share, we still need to determine what it implies about the existence or magnitude of that firm’s ability to charge more than a competitive price. . . . The starting point is the lay intuition that a firm with 100 percent of a market possesses market power. It almost surely has the power to restrict output and to charge more than the perfectly competitive firm. But how much more? Five or 500 percent? Such answers, never supplied, are essential if one is to infer market power from market share.”); *id.* at 498–99 (“[T]he power implications of any particular market share remain obscure. The courts have not stated how much power they believe to be associated with given market shares.”).

Or: “Suppose that a dominant firm is able to charge a price 50% in excess of its marginal cost. What would be your best guess as to its market share (supposing that the market is defined properly)?” The reader can ponder how closely the subjects’ answers are likely to cluster, and on what figures.

In order to specify concretely a standard reference market, two approaches come most readily to mind. One is that the standard reference market would in some sense be the average market. One way to make this notion more precise would be to use formula (4)⁵² and insert averages for the market demand elasticity and rivals’ supply elasticity. This method, of course, would involve great effort. It would be necessary to collect the data; for each data point, one would have to decide how to define the markets in which to make the measurements;⁵³ averages might be weighted in various ways or unweighted; and a time period and jurisdictional scope would need to be selected. Furthermore, whatever answer one obtained, it would need to be revised periodically to reflect shifts in demand, technological change, adjustments to trade policies, and so forth. That no such study has been done and that it is hard to imagine doing so reinforces the point that, although there must be an implicit standard reference market, it is hard to imagine that there actually exists a common, stable view as to its contours (supposing, that is, that when individuals say “typically,” “ordinarily,” and so forth, they have in mind some sense of what is true on average).

The other most natural possibility seems to be that we implicitly contemplate some abstracted, hypothetical, idealized market, perhaps some sort of textbook market. It still is necessary to define what it is. Textbooks do not in fact contain any such standard market. Maybe one would select some focal points, such as by inserting a market demand and a rivals’ supply elasticity in formula (4) that are each 1.0. In that case, a market share of 50%, for example, would imply that $L \approx 0.33$, so price would be 50% above marginal cost. Or one might simply make up a table: perhaps a 10% market share would be associated with price 1% above marginal cost, a 50% share with price 20% above cost, and so forth. Any such specification would be arbitrary, but that would not matter as long as there was common agreement to use the resulting translation table in all discussions relating market share to market power.

⁵² The logic of the present analysis does not depend on formula (4) or on defining market power using the Lerner index.

⁵³ Note a possible circularity if application of the standard reference market, now taken to be the average market, is necessary to decide how to define markets in the first place. See, for example, statement 4 in section A and also the discussion in Part IV.

To reinforce the point about the implicit need for such a standard reference market, reconsider the five statements in section A. (1) Asserting that “80% is quite high” would refer to some specific elevation of price above cost, and we can suppose that the elevation associated with an 80% share in the standard reference market table would indeed be high.⁵⁴ (2) “[T]hat [a] share [of 80%] conveys much less market power than meets the eye” is now clear, for the amount that “meets the eye” would be the amount in the table, and the speaker would be asserting that power is much less than what the table shows for the 80% market share. (3) That a 55% share “conveys as much power as we would ordinarily associate with a much greater market share” is likewise clarified (and would be even more definite if the speaker indicated how much greater a market share). (4) That a share of 30% “more accurately depicts [the defendant’s] actual market power” than a share of 80% would mean that the defendant’s power is closer to the value indicated on the standard reference market table for a 30% market share than for an 80% share. (5) That “the truth is not near either of these extremes [of shares of 80% and 30%] and instead lies somewhere in between” means that the level of market power is not close to the values on the table for either 80% or 30%, but has a value closer to the mean of these two values. In sum, invoking the concept of a standard reference market does in principle render coherent a wide range of typical statements in the market definition / market share paradigm that are incomprehensible without it. This conclusion, by its very phrasing, does require the existence of a *standard* reference market, that is, a common benchmark.

The argument of this Part is entirely formal and semantic, a specification of what understandings are logically presupposed by the language employed within a particular discourse. Nevertheless, the present argument is significant because the need for a standard reference market does not depend on particular word choices but rather on the sorts of ideas that are central to the market definition / market share paradigm. Furthermore, if a concept of a standard reference market is necessary to render the paradigm coherent even at a minimal level of successful communication, yet its existence is almost entirely unrecognized in any explicit manner and its content is wholly unspecified, then at the least there exists a critical chasm in prior analysis of the paradigm and in its innumerable applications in particular cases.

IV. DEFINING THE RELEVANT MARKET

Central to application of the market definition / market share paradigm is the selection of the so-called relevant market. Given the im-

⁵⁴ And, if it were not, individuals would no longer state that “80% is quite high.”

portance of this choice, the extensive academic commentary on the subject, and the number of times that relevant markets have been defined in practice, it is remarkable that the criterion by which one market definition is deemed superior to another has received little direct attention and analysis. That is, there is no canonical, operational statement of the standard for determining what constitutes a relevant market and, a fortiori, no developed underlying rationalization for whatever the principle might be. In this regard, recall from subsection II.B.2 that there is no corresponding concept in industrial organization economics that one might draw on to fill the void. (Some might regard the hypothetical monopolist SSNIP test employed in the U.S. Merger Guidelines and followed in other jurisdictions⁵⁵ to be an exception. However, as will be discussed below,⁵⁶ this test constitutes a mechanism, not a pertinent criterion; it generates a determinate market definition outcome that is not directly addressed to how much market power is increased by the merger under consideration, which is to say the merger's likely effect on price. Its method is also subject to the central criticism developed in this Part.⁵⁷)

This Part examines what seem to be the most plausible candidate criteria. First, the best market is taken to be that which most accurately measures market power. Because this approach discards relevant information and as a consequence can lead to an erroneous conclusion regarding liability, a second, essentially tautological notion of best is also examined: the relevant market is that which leads to the right legal answer.

Although stating possible criteria explicitly should facilitate future analysis and applications, the main purpose of the present exercise is destructive. For each of the proposed criteria, sharply stating it is suf-

⁵⁵ See U.S. MERGER GUIDELINES, *supra* note 5, § 4.1.1 ("The hypothetical monopolist test requires that a product market contain enough substitute products so that it could be subject to post-merger exercise of market power significantly exceeding that existing absent the merger. Specifically, the test requires that a hypothetical profit-maximizing firm . . . that was the only present and future seller of those products ('hypothetical monopolist') likely would impose at least a small but significant and non-transitory increase in price ('SSNIP') on at least one product in the market . . ."); *DG Competition Article 82*, *supra* note 1, ¶ 14 (SSNIP test used by the Commission for market definition, particularly in merger cases); Commission Notice on Market Definition, *supra* note 2, ¶¶ 15–17.

⁵⁶ See *infra* note 99.

⁵⁷ In brief, the SSNIP test asks how much a hypothetical monopolist of variously defined markets could elevate price, not how much the merging parties will be able to elevate price. Even if one were actually considering a monopolization (abuse of dominance) claim, the question should be the extent to which the firm in question can elevate price (say, as a consequence of the alleged abusive practice), not in what market a hypothetical monopolist thereof could elevate price to some specified degree. Moreover, regarding the actual monopolist, once one has estimated how much it can elevate price in its own (homogeneous goods) market, the analysis should be viewed as complete rather than employing this estimate as the input into a market definition process requiring further steps (and, as stated, answering a different question). See *infra* section VI.C.

ficient to make apparent its uselessness. Specifically, in each case, it becomes crystal clear that the only way to apply the criterion is first to know — more realistically, to formulate one’s best approximation of — market power.⁵⁸ (This phrasing is designed to emphasize that the analysis in this Part in no way assumes that information about market power is very good or comes from any particular source or uses any specific method. The estimate might derive from careful and persuasive econometric analysis, or the only evidence may be fairly vague impressions of how consumers are likely to behave. The claim is simply that, whatever is one’s best estimate — or, in cases with limited evidence, “guesstimate” may be more appropriate — the most one can hope for in defining a market by any of the stated criteria is to employ that estimate in making the choice.⁵⁹)

This market power estimate must be used to choose among market definitions under any plausible approach. Yet the *raison d’être* for the market definition enterprise is to provide a basis for inferring market power. Accordingly, why would one ever use one’s best estimate of market power to undertake further effort in order to provide a clue as to market power when that clue contains no information that was not part of the initial (best) estimate?

Put another way, any plausible application of the market definition / market share paradigm assumes the conclusion that it is designed to help reach, rendering it pointless. Moreover, as noted, the seemingly attractive approach of identifying the market that provides the most accurate indication of market power involves discarding some of the initial information and thus can lead to an incorrect outcome. This argument means that one should never engage in the market definition enterprise as part of the process of inferring market power.

A. Choosing the Best Market Definition

This section considers techniques that resemble what is ordinarily done when choosing a relevant market. To begin, the set of available market choices is limited to those that would be regarded as conventionally plausible. This framing obviously raises the question of how the permissible set is selected — a serious problem, but one whose res-

⁵⁸ One could generalize the discussion in the text along other dimensions without affecting the analysis. For example, if price discrimination is possible, the degree of exercised market power might vary across customer groups, so there would be multiple estimates and one could define different markets for different groups if one wished.

⁵⁹ As will be discussed in section VI.D, engaging in a market redefinition inquiry may help to hide (including from oneself) just how poor the information inputs are, but there is no magical process by which the technique is able to create something (better information) that does not exist in the first place.

olution is immaterial for present purposes and thus will not be pursued further in this section.⁶⁰ Instead, the choice set will be taken as given; a reader may suppose that it is determined in whatever way is most appealing.

Furthermore, attention will be confined to the choice between two markets, one narrower (and at least somewhat too narrow) and the other broader (and at least somewhat too broad).⁶¹ If there were three or more available choices, as there often are, one would begin by eliminating all but the closest two on each side of the truth. For concreteness, one may wish to think of the *Cellophane* case,⁶² where the narrower market is cellophane and the broader one is all flexible wrap-wrapping materials, or of a case involving a manufacturer of canoes, where the narrower market is just canoes and the broader market is, perhaps, all small, unpowered boats, which would include row boats, kayaks, and perhaps some sailboats.

1. *Best as Most Accurate Measure of Market Power.* — Choosing that market which most accurately measures market power is probably close to what most analysts, enforcement agencies, and courts have in mind when they attempt to define the relevant market in a particular setting. At the outset, it is important to elaborate the criterion. Specifically, when choosing between two markets (one too narrow and the other too broad), just what does it mean to say that one choice better measures market power than does the other?

This question has received little direct attention, but most would probably answer it in roughly the following manner: granting that any choice is inevitably imperfect, the best choice is that which produces the smallest error, that is, which yields a market power inference that is closest to the actual level of market power. Hence, in choosing between a narrower and a broader market (where, as mentioned, we are supposing that the truth lies somewhere in between), one would ask

⁶⁰ Because the only purpose of defining a relevant market is to make an inference about market power, the only truly pertinent criteria for determining the set of admissible markets would be that they be selected in such a way that the subsequent choice among them yields as accurate an inference as possible, which is question-begging in the same way as is choosing the best market from among the finalists.

⁶¹ It is possible, depending on how the choice set is determined, that all the elements will be too broad or all too narrow. For example, the narrowest plausible market may be viewed as the homogeneous goods market, the firm's share in that market may not be very high, and yet, due to a combination of fairly inelastic market demand and rivals' supply, the firm could have significant market power, much more than indicated by its share (where the indication might be based on use of the standard reference market). In that case, all markets would be too broad in the sense that the firm's share in them would understate its market power. This possibility constitutes a further defect in the approaches examined in this section, but will be set to the side. (Of course, choosing between markets, both of which, say, are too broad, is easy if one knows this fact: one simply picks the narrower of the two.)

⁶² *United States v. E.I. du Pont de Nemours & Co.*, 351 U.S. 377 (1956).

whether the inference from the larger market share in the narrower market overstates market power by more than the inference from the smaller market share in the broader market understates market power.⁶³ If the lesser error lies with the former choice, then the narrower market is the relevant market; if the latter minimizes error, then the broader market is best.

It is also necessary to say what we mean by the inference from a given market share in a particular market. If we had in mind employing formula (4) from Part II if the narrower market is the homogeneous goods market and formula (5) for all other markets, using expressions (6) and (7) to transform all other markets back into the homogeneous goods market, then we know from prior analysis that we would obtain precisely the same (and correct) market power inference for any market we might define. Accordingly, the choice of a market is moot — one would always choose the pure, homogeneous goods market and proceed to estimate the market elasticity of demand and rivals' elasticity of supply in that market.

Because a serious choice of the relevant market can arise only if we are engaging in some other mode of inference, we need a way of saying what market power we associate with the market shares in question (the higher share in the narrower market and the lower share in the broader market). In light of the preceding remarks that draw on the analysis in Part II, the problem is formidable (if solvable at all). However, such problems are being set to the side for purposes of this Part, so let us suppose that some sort of answer (largely taken for granted in conventional analysis) is offered. Regarding the communication of such answers in market share terms, we may refer to Part III, where it is suggested that market shares can have common meaning only if one invokes a standard reference market that provides an accepted way of translating market shares into market power. Thus, when one states, for example, that a firm's market share in the narrower market overstates market power, an aspect of this assertion is that one is drawing some sort of market power inference from the firm's market share in the first place, and the standard reference market was presented merely as a device for making this notion concrete. In any event, for the present discussion all that is required is that there exists some method of inferring market power from the market share in the narrow and broad markets, and the reader should feel free to imagine any preferred way of doing so because the analysis to follow does not depend on the particulars.⁶⁴

⁶³ See Kaplow, *supra* note 22, at 1828.

⁶⁴ As section III.A explains, there must be some notion like a standard reference market for typical statements about the implications of market shares to be made sense of in a manner that communicates. Since the standard reference market table, as explained in section III.B, can be

We are now ready to answer our question of which market definition is best. To begin, our criterion is that we choose that market which produces the least error in our resulting inference. By error, we mean the degree to which a particular market definition mismeasures market power. For concreteness, we can let MP_{narrow} represent our market power inference in the narrow market, MP_{broad} represent our inference in the broad market, and $MP_{estimate}$ indicate our best estimate of the actual degree of market power.⁶⁵ Accordingly, the inquiry asks which error is smaller, $|MP_{narrow} - MP_{estimate}|$ or $|MP_{broad} - MP_{estimate}|$.⁶⁶ Consider the following explicit procedure:

- *Step 1:* Drawing on all available information, formulate one's best estimate of market power, $MP_{estimate}$.
- *Step 2:* For the narrow market, determine what market power, MP_{narrow} , is associated with the higher market share in that market; likewise, determine MP_{broad} for the lower share in the broad market.⁶⁷
- *Step 3:* Compute $|MP_{narrow} - MP_{estimate}|$ and $|MP_{broad} - MP_{estimate}|$ and determine which difference is smaller.
- *Step 4:* If it is the former, choose the narrow market as the relevant market; if the latter, choose the broad market.
- *Step 5:* Make the appropriate market power inference. That is, if we choose the narrow market, then we take market power to be MP_{narrow} , and if we choose the broad market, then market power is deemed to be MP_{broad} .

constructed in any manner one likes, the method really is nonrestrictive. (One could add further layers. For example, if a higher share in a narrower market is known to overstate market power, one could make some downward adjustment in the power one infers; likewise for the lower share in the broader market. At the end of the day, there are two possibilities: Either the inference from the higher share in the narrower market differs from the inference from the lower share in the broader market, as is commonly supposed, in which case the argument in the text to follow is fully applicable because it does not depend on the magnitude of the difference. Or the inferences are the same, in which case market definition is irrelevant and thus pointless, in line with this Article's central claim.)

⁶⁵ Following Part II, each of these measures may be taken to be corresponding values of the Lerner index. One could instead substitute some other chosen measure of the extent of profit-maximizing price elevation without affecting the argument.

⁶⁶ One could refine this statement to incorporate the uncertainty of estimates, asymmetry and nonlinearity of error costs, and so forth without materially altering the argument to follow.

⁶⁷ As mentioned, the analysis does not depend on whether we use the standard reference market or some other process, as long as we do not essentially ignore the differences in the market definitions and use some other means, such as that examined in section II.B, to obtain the same market power inferences in both markets — in which case market definition is moot, as previously noted.

The method of determining which market definition is best in the sense of providing the most accurate measure of market power, we can now see, has the feature that the first, essential step consists of developing our best estimate of market power — that is, using all available information, however good or limited it may be, to formulate the most accurate assessment of market power that we can.⁶⁸ This step is a logical prerequisite because, in choosing the best market, we have to determine the degree to which the market power inference in the narrow market overstated our best estimate of the actual degree of market power and by how much the market power inference in the broad market understated this estimate. For both determinations, it is presupposed that we have in mind some such best estimate. (How can one measure, for example, the extent to which the market power inference in the narrow market overstates the truth without employing our best estimate of the truth?) Without this assessment, there is no meaningful way in which we could say that one measurement error is greater or less than the other.⁶⁹

It follows immediately that choosing a relevant market is a useless exercise. Since the conclusion — a market power inference from a market share in the market deemed to be the relevant one — contains no information helpful in obtaining an estimate of market power that is not already contained in our initial estimate of market power, nothing is gained. That this point is essentially a tautology hardly detracts from its validity.⁷⁰

However, matters are worse. Unless by fluke one of the plausible, conventional markets yields a market power inference that coincides with our best estimate of market power — that is, the market share in one of the markets leads us to infer a level of market power that equals our best estimate — we are left with an inaccurate market power inference.⁷¹ If we selected the narrow market as best, market

⁶⁸ What constitutes a best estimate will vary by context (for example, screening by enforcers, deciding on summary judgment, or full adjudication). For purposes of the present argument, the best estimate simply refers to that which is best given the information set under consideration (that is, the same information one might have used to define the market in any other manner).

⁶⁹ For further discussion, see section VI.D.

⁷⁰ The only way our conclusion could contain information not present in our initial best estimate of market power would be if some information irrelevant to market power went into deciding which market definition was better (or if some information relevant to market power was given more or less weight than that appropriate in forming a best estimate of market power). In that case, however, the inference of market power drawn from the market share in the chosen market could only be worse than our initial estimate. (This claim also is a tautology, in that our initial estimate was stated to be the best estimate we could derive from all pertinent information.)

⁷¹ Even then, the coincidence would hold only for a particular firm with a given share in that market. If the firm's share were different, or if we wanted to assess some other firm that sells the same product (but has a different market share), or if we were considering possible mergers between firms that sell the same product, then only by chance would the same market definition

power is overstated, specifically, by $|MP_{narrow} - MP_{estimate}|$, and if we selected the broad market as best, market power is understated, by $|MP_{broad} - MP_{estimate}|$.

The resulting error is due to the fact that choosing the best market in this fashion throws away information. If one is told only that the narrow market was chosen, all one knows is that market power is somewhere between the level of power indicated by the higher market share in that market and the midpoint between that level of power and the level associated with the lower market share in the broad market. Likewise, if all we are told is that the broad market was chosen, then we know only that market power is somewhere between that midpoint and the lower level of power indicated by the lower market share in the broad market. In any case, even this level of refinement is not employed. Instead, we use whichever endpoint corresponds to the selected market. That is, once the relevant market is defined, we proceed with our market share inference based on that market. In essence, therefore, we have decided to ignore the more refined estimate, $MP_{estimate}$, that we used in choosing one market instead of the other as the relevant market, substituting instead either MP_{narrow} or MP_{broad} . In other words, we have chosen the lesser of two evils when a virtuous alternative was not only available but was necessary to consult in order to assess which mistake was smaller.

That choosing between two markets involves inaccuracy is familiar and is ordinarily seen as unavoidable at least to some extent.⁷² But such is not the case. A more straightforward approach would eschew the market definition / market share paradigm entirely. Since one starts in step 1 with a best estimate of market power, it is both easier and most accurate to stop — to quit while one is ahead. Simply use that best estimate of market power and never define markets.

The error that results from instead choosing the market that most accurately measures market power can, unsurprisingly, have consequences. Suppose, for example, that the broad market is selected and that the market power inference from the lower market share in that market is still fairly high, but barely below the threshold above which

continue to work. (Suppose, for example, that the initial firm in question has a profit-maximizing price that is 5% above marginal cost, that this corresponds to a 20% market share on the reference table, and that there happens to exist some market definition that results in the firm having a 20% share in the resulting market. Now, consider another firm that sells the same product and has a 50% share in the market that we just selected; it might have the ability to raise price 10%, but the reference table might indicate instead that a 25% elevation is associated with a 50% market share.)

⁷² As mentioned in note 64, the only way to avoid the problem is to use a process that interprets shares in the different markets in such a way as to reach the same conclusion about market power — and one corresponding to our original best estimate. In that event, market definition is playing no real role in the process of inferring market power.

action is appropriate and called for by the legal rule. It may well be that one's best estimate of market power is above this threshold. Hence, one would abstain from action that one's best estimate indicates is appropriate. Likewise, when the narrow market is best, one might be led to intervene when one's best estimate indicates that such action is a mistake.

For concreteness, suppose that we wish to condemn a horizontal merger if and only if it will result in a price increase of at least 5%. Suppose further that our best guess of the price increase made possible by the augmentation of market power is 10%, that in the narrow market we would have inferred sufficient market power to raise price by 25%, and that in the next broader market (which is much larger, it so happens) power seems sufficient to raise price by only 1%. The broader market is best in the sense that it involves less inaccuracy in measuring market power, but deeming it the relevant market produces the wrong legal outcome.

In sum, choosing as the relevant market that which most accurately measures market power has three strikes against it: It presumes that we already have our best estimate of market power, when the whole point of the exercise is to generate a market power inference. It leads to results that are sometimes erroneous (and predictably and avoidably so). And it is entirely unnecessary to undertake the additional costly effort to determine MP_{narrow} and MP_{broad} . Instead of market definition and redefinition being pointless, it is worse.

2. *Best as Yielding the Right Legal Answer.* — The unnecessary error identified in subsection 1 has an obvious remedy: do not deem the best market definition to be that which measures market power most accurately but instead take it to be that which gets the right legal answer. How would one go about doing this? In place of the procedure outlined in subsection 1, employ the following alternative:

- *Step 1:* Drawing on all available information, formulate one's best estimate of market power, $MP_{estimate}$.
- *Step 2:* Use this market power estimate (along with whatever other information and application of legal rules are required) to determine the right legal answer. (As with market power, right is not any absolute truth, but simply the best one can do given available information.)
- *Step 3:* For each available (plausible) market definition, determine what market power would be inferred from the share in that market — for example, MP_{narrow} and MP_{broad} .
- *Step 4:* For each such market power inference, determine what legal conclusion would ultimately follow.

- *Step 5*: Deem the relevant market to be the market that yields the right legal answer.⁷³

This procedure works.⁷⁴ In particular, it gets the right legal answer and thus avoids the shortcomings of the more natural method examined in subsection 1. But it works due to its circular construction. This method renders market definition transparently superfluous (and involves additional effort). It is obvious that we should declare victory at the end of step 2, wherein we have used our best estimate of market power (along with whatever else is pertinent) to determine the correct legal conclusion. All the steps that follow are a blatant attempt to reverse engineer a market definition solely for the purpose of being able to say that we have defined a market (a motivation examined further in section VI.E). To undertake them is senseless. To abstain is to accept that one should never define markets.

B. *What to Do?*

Each of the candidate criteria for defining the relevant market leads to the same conclusion: abandonment of the market definition / market share paradigm. Moreover, a common reason supports this judgment, namely, that the first step under both approaches involves deriving one's best estimate of market power, which is obtained prior to and without regard to any subsequent market definition exercise. Furthermore, the first criterion discards information and thus will sometimes produce inferior decisions, and the second is patently tautological. And each requires additional effort.

Are there any alternatives? It seems fairly clear that there are none that could avoid the central critique.⁷⁵ Any choice among preexisting

⁷³ Multiple choices may work (for example, if the correct conclusion is that market power is insufficient, many broad markets may yield this result), in which case any of these may be selected.

⁷⁴ Actually, not necessarily. *See supra* note 61 (mentioning the atypical case in which even the narrowest homogeneous goods market is too broad).

⁷⁵ Part III's invocation of a standard reference market supplies another possible criterion for defining relevant markets, but one subject to essentially the same critique. Consider the following procedure:

- *Step 1*: Drawing on all available information, formulate one's best estimate of market power, $MP_{estimate}$.
- *Step 2*: Use a standard reference market table to look up that level of market power and read off the corresponding market share S^* , that is, the market share in the standard reference market that implies the level of market power identified in step 1.
- *Step 3*: Find, or more likely create — jerry-rigging as necessary — a market in which the firm has the stated market share S^* . For example, it may be necessary to draw lines between virtually identical products, include distant substitutes while omitting close ones, or include fractions of some substitutes in order to hit the target. This market would be deemed the relevant market.

markets, or even the creation of a market for market definition purposes, must be guided by the desire to make as accurate a market power inference as possible — or, even more to the point, to lead to a correct result. If one starts with one's best estimate of market power, it is difficult to understand how the abstract act of defining a relevant market (or anything else for that matter) can do better.

Note that it does not follow from the analysis in this Part that thinking in terms of market abstractions is never helpful. The most useful application pertains to the homogeneous goods market examined in subsection II.B.1. In such a setting, there is a formula, expression (4), that indicates what information we need to infer market power. As explained in subsection II.B.2, any attempt to redefine the

-
- *Step 4:* Having completed the market definition process, we must still make a market power inference. How is this done? We take the market share in what we have determined to be the relevant market, which is S^* , and ask ourselves what level of market power it implies. We can answer that question by looking up S^* on our standard reference market table and finding out what level of market power is associated with that share. And that is our answer. It will, of course, be $MP_{estimate}$.

The answer we obtain at the end of step 4 is precisely the answer with which we began in step 1, our now-familiar best estimate of market power (and, as before, it is apparent that there is no plausible way to proceed to step 2 without first formulating such an estimate). And the market share, S^* , that we contemplate at the beginning of step 4 is just the share we read off the standard reference market table in step 2. Unless we make a careless error, we entirely undo in the latter portion of the technique precisely what we did in the first part.

Note, however, that this process, like that in subsection A.2, does have the virtue that it obtains a correct answer — that is, as good an answer as is feasible given the available information. In essence, it substitutes, for the familiar all-or-nothing choice whether to combine markets (*see, e.g., AREEDA, KAPLOW & EDLIN, supra* note 14, at 492; Carlton, *supra* note 13, at 16; Fisher, *supra* note 49, at 16; Kaplow, *supra* note 22, at 1828; Kaplow & Shapiro, *supra* note 12, at 1092; Pitofsky, *supra* note 2, at 1811), a jerry-rigged market definition that is entirely arbitrary except for the fact that, on the standard reference market table, the firm's share yields the correct market power inference. Thus, the fact that this process produces the right answer is made possible by its lack of resemblance to ordinary methods of drawing market boundaries. (Indeed, it does not even matter under this criterion which of many other possible products are combined if, for example, there is a need to generate a lower market share.)

One might imagine that, instead of using all available information on extant market power, which all these procedures require, we might jiggle and adjust the market definition, comparing alternatives, until we obtain a market that “looks” like the standard reference market. Yet it should be apparent that such is not possible. Market definitions are abstractions to begin with, and either (A) the standard reference market is a table reflecting empirically generated averages across all markets (averaging over not only apples and oranges, which are quite similar and might even be in the same market, but also cell phones, steel, candy, and houses) — and it is hard to give meaning to the determination of which of two or more market definitions more resembles, under some Rorschach test, such an average — or (B) the standard reference market is a table that is derived in an entirely arbitrary manner, by writing down a series of numbers generated in some fashion to have the desired monotone relationship (higher market share, more market power). Furthermore, if one employed any such resemblance test, it has the obvious deficiency relative to the more explicit techniques considered here of being unlikely to reach as good an answer to the original market power question. *See also infra* section VI.D (further elaborating the point that markets cannot be defined by how they may appear to our senses).

market only generates confusion;⁷⁶ to render the exercise coherent, one must entirely undo the redefinition. Other ways of thinking about markets, substitutes, and the like will be discussed below, both in Part V and in sections VI.A and VI.C. It should be clear throughout, however, that using these familiar abstractions in thought in order to come up with a best estimate of market power in no way involves defining a relevant market. If one finds that one has stumbled into such an exercise, one should turn back to avoid wasting time and leading oneself astray.

In prior writing, many commentators and courts have endorsed, in varying degrees and using various methods, the idea that we should interpret market shares in a manner that takes into account the particular features of the products and markets in question. A few have pushed especially hard on this front, arguing that, whatever market one chooses, one should obtain precisely the same market power inference from the different market shares in each different market.⁷⁷ Just how one would make such inferences is obscure. (As just mentioned, the analysis in section II.B indicates that one can interpret market shares meaningfully only in the homogeneous goods market, and any rigorous attempt to do so in other markets involves essentially undoing the market broadening in order to put oneself back into the homogeneous goods market.) It is much harder, however, to understand why one would bother to define markets — and advocate some market definitions over others — if the chosen definition should not matter, and it is also hard to identify just what role, if any, is played at the end of the day by the market one has chosen. The present discussion suggests that there is no useful way to choose the market, and, since any choice presupposes that one has in hand a best estimate of market power, there is no need to create some careful, subtle process by which one adjusts market power inferences from the market share in the chosen market in order to reach an ultimate, correct conclusion.

Prior commentary varies greatly in its attitudes toward how critical or sanguine one should be about the market definition process and under what circumstances one should pursue alternatives. Landes and Posner are among the earliest and strongest critics of market definition practice, although they seek to reform rather than replace it.⁷⁸ Other

⁷⁶ See, e.g., Kaplow, *supra* note 22, at 1827–28; Schmalensee, *supra* note 14, at 1804.

⁷⁷ See *supra* note 33.

⁷⁸ One of their claims that goes furthest appears in their introduction, where they state: “Among other things, we show how the inference of market power can be adjusted so that defining a market broadly or narrowly will not affect the inference [of market power] that is drawn” Landes & Posner, *supra* note 13, at 938; see also *id.* at 978. However, although they do discuss at length factors that bear on the magnitude of market power (a major focus of their article), they nevertheless advocate particular market definitions as leading to better — and thus presumably different — inferences than others produce, and they do not in fact develop a way to achieve

commentators have been critical as well.⁷⁹ Still others have urged caution but adopt a more moderate or mixed stance toward the market

the asserted neutrality. (Throughout, as Schmalensee emphasizes in his commentary, Landes and Posner seek to employ formula (4), which is valid only for the homogeneous goods market, to allow inferences in all manner of markets, an approach explained in subsection II.B.2 to be invalid. Schmalensee, *supra* note 14, at 1797.) Despite the force of many of their arguments, their conclusion is rather tame:

When those elasticities are known or knowable, our analysis provides a method of estimating market power in quantitative terms. The analysis thus should be helpful to enforcement agencies in setting priorities and allocating their resources, and to courts in those cases, which may be few, where estimates of the elasticity of demand and supply are obtainable in a form usable in the litigation process.

Landes & Posner, *supra* note 13, at 983. The latest edition of Posner's book on antitrust law also sends mixed signals. Compare RICHARD A. POSNER, *ANTITRUST LAW* 147 (2d ed. 2001) ("The importance that antitrust law attaches to defining a market is another consequence of the law's failure to have developed an approach at once genuinely economic and operational to the problem of monopoly. If we knew what would happen if a group of sellers raised their prices — if we knew how rapidly the price increase would be undone by the response of other sellers — it would be redundant to ask whether the group constituted an economically meaningful market. In other words, if we knew the elasticity of demand facing those sellers, we would know almost all we needed to know."), with *id.* at 148 ("It is lack of confidence in the ability to measure elasticities reliably by the methods of litigation that necessitates definition of the market.")

⁷⁹ Among the sharpest critics are Farrell & Shapiro, *supra* note 9, at 1 (arguing that the market definition "approach can be clumsy and inaccurate in industries with differentiated products where the theory of harm is related to unilateral . . . effects"); Fisher, *supra* note 49, at 16 ("If one always remembers this, there is no positive harm in engaging in the market definition exercise. Indeed, viewed correctly, arguments concerning whether products are in or out of the market which are made in terms of demand and supply substitutability and hence in terms of constraints are exactly the arguments which one would have to decide in looking at the constraints directly. The trouble is that it is too easy to forget what the analysis is all about. By focusing on whether products are in or out of the market, one converts a necessarily continuous question into a question of yes or no. The temptation is to regard products which are in as all counting equally and products which are out as not counting at all."); Fisher, *Economic Analysis*, *supra* note 41, at 132–33 (discussing *Cellophane*: "Nothing is to be gained, and much information may be lost by an attempt to force these facts into the Procrustean bed of market definition, insisting on an answer that other flexible wrapping papers are either in or out of the market. If one insists on expressing the facts in such language, then the correct answer is that such papers are 'in' at high prices for cellophane and 'out' at low ones. Evidently, this is not a very useful way to proceed."); and Schmalensee, *supra* note 14, at 1798–99 ("This approach focuses attention on market definition, not on the fundamental question of market power. It has a number of intrinsic weaknesses, some of which are noted by Landes and Posner, and thus is not always reliable or even usable." (footnote omitted)); *id.* at 1808 ("Other approaches to proving the existence of substantial market power are no less valid than the market share approach."). I have also been critical in my own prior writing, beginning with Kaplow, *supra* note 22, at 1827 n.45 ("In principle, of course, the most natural approach to determining extant market power on the basis of a thorough economic analysis would be to cast aside a priori impressions of market power associated with particular market shares and to derive the measure of market power from scratch, using all the relevant factors." (emphasis omitted)); *id.* at 1828–29 ("But if sufficient information were available to make that comparison [between errors in market power inferences from the two markets] possible, it could be put to better use: a more sensible and straightforward approach would be to discount directly the perceived market power by the degree to which failure to combine markets overstates market power."); *id.* at 1830 ("Combining markets and then measuring elasticity is no less difficult than directly determining the proper adjustment to one's inference of market power in light of the existence of close substitutes. In fact, the two inquiries are theoretically equivalent." (footnote omit-

ted)). See also AREEDA, KAPLOW & EDLIN, *supra* note 14, at 492; Kaplow & Shapiro, *supra* note 12, at 1171 (“[O]ne often needs to know the right answer — that is, how much market power exists — in order to know which market definition is best.”). On geographic market definition in particular, see Kaplow, *supra* note 22, at 1839 (“A plaintiff cannot refute the hypothesis that foreign production costs are lower than domestic production costs [and thus overcome Landes and Posner’s geographic market definition presumption] without reliable information concerning both sets of costs. If information on domestic production costs were available, one could *directly* compute the measure of market power suggested by Landes and Posner . . .”), and AREEDA, KAPLOW & EDLIN, *supra* note 14, at 497 (“Imports may have occurred only because [the defendant’s] price exceeds the competitive level, which would make clear that actual or future imports do not prevent supracompetitive pricing. Of course, if we know that, we don’t need any market definition.”).

For additional strong remarks in works not focused on criticism of market definition, see Timothy F. Bresnahan, *Empirical Studies of Industries with Market Power*, in 2 HANDBOOK OF INDUSTRIAL ORGANIZATION 1011, 1049 n.44 (Richard Schmalensee & Robert D. Willig eds., 1989) (stating in passing that “[t]hese [market power] questions are well-posed, even where the usual method of answering them, defining a ‘relevant market’ and calculating market shares in it, is senseless”), and Frank H. Easterbrook, *The Limits of Antitrust*, 63 TEX. L. REV. 1, 22 (1984) (“An inquiry into power does *not* entail the definition of a ‘market,’ a subject that has bedeviled the law of mergers. Usually the search for the ‘right’ market is a fool’s errand. The seller of 100% of a particular good may have no power if consumers have substitutes or if rivals can make the good as cheaply. On the other hand, there may be tens of possible markets, each offering a little insight into conditions of competition. Market definition is just a tool in the investigation of market power; it is an output of antitrust inquiry rather than an input into decisions, and it should be avoided whenever possible.”). See also Janusz A. Ordover & Robert D. Willig, *The 1982 Department of Justice Merger Guidelines: An Economic Assessment*, 71 CALIF. L. REV. 535, 536 (1983) (in an article largely praising the 1982 Guidelines, including its approach toward market definition, observing: “Arguments for and against a merger that turned upon distinctions between broad and narrow market definitions have always seemed theologic rather than economic. The focus on market definition has seemed to be an inadequate substitute for, and a diversion from, sound direct assessment of a merger’s effects.”).

A particularly harsh attack on market definition appears in Richard S. Markovits, *Predicting the Competitive Impact of Horizontal Mergers in a Monopolistically Competitive World: A Non-Market-Oriented Proposal and Critique of the Market Definition-Market Share-Market Concentration Approach*, 56 TEX. L. REV. 587, 599–603, 698–728 (1978). His extensive critique, which focuses on horizontal mergers in differentiated products markets, is difficult to interpret. Although some of his bottom lines overlap with some of those reached here, the internal analysis largely differs, in part because many of his arguments draw on his idiosyncratic theories of imperfect competition that he develops in the article (which does not cite the then-leading, modern treatments of the subject by Michael Spence and by Avinash Dixit and Joseph Stiglitz). Some indications of the disconnect with standard economic analysis are that, in 145 pages, he does not use the terms elasticity or cross-elasticity (even though his analysis focuses on these ideas), that a large portion of his argument is predicated on firms’ pricing individually to consumers (that is, highly refined price discrimination, without regard for problems of identification or arbitrage), and that, when he writes about coordinated oligopolistic pricing at length, he omits the central point that mergers reduce the parties’ incentive to cheat. And some of his other reasons are fairly familiar imperfections in the market definition process that induce most to be careful and seek alternatives when available rather than to reject the method entirely. That said, Markovits’s emphasis in differentiated products mergers — on how many buyers whose first choice is one merging party’s product have the other party’s product as their second choice — anticipates the more formal developments on the subject in subsequent decades. See sources cited *infra* note 107. An internal summary of his main points on issues pertaining to this Article is as follows:

The failure of the traditional approach is partially attributable to the inability of market share data to reveal much information about a large number of highly relevant factors

definition / market share paradigm.⁸⁰ But none offers the sort of case presented here against the foundations of the entire enterprise.

to which it does not even purport to relate. Thus, market shares have little bearing on (1) the amount of potentially pro-competitive efficiencies the merger will generate, (2) many of the factors that influence the effectiveness of potential competition (*e.g.*, the rate of growth of market demand), and (3) many reputational, informational, and other factors However, the argument [here] justifies a far more radical departure from the traditional approach. The market share-market concentration approach ultimately must be rejected because even when a clear break in the chain of competitors exists (1) some firms outside the associated market will be in a position to compete for some of its buyers, (2) various firms within the market will not be universally or equally competitive with each other, and (3) different firms in the same market, and a fortiori in different markets, will have very different distributions of competitive ranks, advantages, and disadvantages.

Markovits, *supra*, at 727; *see also id.* at 590, 603, 699.

⁸⁰ *See, e.g.*, AREEDA, HOVENKAMP & SOLOW, *supra* note 2, at 108 (“Chapter 5B then explores the ‘alternative’ or ‘direct’ indicators of market power, which can be independent of market definition and are sometimes superior to it. . . . While the data in actual cases are often too incomplete to permit the more technical economic procedures, better understanding of the economics illuminates the less technical — but also typically less accurate — market definition methodology that antitrust litigants and tribunals customarily employ. . . . We will see that market definition may not be necessary to prove market power; may not itself suffice to establish the degree of market power; and may be redundant or misleading once the facts necessary for an accurate definition have been obtained.”); *id.* at 136 (“Still, there may be no alternative to the market-definition approach when no other observable facts establish the existence and degree of market power more directly and with tolerable accuracy.”); Areeda, *supra* note 7, at 565 (“We have no choice but to do the best we can in defining the market and to moderate the inferences we draw from prospective market shares”); WHINSTON, *supra* note 43, at 78–79, 99–100; Baker, *supra* note 2, at 131 (“Market definition is least useful when market shares would not be strongly probative of market power or anticompetitive effect, while direct evidence as to market power or anticompetitive effect is available and convincing.”); *id.* at 162 (“But if direct evidence is unavailable or not strongly probative, making market definition necessary or useful, the court will likely need to define the market through undertaking a hypothetical prospective market definition analysis (albeit relying importantly on retrospective evidence).”); Carlton, *supra* note 13, at 3 (“Market definition is a crude though sometimes useful tool for identifying market power.”); *id.* at 17 (“Although the clear theoretical construct of market definition can guide one, the absence of estimates of the demand (or cost) system subject this exercise to possible error and arbitrary judgments.”); Pitofsky, *supra* note 2, at 1834 (“The remainder of this Article is devoted to an effort to present an approach to relevant market definition that neither understates nor overstates market power.”). Finally, some have proposed to me that at least a number of economists hold more critical views of market definition than suggested by the literature, although the particular reasons for such skepticism are not apparent. In assessing merger review in the European Union, Kai-Uwe Kühn suggests that:

[E]mpirical research on the problem of market definition and the effects of mergers has made economists realize that the information needed to determine the unilateral effects of mergers requires as much or less information than an economically rigorous estimation of the market boundaries. This means that it is in principle no more difficult to estimate price effects of mergers directly, skipping the market definition step.

Kai-Uwe Kühn, *Reforming European Merger Review: Targeting Problem Areas in Policy Outcomes*, 2 J. INDUSTRY, COMPETITION & TRADE 311, 315 (2002). However, he concludes his brief discussion of the subject on a more moderate note: “But market definition should be regarded as part of the general analysis of the competitive effects of the merger.” *Id.* at 316.

V. ELASTICITY VERSUS CROSS-ELASTICITY

The analysis in Part IV establishes that, once a best estimate of market power is formulated — however good or bad such an estimate might be — it should be treated as the market power conclusion rather than used as an input into a market definition process that, as ordinarily conducted, loses information and can produce inferior ultimate outcomes. Unfortunately, actual practice tends to be even worse because the fixation on market definition redirects our attention from the evidence that bears most directly on market power (and how that evidence is economically relevant) to other evidence and relationships that provide inferior guidance on the question at hand.

Specifically, for over half a century, U.S. antitrust law (now mimicked by other regimes) has focused on the cross-elasticity of demand.⁸¹ The Supreme Court in *Cellophane* famously pronounced that, in defining markets, one should focus on the reasonable interchangeability, also known as the cross-elasticity, between products in the two markets one contemplates combining.⁸² Nor is this attention surprising: if one takes for granted that the question is whether two product groupings should be treated as if they constitute a single market or two separate markets, it seems natural to focus on the extent to which con-

⁸¹ For details of the historical development, see Werden, *supra* note 13, at 398–401. For an important example in the European Union, see Commission Regulation 802/2004 of 7 April 2004 Implementing Council Regulation 139/2004 on the Control of Concentrations Between Undertakings, Annex I, 2004 O.J. (L 133) § 6.I, at 15 (EC) (“A relevant product market comprises all those products and/or services which are regarded as interchangeable or substitutable by the consumer, by reason of the products’ characteristics, their prices and their intended use. . . . Factors relevant to the assessment of the relevant product market include the . . . cross-price elasticity of demand . . .”).

⁸² “What is called for is an appraisal of the ‘cross-elasticity’ of demand in the trade.” United States v. E.I. du Pont de Nemours & Co., 351 U.S. 377, 394 (1956); *see id.* at 400 (further use of cross-elasticity); *id.* at 395 (“reasonably interchangeable”); *id.* at 404 (“reasonable interchangeability”); *see also* Brown Shoe Co. v. United States, 370 U.S. 294, 325 (1962) (“The outer boundaries of a product market are determined by the reasonable interchangeability of use or the cross-elasticity of demand between the product itself and substitutes for it.”). “Over the last decade, most courts of appeals have indicated that market delineation should focus on cross elasticity of demand, generally citing *Brown Shoe*.” Werden, *supra* note 13, at 400; *see also* ANTITRUST LAW DEVELOPMENTS, *supra* note 1, at 556–66 (discussing cases on reasonable interchangeability of use and cross-elasticity of demand); Malcolm B. Coate & Mark D. Williams, *A Critical Commentary on the Critical Comments on Critical Loss*, 53 ANTITRUST BULL. 987, 1019 (2008) (“The case law is clear that market definition turns on the closeness of substitutes, including cross-elasticities of supply and demand for nearby products.”); Pitofsky, *supra* note 2, at 1817 (“Virtually all courts started with the *Cellophane* cross-elasticity of demand test . . .”). Ironically, in *Eastman Kodak Co. v. Image Technical Services, Inc.*, 504 U.S. 451 (1992), the Court in text endorsed *Cellophane*’s focus on cross-elasticity, but then cited the fourth edition of the (then) Areeda and Kaplow text, followed by an elaboration in the footnote: “What constrains the defendant’s ability to raise prices in the service market is ‘the elasticity of demand faced by the defendant — the degree to which its sales fall . . . as its price rises.’ Areeda & Kaplow ¶ 342(c), p. 576.” *Id.* at 469 n.15 (alteration in original) (emphasis added).

sumers substitute between products in each market as the price of one product rises.

Yet as some commentators have pointed out — with varying degrees of analysis and force but unfortunately with limited success in changing practice — this emphasis on cross-elasticities is misconceived.⁸³ Assuming that the purpose of the inquiry, even if nominally directed at market definition, is ultimately to enable valid inferences about market power, the analysis in Part II indicates that what matters is the market elasticity of demand, not the cross-elasticity between a given product and one or more particular substitutes.⁸⁴ (Review formula (4).) The simple reason is that it is the market elasticity of demand which indicates the rate at which sales are lost as price rises and thus the extent to which a firm will find price elevation to be profit-maximizing.

Section A states explicitly the relationship between the elasticity of demand and cross-elasticities with other products. Section B discusses implications for market power inquiries, emphasizing some common shortcomings of the market definition process that are in addition to those analyzed in Parts II–IV. In particular, since the emphasis on elasticity is usually correct — and using the market elasticity of demand to make market power inferences does not entail defining markets — whereas the standard focus on specific cross-elasticities is incorrect — even though such is required to define markets — there

⁸³ The development presented in this Part does not appear elsewhere. One of the first discussions of the issue appears in Landes & Posner, *supra* note 13, at 961 n.43, in which they indicate a preference for elasticities without expressing hostility to the use of cross-elasticities. Posner's book takes an even more benign view of standard practice. See POSNER, *supra* note 78, at 149–50. For stronger statements, see AREEDA, KAPLOW & EDLIN, *supra* note 14, at 495; Kaplow, *supra* note 22, at 1829 n.52; Joseph J. Simons & Michael A. Williams, *The Renaissance of Market Definition*, 38 ANTITRUST BULL. 799, 826–27 (1993); Schmalensee, *supra* note 14, at 1803; and Werden, *supra* note 13, at 401. Some courts and commentators affirmatively state the relationship backwards, suggesting that the reason to ascertain elasticities is to help assess cross-elasticities (which, as the text to follow explains, is not that surprising given the focus on market definition and thus the apparent need to identify which products are important substitutes). See *New York v. Kraft Gen. Foods, Inc.*, 926 F. Supp. 321, 333 (S.D.N.Y. 1995) (“Cross-price elasticity is a more useful tool than own-price elasticity in defining a relevant antitrust market. Cross-price elasticity estimates tell one where the lost sales will go when the price is raised, while own-price elasticity estimates simply tell one that a price increase would cause a decline in volume.”); MOTTA, *supra* note 20, at 107 (“When own-price elasticity for the product considered, say product *A*, is high enough to lead us to believe that a hypothetical monopolist would not profitably raise prices of *A* in a small but significant way, it becomes important to identify which products exercise a constraint on *A*. Cross-price elasticities might help us to rank the closest substitute (which, together with *A*, will become the object of the next step of the hypothetical monopolist test.”); Michael L. Glassman, *Market Definition as a Practical Matter*, 49 ANTITRUST L.J. 1155, 1163–64 (1980).

⁸⁴ The market elasticity of demand is sometimes called the price elasticity of demand, and the cross-elasticity is sometimes referred to as the cross-price elasticity. In this Part, the use of the term elasticity by itself refers to the market elasticity of demand, and the simpler term cross-elasticity will be employed.

exists yet another reason that the market definition / market share paradigm should be rejected. Section B also notes that empirical evidence on elasticity tends to be more available and reliable than that on cross-elasticities, which makes it all the more puzzling that one would insist on examining the latter while neglecting the former. Section C asks whether cross-elasticities are ever useful. In short, the answer is affirmative, notably, in examining mergers between producers of differentiated products, but the proper use is not to define markets. Section D offers summary reflections.

A. Relationship Between Elasticity and Cross-Elasticity

To ground further discussion, it is useful to state precisely the relationship between elasticity and cross-elasticity. In subsection II.B.1, expression (3), the market elasticity of demand, ε_d , is first introduced. As explained, it denotes the percentage reduction in the quantity demanded of our good in the homogeneous goods market caused by a one percent rise in the price of that good. As individuals adjust their spending on this good — which it will now be convenient to designate as good 1 — they must necessarily make offsetting adjustments in their spending on other goods, taking their income as given. Suppose that there are N goods in total. The formula stating this relationship is⁸⁵:

$$(8) \quad |\varepsilon_d| = 1 + \sum_{i=2}^N \frac{R_i}{R_1} \varepsilon_{i1}.$$

⁸⁵ For a derivation, see, for example, Kaplow & Shapiro, *supra* note 12, at 1091–92. (In the limited other writing on competition policy that mentions this decomposition of the elasticity into a weighted sum of cross-elasticities, different formulas sometimes appear. See AREEDA, HOVENKAMP & SOLOW, *supra* note 2, at 130 n.4; Landes & Posner, *supra* note 13, at 961 n.43. For explanatory comments, see Werden, *supra* note 13, at 413 n.172.) The method of derivation tracks the preceding explanation in the text: one differentiates a representative individual's budget constraint (indicating that the sum of the quantity purchased of each good, weighted by that good's price, equals total income, which is taken to be fixed) with respect to the price of good 1, holding other prices fixed, and then rearranges terms and makes substitutions using definitions of other terms to produce expression (8). Because the prices of all other goods are held fixed, this formula embodies partial equilibrium analysis. See *supra* note 38. For goods that are perfectly elastically supplied (have constant marginal cost), this simplification is unproblematic. However, if the supply curve for some other good is steeply rising — and, moreover, if we are considering a substitute for which the cross-elasticity is rather high and the relative market size is low — then a full general equilibrium analysis would reflect a significant rate of price increase for the substitute, which would to that extent dampen the magnitude of the substitution, implying a greater degree to which price could profitably be elevated in the original market. Such effects are ordinarily set to the side and will not be considered further here, although one can conceive of situations in which this factor would be important. (If one focuses on elasticity rather than cross-elasticity, as argued here, the phenomenon arises implicitly. Note that if one is measuring the elasticity by actual market behavior, such feedbacks might be captured, although they would not be if the analysis held constant — controlled for — prices of such substitutes.)

The term ε_{i1} on the right side is the cross-elasticity of demand: specifically, the percentage change in the demand for good i as the price of good 1 increases by one percent.⁸⁶ R_i indicates the revenue expended on good i , which means that R_i/R_1 is the ratio of revenue spent on good i to that spent on good 1 — hereinafter referred to as a revenue share or revenue weight.

To elaborate the intuition behind formula (8), consider first the case in which all of the cross-elasticities equal zero; that is, as the price of our good (good 1) rises, individuals do not modify at all the quantity that they purchase of any other good. Expression (8) informs us that, in this case, $|\varepsilon_d| = 1$. This result (rather than $|\varepsilon_d| = 0$) is correct because, if the price of good 1 rises and purchases of all other goods stay the same — and we have assumed that disposable income is constant — it must be that the quantity purchased of good 1 falls. And, in particular, for a small percentage increase in the price, the quantity falls by the same percentage, which explains the unitary elasticity.⁸⁷ Accordingly, the extent to which the magnitude of the elasticity exceeds 1 reflects the degree to which consumers transfer expenditures from good 1 to all other goods.⁸⁸

⁸⁶ It is conventional to use ε_{i1} on the left side (thus referring to the percentage change in the demand for good 1 as the price of good 1, its own price, increases), but ε_d is used to maintain consistency of notation with that in Part II.

⁸⁷ Some readers may be familiar with the fact that, say, a 1% rise in x and a 1% fall in y does not leave the product xy unchanged (instead it equals $(1.01x)(0.99y) = 0.9999xy$). Since the formal definition of the elasticity involves the derivative, however, the elasticity is the limiting value of this factor as the magnitude of the change approaches zero, so formally the elasticity is 1.0.

⁸⁸ To explore this way of stating the idea further, one can subtract 1 from each side:

$$|\varepsilon_d| - 1 = \sum_{i=2}^N \frac{R_i}{R_1} \varepsilon_{i1}.$$

The left side thereby becomes the percent change in the quantity of good 1 relative to (net of) a baseline in which the amount of revenue spent on the good stays at its preexisting level. Then one can further multiply both sides by R_1 , which yields:

$$R_1(|\varepsilon_d| - 1) = \sum_{i=2}^N R_i \varepsilon_{i1}.$$

To make this point concrete, suppose now that some good, say good 2, is a substitute, so that expenditures on it rise as the price of good 1 increases. If nothing else changes, we will obtain a higher value for $|\varepsilon_d|$, in particular a value in excess of 1. In the formula, term 2 in the summation is now positive. Additional positive terms (more substitutes) would, all else equal, further raise the elasticity.⁸⁹ Likewise if the magnitude of such positive terms is larger. The latter is often referred to as involving closer substitutes, but note that each cross-elasticity term is weighted by the revenue ratio, the good's importance in overall expenditures relative to that of good 1. Because this latter point is widely ignored yet proves to be quite important, the underlying intuition will be explored in depth in subsection B.2.

B. Implications

In examining expression (8), it is immediately obvious that focusing on particular cross-elasticities rather than on elasticity is a mistake. Recall first, from formula (4), that it is the market elasticity of demand (in the homogeneous goods market) that has immediate market power implications. Taking the simplified case in which rivals' supply elasticity is zero, the Lerner index is simply given by $S/|\varepsilon_d|$. No cross-elasticities appear directly in this formula.

How, then, are cross-elasticities even relevant? As expression (8) indicates, the elasticity — about which we care directly — equals (one plus) the weighted sum of the cross-elasticities. However, the standard use of cross-elasticities (and related notions) in defining markets does not properly reflect this basic relationship.

1. *Summation.* — First, it is plain that this summation (from 2 to N) is over all other products. The elasticity, as explained in section A, captures this aggregate of all demand response channels. A single cross-elasticity, by definition, addresses only one. And a handful allows an assessment of only a few. Even if one accurately identifies that substitute or those two or three that have the largest cross-elasticities, their combined effect is less⁹⁰ and often will be substantially (possibly overwhelmingly) less than the total effect. Therefore, to use cross-elasticities, one needs to determine all the cross-elasticities, and then take the weighted sum.⁹¹ But why do that rather than simp-

At this point, the left side indicates the change in revenue spent on good 1, which must be equal to the change in revenue spent on all the other goods, which in turn is given by the sum of the cross-elasticities, each weighted by the amount of revenue spent on the respective good.

⁸⁹ By similar reasoning, negative terms (complements) reduce the elasticity.

⁹⁰ But see *infra* note 94.

⁹¹ One might think that omitting small values would involve little error. This would be true if, for example, there were only two other goods in existence, the cross-elasticity on one was large and the other tiny, and there were not offsetting expenditure shares (further discussed in subsection 2). However, in a developed economy in which products number in the thousands (or vastly

ly looking at the elasticity in the first place? After all, “from the firm’s viewpoint, it does not matter why customers turn away or where they turn but only the rate at which sales decrease when that firm raises its price.”⁹² For example, “if home cooks react to higher prices for electric mixers by stirring manually and spend their mixer money on holidays, movies, shirts, or gin, no market definition can reflect those choices.”⁹³

In some instances, it is possible that looking at one or a few cross-elasticities might be sufficient: if they were very high (and if their revenue weights were also high, on which more in subsection 2), then we could use those figures to form a lower bound on the elasticity. That is, even if all the other terms in the summation were zero, we would know that the elasticity was at least one plus the weighted sum of those cross-elasticities we measured.⁹⁴ Then, if it were also true that the resulting elasticity was sufficiently high to demonstrate that market power was too low to be of concern, there would be no need to investigate further. Thus, if direct evidence on the market elasticity were hard to find and good evidence on key cross-elasticities were available (an atypical occurrence, as subsection 3 notes), looking at those cross-elasticities would make sense. Note, however, that we would be looking at them not to define a market, but to obtain a (lower bound) estimate of the market elasticity of demand, from which a direct market power inference would then be made.⁹⁵

In other instances, looking only at close substitutes can be highly misleading. Suppose that in one case there is a single, excellent substitute (all other cross-elasticities are zero). In another, there are twelve notably poorer substitutes, with cross-elasticities only a third as high. If all the revenue shares were equal, then the twelve poor substitutes would have four times the impact of the one excellent substitute.⁹⁶ Yet the normal practice of confining attention to close substitutes and counting them in full, while ignoring more distant substitutes, no mat-

more, depending on how one groups them), looking at the few with the largest cross-elasticities and ignoring myriads of others, even if they each individually have small cross-elasticities, could lead to substantial underestimates of the elasticity.

⁹² AREEDA, KAPLOW & EDLIN, *supra* note 14, at 488–89.

⁹³ *Id.* at 489 n.17.

⁹⁴ If some other goods were complements, for which the cross-elasticity is negative, then it is possible that the sum of omitted terms would be negative, in which case the method in the text would not provide a lower bound. As a practical matter, the technique would likely be problematic only if strong complements were omitted, which omission should therefore be avoided.

⁹⁵ Note further that we really cannot tell how tight this lower bound is (whether the true elasticity is only slightly above or far higher) unless we know, in aggregate, what we have omitted. The magnitude of the omission could be estimated if we knew the extent to which consumers switched to all other goods, but in that case we would already know the elasticity.

⁹⁶ See Kaplow, *supra* note 22, at 1831 n.56; AREEDA, KAPLOW & EDLIN, *supra* note 14, at 495; Werden, *supra* note 13, at 402; AREEDA, HOVENKAMP & SOLOW, *supra* note 2, ¶ 507a, at 131.

ter how numerous, leads to a strong conclusion in the opposite direction. There is, in general, no way to know the importance of all these other substitutes without actually looking. One way of doing so is to enumerate dozens of substitutes, or even every good in the economy, measure the cross-elasticities (and revenue weights), and apply expression (8). The other method, of course, is just to examine the market elasticity of demand — the thing we actually care about — in the first place and often the last because, once that is done, we do not need to know particular cross-elasticities for most purposes (on which more in section C).

2. *Revenue Weights.* — There is an additional problem with the manner in which cross-elasticities are ordinarily taken into account; namely, courts and commentators perform their analysis about whether to combine markets as if the revenue share weights in expression (8) did not exist.⁹⁷ To begin, consider further the intuition behind why revenue weights appear in this formula and what their significance is. Suppose, once again, that there is only one good with a non-zero cross-elasticity, good 2. Furthermore, assume that this cross-elasticity, ε_{21} , equals 1 and that the amounts consumers spend on goods 1 and 2 are equal (that is $R_2 = R_1$, which in turn implies that $R_2/R_1 = 1$). In that case, the market elasticity of demand, $|\varepsilon_d|$, will equal 2. But what if $R_2 = 2R_1$, so that $R_2/R_1 = 2$? Then $|\varepsilon_d| = 3$. Why is there such a large difference in market power even though in both cases we have substitutes with the same cross-elasticities? If the revenue share of good 2 is twice as high, then a given *percentage* increase in the number of units of good 2 that are purchased (which is what the cross-elasticity measures) will correspond to twice the *absolute quantity* by which demand for good 2 increases as the price of good 1 rises. Hence, the impact of substitution to good 2 — put another way, the number of purchases of good 1 that switch to good 2 as the price of good 1 rises — will be twice as great.

Accordingly, what may seem to be a mere technical refinement is actually quite important. Two goods with equal cross-elasticities can have very different implications for the overall elasticity. Similarly, a single good with only a moderate cross-elasticity can have a far greater impact than a good with a much larger cross-elasticity — or even a group of goods all with larger cross-elasticities. Specifically, this result will arise whenever the revenue share of that single good significantly exceeds that of the good or goods with higher cross-elasticities.

⁹⁷ For an exception, see Kaplow & Shapiro, *supra* note 12, at 1092. The U.S. Merger Guidelines might also be viewed as an exception, although their approach to market definition addresses a different question from what is conventional and from what is most relevant to assessing the competitive effects of mergers. See *infra* note 99.

As a general matter, the relative revenue shares of goods vary dramatically. After all, it is obvious that some goods in the economy have total expenditures that are orders of magnitude different from those on other goods. Therefore, failing to adjust one's analysis of cross-elasticities to account for revenue shares — in particular, to multiply them, as expression (8) shows to be required — can lead one badly astray. Nevertheless, these sorts of differences — taking into account revenue shares, as well as the need to aggregate groups of substitutes (that may collectively have a high revenue share, even if each has only a modest cross-elasticity) — are routinely ignored when defining markets, as the familiar injunction to focus on cross-elasticities omits this consideration.

Even though the examination of cross-elasticities in deciding whether to redefine a market gives no account of revenue weights, the practice of market redefinition does implicitly give them weight in those instances in which impressions of cross-elasticities are sufficiently high that markets are ultimately combined. The reason is that, once the markets are aggregated, one then computes the firm's share (which equals the firm's sales divided by total sales in the newly defined market), and when one does so, the denominator, being the sum of sales in both markets, will necessarily reflect the market size of the market that was added in. In this respect, when substitutes involve greater revenue, they will be given more weight — that is, if and only if one does decide to combine markets. It is, however, important to keep in mind the familiar point that, when markets are thus combined, imperfect substitutes are treated as if they were perfect substitutes, with the result that market power tends to be overstated — and this problem will be worse the larger the market that is added — unless one adjusts inferences in a manner that, in essence, fully undoes the market redefinition (as discussed in subsection II.B.2).⁹⁸ A version of this problem arises with the hypothetical monopolist test employed in the U.S. Merger Guidelines and in other jurisdictions, as elaborated and illustrated in some detail in the margin.⁹⁹

⁹⁸ Suppose, for example, that we compare two cases: in one, the cross-elasticity is 2.5, moderately over the minimum level of 2.0 we might demand for inclusion, and in another, the cross-elasticity is 10.0. If the revenue shares in each case were the same, the actual impact of the substitute on market power would be only a fourth as high in the former case, even though both are treated the same in conventional analyses.

⁹⁹ See U.S. MERGER GUIDELINES, *supra* note 5, § 4.1.1; *DG Competition Article 82*, *supra* note 1, ¶ 14 (SSNIP test used by the Commission for market definition, particularly in merger cases). It might be that, in a narrow market, the hypothetical monopolist would not quite muster a 5% price increase (SSNIP), so the market is broadened to include a substitutes market that, although being closest, has only a modest cross-elasticity but whose size is very large relative to that of the original market. Levels of and increases in the Herfindahl-Hirschman Index (HHI) in the broader market could readily be very low even if the merger significantly enhances market power. "Significantly" is used in a relative sense: after all, we know that the price elevation must be less

To round out one's understanding of this topic, it is helpful to contemplate the relationship between the magnitude of the cross-elasticity for a particular substitute and that good's revenue weight. Consider the following illustration, a variant on the facts in the *Cellophane* case. Suppose that the only flexible wrapping material other than cellophane is aluminum foil and that one had computed the cross-elasticity and revenue weight to determine its contribution to the elasticity of demand for cellophane. Suppose further that, the next day, a new product, the oven, is invented and that, accordingly, the demand for foil doubles, whereas that for cellophane is unchanged because it is useless in this domain (it melts). Since the revenue weight for foil is now twice as high, does this mean that the elasticity of demand for cellophane is accordingly higher? Not at all, because in this hypothetical scenario, the relevant cross-elasticity will be halved: a one percent in-

than 5%; when the *hypothetical monopolist* can raise prices at least that much, say 7%, what would be viewed as significant HHIs and increases in the HHIs may readily translate into price increases that are much less than 5%.

To make this point concrete, consider the following example. Suppose that the elasticity of demand is as high as 10, in which case a firm with a 100% market share (hypothetical monopolist) would have a Lerner index of 0.10 (simplifying throughout by taking rivals' supply elasticity to be zero). (As explained in section II.A, this translates into a price elevation of just over 11%, but it will suffice for present purposes to use rounded approximations, in this case 10%.) If, prior to the merger, we had Cournot competition and the premerger HHI was 2000, then the premerger margin is 0.2/10 or 2% (using the formula appearing, for example, in Ordover, Sykes & Willig, *supra* note 25, at 1865; and Kaplow & Shapiro, *supra* note 12, at 1085, under which the industry-wide average, output-weighted margin equals the HHI of market concentration divided by the elasticity of market demand (and keeping in mind that HHIs are typically measured on a scale of 0 to 10,000 rather than 0 to 1, so conventional HHIs, in terms of the relevant formula, indicate tenths rather than thousandths)). Therefore, a hypothetical monopolist would raise its price about 8% (the difference between the monopoly elevation of approximately 10% and the premerger elevation of 2%). This passes the Merger Guidelines' SSNIP test. Suppose that the actual merger raises the HHI from 2000 to 2600, which would be regarded as a large increase from an already substantial level. See U.S. MERGER GUIDELINES, *supra* note 5, § 5.3 ("Mergers resulting in highly concentrated markets [with an HHI above 2500] that involve an increase in the HHI of more than 200 points will be presumed to be likely to enhance market power."). But how much will the price increase? Using the formula, the postmerger HHI corresponds to a postmerger margin of 0.26/10, or 2.6%. (This mode of postmerger analysis, although common, is incomplete because it ignores that the share of the merged firms will not in general equal the sum of their premerger shares and other firms' shares will change as well; for a rigorous analysis, see Joseph Farrell & Carl Shapiro, *Horizontal Mergers: An Equilibrium Analysis*, 80 AM. ECON. REV. 107 (1990).) Therefore, the merger that is solidly in the danger range raises price by only about 0.6% (just over half a percentage point). (We are not assuming the presence of merger efficiencies or other factors that under the Merger Guidelines might modify the bottom line.) By contrast, if a merger substantially raised the possibility of coordinated behavior (or was a merger to monopoly), producing the hypothetical monopoly result, but the SSNIP test in the homogeneous goods market suggested an elevation of only 4% (nearly seven times as much), the market would be expanded, and if the "next" market had a high level of revenue relative to that in the original, narrow market, the merger may well pass under the radar screen. This juxtaposition shows how, to a significant degree, the Merger Guidelines provide determinacy at the expense of coherence. The problem is that available information is not used directly to ascertain how much the merger will increase price.

crease in the price of cellophane will have the same quantitative impact on the demand for foil as before (let us suppose), but since the total amount of foil sold has doubled, as a percentage of foil sales the effect of the demand increase will be only half as great. Hence, the cross-elasticity (which, recall, is defined in percentage terms) is only half as high. Combined with the fact that the revenue weight is twice as high, we can see that the product of the two is unaltered. This result is in accord with intuition since, by assumption, nothing has changed about the importance of foil as a substitute for cellophane.

However, this example does suggest an important dilemma for the market definition / market share paradigm: Suppose that previously one thought that the cross-elasticity was high enough to justify combining markets. If one still thinks so, the impact of the redefinition on the market share of the cellophane firm will now be twice as great even though the importance of the substitute has not changed one iota. The resulting inference of much less market power would be a mistake. But if one decides, on account of the reduced (halved) cross-elasticity, to change one's mind and exclude foil, then (under conventional analysis) one is entirely ignoring the effect of foil on market power in the cellophane market even though the effect is unchanged and one had previously concluded that it was important. Accordingly, this example further shows how redefining markets to produce different market shares is a very blunt instrument for reflecting the impact of substitutes on market power.¹⁰⁰ By contrast, these different cases are treated identically — and correctly — if one confines attention to the elasticity of demand (which has not changed) and forgoes the market definition process altogether.

In summary, fairly simple economic analysis yields expression (8), which instructs us to look (for each other good) at a product of two numbers (cross-elasticity and revenue shares) — call it $A \times B$ — whereas the standard rule and practice considers only the magnitude of A , even though it is obvious that the value of B varies tremendously from one context to another (and even across potential substitutes in a single case¹⁰¹). Combined with the first argument in this section, that

¹⁰⁰ Note that the problem has nothing to do with the change in the revenue weight occurring the next day; if ovens had been present from the outset, the dilemma would be the same. Indeed, this dilemma is simply one of the standard problems with using market definition: in general, no market definition captures the strength of substitution, a point that has already been made in various ways in prior Parts of this Article.

¹⁰¹ Indeed, it might seem natural to consider as the closest substitute that with the highest cross-elasticity. However, the most important substitute — the one responsible for most restraining market power — may well be another good with a lower (possibly much lower) cross-elasticity but a sufficiently greater revenue share to result in a product of the two numbers that is higher than that for the substitute with the highest cross-elasticity. See generally Werden, *supra* note 13, at 402–06 (discussing cross-elasticities and different ways to rank substitutes).

one needs to aggregate all of the weighted cross-elasticities rather than focus on one or a few, it is clear that conventional analysis that focuses on particular cross-elasticities rather than on the overall elasticity will often be badly off the mark. Of course, a central motivation for doing the former is that it seems necessary if one must define markets, a process of inferring market power that we can now see has yet another striking deficiency.

3. *Evidence Availability.* — Subsections 1 and 2 explain why focusing on one or a few cross-elasticities is a mistake, one engendered by the quest for the so-called relevant market. We should instead inquire into the market elasticity of demand, which captures all channels of substitution and weights them in the appropriate fashion — and which does not require us to define markets. A further irony is that, however difficult such estimation might be, the elasticity is ordinarily more easily and more reliably estimated than cross-elasticities because the latter is more sensitive to how well changes can be related between two markets rather than within one.¹⁰²

The comparative ease of measuring elasticity versus cross-elasticity may well characterize some impressionistic evidence as well. Firms' internal evidence (however obtained) used to formulate pricing policy, just like formula (4) — which, recall, is based on firms' profit-maximizing decisions — will need to reflect the elasticity, not some particular cross-elasticity.¹⁰³ Likewise, if consumers are surveyed, they may be able to give a better read on whether and how much they would reduce purchases than on the details of how they would allocate their resulting shift among all other goods.¹⁰⁴

¹⁰² See, e.g., AREEDA, KAPLOW & EDLIN, *supra* note 14, at 495 n.29; Simons & Williams, *supra* note 83, at 827; Werden, *supra* note 13, at 402 n.152; see also Jerry A. Hausman & Gregory K. Leonard, *Competitive Analysis Using a Flexible Demand Specification*, 1 J. COMPETITION L. & ECON. 279 (2005) (surveying methods of estimating elasticities and cross-elasticities, often noting greater obstacles regarding the latter); Kaplow, *supra* note 22, at 1830 n.55 (“But empirical estimates of elasticity are usually available only for conventionally defined product markets. Because elasticity already incorporates substitution (cross-elasticity) . . . , empirical estimates of demand elasticity could be used directly when product markets are not combined. When product markets are combined, however, no immediately usable empirical data are generally available.”); Werden, *supra* note 13, at 363 (“Demand elasticities are actually being estimated, and the estimated demand elasticities are being used to delineate markets, to measure market power directly, and to predict the competitive effects of mergers.”). In fact, it sometimes may be best to estimate the *firm's* elasticity of demand — which, as formula (2) indicates, may enable an even more direct measure of market power than that provided through estimation of market elasticities, much less numerous cross-elasticities. See, e.g., Baker & Bresnahan, *Residual Demand*, *supra* note 43.

¹⁰³ This point is not always true. An advertising campaign aimed at a particular substitute would be based on evidence of perceived cross-elasticity (and the revenue weight).

¹⁰⁴ Although not as logically connected to the choice between the two measures, it is notable that many inquiries are conceptually disconnected from the relevant question. Both elasticity and cross-elasticity are concerned with a (percentage) *rate of change* in purchasing behavior in response to *changes* in prices. Observing (a static picture of) what consumers are purchasing at

This Article does not focus on empirics, and in various settings one or another source of evidence may prove more reliable. Nevertheless, it seems likely that elasticity will usually be easier to measure (or guesstimate) than is even a single cross-elasticity, much less many of them. Since the elasticity is also what we wish to know, whereas a cross-elasticity is only a component of the elasticity, the case for focusing on elasticities rather than cross-elasticities is even stronger.

C. Are Cross-Elasticities Ever Useful?

Subsection B.1 indicates that, if there were no direct way to assess the elasticity of demand and if, moreover, we had good evidence on some important substitutes (those for which the product of the cross-elasticity and the revenue weight is high), we could use the latter to form a lower-bound estimate of the former. Subsection B.3 suggests that this scenario will not arise frequently. This section considers whether there are other uses of cross-elasticity information.

The answer is affirmative. In particular, there is an important use for cross-elasticities when evaluating the prospect of unilateral¹⁰⁵ price increases in horizontal mergers in differentiated product industries. In such settings, the question is not what the level of market power is, either before or after the merger, but how much it will change (this increment, of course, can be calculated by subtracting the before-level from the after-level). Often, the best way to estimate this change (or to predict the level of market power after the merger) is to determine the

given prices does not address this question. It may in our imaginations: we might assume that the reason so many are purchasing y rather than x is that the price of x is elevated. But the quantities purchased of each good do not in themselves answer that question and certainly do not indicate the *rate* at which consumers' purchases would shift if price *changes* slightly, which as just stated is central to measuring both elasticities and cross-elasticities. (This point is distinct from the so-called *Cellophane* fallacy, which involves asking whether a firm, whose status as a monopolist we seek to ascertain, could profitably raise prices even more than it already has — which, if it is a profit-maximizer, it certainly cannot, for if it could have it already would have — rather than whether the firm's prices already entail significant elevation above a competitive level.) As an illustration of how market power analysis can go awry, see Capps, Dranove, Greenstein & Satterthwaite, *supra* note 43, who criticize the Elzinga-Hogarty geographic market definition test, which is based on shipment flows, and find that, in a hospital merger setting, it can lead to broad market definitions that imply negligible market power whereas more direct approaches demonstrate that significant price increases may well result.

¹⁰⁵ That is, the discussion focuses on how much the merging parties would find it in their interest to raise prices, taking as given their competitors' behavior and thus ignoring how the merger may affect the ability to elevate price in a coordinated manner. Likewise, the discussion presupposes a lack of prior coordination (for if, say, there were already complete coordination, then both firms would already be charging essentially the monopoly price, so the merger would not result in a further price increase).

extent to which the merger relaxes the pricing constraints on the merging firms.¹⁰⁶

Examining this problem has become the focus of a significant body of literature, much of it under the rubric of “critical loss.”¹⁰⁷ Abstracting from the details of particular arguments in that work as well as from how it is commonly used, it suffices to focus on a single, key point. Supposing that we wish to know how much, say, the producer of good 1 will want to raise the price of its product after merging with the producer of good 2, of central importance will be the cross-elasticity of demand between these two products (as well as the relative revenue weight, as previously explained). In a sense, the merger reduces the “effective” elasticity of demand for each good by an extent indicated by the cross-elasticity (and revenue weight) because the firm is not bothered by the sales it loses to itself (assuming, for simplicity in this discussion, that the price-cost margins for the two products are the same¹⁰⁸). Viewed another way, we might imagine that the merged firm would raise the price of both products. Before the merger, sales lost to each other mattered — and thus price elevation was moderated on that account — whereas after the merger, these losses are internal and thus offset by internal gains.

Observe that this important use of cross-elasticities does not involve market definition.¹⁰⁹ That is, determination of the extent to

¹⁰⁶ As will be mentioned in section VI.C, analogous considerations are applicable if one wishes to determine, for example, the extent to which an alleged act of monopolization directed at producers of a substitute would, if successful, augment the monopolist’s market power. See Kaplow & Shapiro, *supra* note 12, at 1185–86.

¹⁰⁷ 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); Michael G. Baumann & Paul E. Godek, *A New Look at Critical Elasticity*, 51 ANTITRUST BULL. 325 (2006); Coate & Williams, *supra* note 82; Farrell & Shapiro, *supra* note 9; Barry C. Harris & Joseph J. Simons, *Focusing Market Definition: How Much Substitution Is Necessary?*, 12 RES. L. & ECON. 207 (1989); Michael L. Katz & Carl Shapiro, *Critical Loss: Let’s Tell the Whole Story*, ANTITRUST, Spring 2003, at 49; Daniel P. O’Brien & Abraham L. Wickelgren, *A Critical Analysis of Critical Loss Analysis*, 71 ANTITRUST L.J. 161 (2003); see also Kaplow & Shapiro, *supra* note 12, at 1085–86, 1173–75 (discussing the diversion ratio and commenting on the critical loss literature).

¹⁰⁸ If the margins differ, then the merged firm is not indifferent to sales shifts. Sales moving from the high-margin good to the low-margin good reduce profits (thus the incentive to raise price of the initially high-margin good is lower than the discussion in the text suggests) but sales moving in the opposite direction increase profits (thus raising the incentive). Note also that, as one changes the price of good 1 while holding the price of good 2 constant, the relative margins will change.

¹⁰⁹ This point is a central theme of Farrell and Shapiro, *supra* note 9. See also Werden, *supra* note 13, at 407 n.164 (“The central role of demand elasticities in merger simulation may be demonstrated most powerfully by the observation that the use of merger simulation eliminates the need for market delineation because the predictions of merger simulations are not sensitive to the make-up of the product group used in the simulations.”). Nevertheless, much of the development and use of the critical loss methodology and related inquiries is for purposes of defining markets, often under the U.S. Merger Guidelines. See, e.g., Malcolm B. Coate & Jeffrey H. Fischer, A

which such a merger induces the merging parties unilaterally to elevate the prices of their differentiated products does not require determining whether their two goods are deemed to be in the same market or which other goods are in or out of the market. The cross-elasticities between their products (and also the revenue weights and the pre-merger margins on each) are relevant to their incentive to elevate price, but these measures are not used, as in ordinary market definition inquiries, to determine whether the two products are in the same market. (In many such merger cases, it may be conceded that the two products indeed are in the same market; market definition disputes would involve what other substitutes are also in the same market, another question that does not need to be addressed using the approach under consideration.)

D. Reflections

Having explained the relationship between elasticity and cross-elasticity in section A and analyzed it in subsequent sections, let us now step back and view the whole from the perspective of the various formulas. Expression (1) defines market power as the Lerner index or price-cost margin. One might try to measure those terms (price and marginal cost) directly. But this process may be too difficult. Expression (2) indicates that market power will equal one over the firm's elasticity of demand. One could try to measure that directly (over recent decades, techniques have been developed to do so¹¹⁰). Sometimes that will be too difficult as well. Expressions (3) and (4) show, respectively, how this firm elasticity and the Lerner index are related to market share, the market elasticity of demand, and rivals' supply elasticity, all in the homogeneous goods market. One could attempt to measure these directly or make inferences about their magnitudes in various ways, and then use these expressions to assess market power. Suppose that this method as well is imperfect and costly.

The market definition / market share paradigm is yet two steps further removed. Rather than measuring the market elasticity of demand in expression (3) or (4) directly, it focuses on cross-elasticities of one or a few substitutes. Expression (8) indicates the relationship between elasticity and cross-elasticity. The analysis in this Part makes it quite clear that, by focusing on only a limited number of cross-elasticities (sometimes only one) and by failing to employ revenue weights, this method is not only indirect but erroneous, probably quite significantly so in many cases. Finally, even when one has in hand a best estimate

Practical Guide to the Hypothetical Monopolist Test for Market Definition, 4 J. COMPETITION L. & ECON. 1031, 1040–44 (2008) (section entitled “Critical Loss Analysis for Market Definition”).

¹¹⁰ See, e.g., Baker & Bresnahan, *Residual Demand*, *supra* note 43.

of one or more cross-elasticities, the market redefinition approach does not sum them (properly weighted or not), as demanded by formula (8), in order to obtain an estimate of the market elasticity of demand — which then in turn, according to formula (4), would facilitate a direct inference about market power. Instead, the approach goes through substitutes one by one and makes an on-off (all-or-nothing) decision about whether the cross-elasticity is sufficiently high. If not, the substitute is ignored (contrary to formula (8)). If so, the substitute is treated as if it were perfectly homogeneous (which it is not, and thus also contrary to formula (8)), and then market power inferences are made through some sort of inference process that is not ordinarily specified (but has been the subject of much of the prior analysis in this Article).

Reflection on the step in which it is decided whether to include or exclude a substitute in defining the relevant market suggests that the analysis in Part IV gave the market redefinition technique an important benefit of the doubt. Specifically, in section IV.A, it was imagined that one chooses between two competing market definitions (a narrower one that is at least somewhat too narrow and a broader one that is at least somewhat too broad) in a way that generates the best outcome (where best is alternatively defined as most accurate in measuring market power or as yielding the right legal answer). This decision procedure might be what an analyst, enforcer, or court means when it asks whether two substitutes are good enough that the markets should be combined, that is, whether the cross-elasticity (degree of interchangeability) is sufficiently high. Observe, however, that it is entirely possible that in one case the cross-elasticity might be 5.7, yet failing to combine the markets is best, whereas in another the cross-elasticity is only 1.2, yet combining the markets is best. If the criterion is the best market definition, then it does not matter per se what the cross-elasticity happens to be, and the many considerations adduced here make it clear that there is neither a particular cross-elasticity nor any narrow range of cross-elasticities that provides the basis for some reasonably uniform threshold test.

By contrast, one might suppose that most who have addressed market definition and focused on cross-elasticities (and cognate concepts) imagine that they are applying some fairly consistent rule. Indeed, that seems the most natural interpretation of the sort of language ordinarily employed. Interestingly, neither analysts nor decision-makers have been forthright in stating what the magic number (or typical number, or standard range) happens to be. To push this point, suppose that expert reports from both sides in a case state that the

cross-elasticity in question is 2.46. (Or 1.34. Or 3.07.) Who wins?¹¹¹ (Compare the two hypothetical survey questions relating market share and market power in section III.B's discussion of the concept of a standard reference market.) Of course, even if there were an answer, we know it would be wrong, very wrong. As just stated, if one is to insist on defining markets, the best definition is determined by all information that bears on market power, and the value of the cross-elasticity between the products in the two markets in question will give us little clue about what that answer is. Therefore, although Part IV indicates that there are coherent senses in which superior answers to the market definition question exist (although answers that render the market definition process superfluous at best), the present discussion reinforces the conclusion of this Part that the standard manner in which markets are defined, which emphasizes the magnitude of particular cross-elasticities, is very far removed from the ultimate target of the inquiry.

In summary, Parts II–IV indicate a number of deep, conceptual flaws with the market definition / market share paradigm. Examining the elasticity of demand and cross-elasticities, and the relationship of the latter to market redefinition, shows further and from another perspective how poor is the relationship between the paradigm and a coherent approach to the assessment of market power. The gap is not primarily practical, depending on what sorts of evidence might be available (although it certainly is that too). It is logical. As emphasized in prior Parts, whatever evidence might be available that bears on market power — and however good or bad it may be — there is a gulf between making the best use of such evidence and the manner in which it is used (abused) when one employs it instead to choose a relevant market and then attempts to make market power inferences from market shares therein.

VI. ADDITIONAL TOPICS

Parts II–V present a multidimensional, conceptual critique of the market definition / market share paradigm. Before concluding, it is

¹¹¹ Even when focusing, more appropriately, on the elasticity rather than on particular cross-elasticities, problems of determining the proper cutoff for purposes of market definition still arise. For example, in their econometric analysis of margarine and butter, David Scheffman and Pablo Spiller find that the elasticity for margarine is -3.5 , which they deem “high . . . indicating that margarine is not a relevant market.” David T. Scheffman & Pablo T. Spiller, *Econometric Market Delineation*, 17 *MANAGERIAL & DECISION ECON.* 165, 175 (1996). Yet they do not explain why this elasticity should be viewed as high rather than low or medium, or why it indicates that the relevant market is broad rather than narrow (just margarine). Indeed, if one applies the hypothetical monopolist test (so the share is 100%) using formula (4), the Lerner index is approximately 0.29 and the implied degree to which price would exceed marginal cost is 40%, which most would take to be quite substantial market power.

useful to consider a number of related issues. Section A examines the applicability of the analysis, which often focuses on substitution between different products, to other sources of supply of the initial good. Section B remarks on the intermingling of assessments of market power and of legal rules for condemning various acts. Section C addresses specific contexts in which market definition is thought to be useful and finds in each that, in light of the preceding analysis, it is not. Section D asks why the market definition process has enjoyed such longstanding appeal in light of its deficiencies. Finally, section E considers briefly the law on market definition, inquiring in particular whether the ideas presented here can be incorporated into existing doctrine that often is seen to require the identification of a relevant market.

A. Rivals' Elasticity of Supply

It is convenient to begin by restating expression (4) from subsection II.B.1 on the determinants of market power in the homogeneous goods market:

$$(4) \quad L = \frac{P - MC}{P} = \frac{1}{|\epsilon_f|} = \frac{S}{|\epsilon_d| + (1 - S)\epsilon_r}.$$

As emphasized in the original assessment of this formula, it makes clear that market power depends not only on market share (in the homogeneous goods market) but also on the market elasticity of demand, ϵ_d , and rivals' elasticity of supply, ϵ_r . Two observations are in order.

First, even though some of the discussion thus far has focused on the elasticity of demand (especially that in Part V on elasticity versus cross-elasticity), the supply elasticity — as mentioned, notably including expansion by existing rivals but also, if interpreted broadly, covering supply substitution and entry — is an additional important determinant of market power. This point is emphasized in Landes and Posner's article and in much of the other literature, and it is featured as well in the U.S. Merger Guidelines' methodology for assessing the impact of horizontal mergers.¹¹² Failure to attend to this factor can result in significant mismeasurement of market power.¹¹³

¹¹² Landes & Posner, *supra* note 13, at 948–50; U.S. MERGER GUIDELINES, *supra* note 5, §§ 5.1, 9; *see also* AREEDA, KAPLOW & EDLIN, *supra* note 14, at 498; Fisher, *supra* note 49, at 15–16; Kaplow, *supra* note 22, at 1847; Schmalensee, *supra* note 14, at 1801. There is some controversy about whether, as a practical matter, it makes more sense to consider supply responses as part of the market definition exercise or in a separate step. *See, e.g.*, Baker, *supra* note 2, at 133–38 (advocating that the steps be separate, as under the U.S. Merger Guidelines). Such pragmatic questions are largely distinct from the focus of this Article (although note 116 indicates a possible connection).

¹¹³ One might suppose that market power would be systematically understated, as argued by Landes and Posner, *supra* note 13, at 950. However, this view assumes that those not addressing

Second, and more relevant for present purposes, is the relationship between this factor and the market definition / market share paradigm. The previous analysis in this Article, for the most part, stands independently of how rivals' supply elasticity is handled. However, it is worth noting that many of the same problems arise when one considers supply elasticity using the standard paradigm. Subsection II.B.2 already explained that the need to essentially undo any market redefinition if formula (4) is to be applied properly — because the formula is valid only for the homogeneous goods market — applies equally to the factor of rivals' elasticity of supply.¹¹⁴

Likewise, the discussion in Part III of the implicit need for a standard reference market pertains to the overall market power inference from a given market share. This inference, in turn, depends (as indicated by expression (4)) on both the elasticity of demand and rivals' supply elasticity. Hence, the analysis in that Part is broadly applicable to both considerations — really, it applies to them as a whole rather than to each separately.¹¹⁵

The same logic is applicable to Part IV. In asking whether one market definition is better than another and when applying criteria for the best market definition, we considered the market power implications from each market definition. Because the implications involved a market power inference from a resulting market share — which implicitly relies on a construct like the standard reference market — that logic also takes account of the aggregate effect of the market demand elasticity and rivals' supply elasticity. Furthermore, any use of market redefinition to capture supply issues is subject to Part IV's core argument that choosing the best market presupposes that one already has a best estimate of market power, all things considered.

supply responses directly must be implicitly employing something like expression (4) and setting ε_i equal to zero. But if they instead employ the market definition / market share paradigm and make inferences from market shares, they may well be implicitly taking ε_i to be at some standard, background level, in which case market power is understated only if ε_i is atypically high, and market power actually is overstated when it is unusually low. See *supra* note 48.

¹¹⁴ Indeed, it can be confusing to make sense of the rivals' supply elasticity in a combined market. As price is increased, more of each of the, say, two goods will be produced by rivals. However, these two sources of supply will, except by coincidence, increase at different rates. The rates could differ widely; for example, there may be capacity constraints in one market and nearly constant marginal cost with no such barrier to expansion in the other. If what we care about is the firm's ability to elevate the price of the initial good, then these two different supply responses have significantly different implications. (Even a hypothetical monopolist of the markets for both goods, as imagined in applying the U.S. Merger Guidelines' methodology, would care about the differential responses and, as a consequence, would elevate the prices in the two markets by different amounts.) Cf. *supra* notes 38 & 85 (discussing how the use of partial versus general equilibrium analysis is related to making particular assumptions about the supply response in the market for the other good).

¹¹⁵ See *supra* note 48.

Consideration of rivals' elasticity of supply leaves one with essentially the same view of the deficiencies of the market definition / market share paradigm as results from an analysis that, for concreteness, focuses on demand substitution. The primary exception is Part V, which addresses the decomposition of the market elasticity of demand. Even there, however, one could draw some analogies.¹¹⁶

B. Market Power and Condemnation

Some examinations of principles related to market power are complicated by concerns that acknowledgment of the existence or significance of market power in a particular case may lead to inappropriate legal condemnation, particularly in the monopolization context. The hesitation starts with the very definition of market power in terms of the Lerner index, as articulated in Part II. That is, market power is taken to be the extent to which price can profitably be elevated above marginal cost, which in turn is often but not always viewed as the competitive price, for in perfect competition it is familiar that price equals marginal cost. Because this Article focuses on whether the market definition / market share paradigm is a sensible means of making inferences about the extent of market power, not about substantive prohibitions in competition law, essentially no attention has been devoted to the purposes for measurement, which is to say the implications of finding substantial market power or concluding that it is insignificant. The basic reason for this division of labor is that the present analysis aims to show that the market definition process is an inherently unsound means of achieving its ends, so the question of the use to which these ends are put is essentially irrelevant to the inquiry. Nevertheless, given the anxiety in the literature and the resulting potential for confusion, a few comments are in order.

Focusing for concreteness on monopolization cases, standard doctrine requires two elements for liability: monopoly power (which is to say a significant degree of market power) and an exclusionary practice by which such power is created or sustained.¹¹⁷ Motivated by Judge

¹¹⁶ For example, it would be a mistake to focus only on certain sources of supply response when there are many, and if one examines, say, the supply elasticity of a particular rival, it is important to weight it by the rival's size. (Thus, if a given rival will expand its supply 3% in response to a 1% increase in price, but that rival supplies only 1% of the market, the aggregate supply response due to this rival would be only 0.03%, which is negligible.)

¹¹⁷ See, e.g., *United States v. Grinnell Corp.*, 384 U.S. 563, 570-71 (1966). EU Article 102 (formerly 82) prohibits the abuse of a dominant position, which clearly requires both a dominant position (taken here roughly to correspond to significant market power) and some form of abuse. Both elements are heavily contested; that is, there is no strong consensus on how much market power is required or on how to define exclusionary practices (or on the possible relationship between the two, notably, whether and how the requisite power should depend on the practice un-

Learned Hand's famous *Alcoa* opinion, which is often seen as condemning success despite his assertions to the contrary,¹¹⁸ and other cases, many worry that a finding of substantial market power will automatically, or at least too readily, lead to condemnation, even when no improper behavior is present. As a consequence, some commentators seem reluctant to find market power even when prices are elevated significantly. And one means of making market power harder to find is to define it so that it is present less often.

One particular alternative specification of the concept of market power is to deem it nonexistent when behavior is benign or socially desirable — that is, to make a finding of an abusive practice a prerequisite to a finding of market power.¹¹⁹ A gentler approach is to press on the use of marginal cost in the Lerner index. Some suggest that marginal cost is ordinarily measured in the short run whereas the long-run marginal cost (often imagined to be higher¹²⁰) is more appropriate.¹²¹ One rationale for this view is that markets are sometimes out of equilibrium, undergoing adjustment, such as when firms expand capacity over time to meet an unanticipated rise in demand.¹²² Others argue

der scrutiny). These debates are tangential to the analysis of market definition as a means of measuring market power.

¹¹⁸ See, e.g., *United States v. Aluminum Co. of Am. (Alcoa)*, 148 F.2d 416, 430 (2d Cir. 1945) (“The successful competitor, having been urged to compete, must not be turned upon when he wins.”).

¹¹⁹ See, e.g., *Carlton*, *supra* note 13, at 3 (expressing a concern about inferring market power from market share on the ground that it is difficult to weigh increased market power against possible efficiency gains); *id.* at 19 (“My experience is that courts ask whether market power exists in the presence of the alleged bad act, a question with the potential to be answered in a misleading way if one ignores the efficiency justification for the alleged bad act . . .”); *Fisher*, *supra* note 49, at 18 (“If the share is maintained solely because of low prices or better products, then we are looking at what competition is supposed to do and not at a monopoly. This is, of course, closely related to the legal position that a monopoly acquired by ‘superior skill, efficiency, or foresight’ does not violate the antitrust laws. I would prefer to say that a large share acquired in such ways is not a monopoly at all.”); *id.* at 28 (“A firm which maintains a large share of the market because of behavior forced on it (‘economically inevitable’) or solely because of being better (‘superior skill, efficiency, and foresight’) is a firm which does not have monopoly power at all. Monopoly power is the power to maintain a high share and earn supranormal profits *without* being better.”).

¹²⁰ Long-run marginal cost may be higher or lower than short-run marginal cost. It may be higher because fixed costs are variable in the long run and thus are included, but it may be lower because capacity may be expanded or technology updated.

¹²¹ See, e.g., *Fisher*, *Economic Analysis*, *supra* note 41, at 139; *Werden*, *supra* note 13, at 381–82.

¹²² See, e.g., *Fisher*, *supra* note 49, at 10–12, 20–22. This scenario is questionable, for in a competitive market, price would equal short-run marginal cost even after an unexpected demand spike. Each firm would equate its marginal cost to the higher price; with rising marginal cost curves, this involves some expansion even before new capacity is created. (If capacity is literally constrained, and marginal cost does not rise rapidly as the capacity limit is approached, then firms would produce at capacity, at least in the short run. But a proper definition of marginal cost would have its level rising discontinuously from its prior finite level to infinity at that level of output, and one could thus technically say that the resulting equilibrium was nevertheless consis-

that prices above marginal cost may be consistent with competition in some appropriate sense — that is, the competitive price, thus interpreted, exceeds marginal cost — perhaps due to the presence of fixed costs.¹²³ As a consequence, it is suggested that markups should be judged relative to a higher benchmark.¹²⁴

Despite these concerns, clarity in analysis is best served by maintaining the traditional distinction between the definition and measurement of market power, on one hand, and the determinants of liability, on the other. A useful illustration (often noted in the literature by commentators taking opposing views on the present question) in-

tent with price equal to short-run marginal cost. Realistically, marginal cost may well rise quickly but not at an infinite rate as “full capacity” is approached, and firms do not often produce at literally 100% of feasible capacity in any event.) Under these circumstances, the imagined temporary divergence — and thus the temporary indication of market power via use of the Lerner index — simply does not arise. (It is true that firms that otherwise would have earned no profits will make some profit temporarily.) Another case sometimes mentioned involves rents on a scarce input owned by a firm. However, if one imagines the resource as if it were owned by a separate entity and thereby treats the opportunity cost of consuming the resource as an input cost, which is theoretically proper, then pricing is not really above marginal cost (although it may exceed a historically based accounting measure of incremental cost).

¹²³ See, e.g., Baumann & Godek, *supra* note 107, at 329; Lawrence J. White, *Market Power and Market Definition in Monopolization Cases*, in 2 ISSUES IN COMPETITION LAW AND POLICY 913, 921–22 (Wayne Dale Collins ed., 2008); cf. Fisher, *Economic Analysis*, *supra* note 41, at 137 (expressing a concern about inferring significant market power from a high market share because “a large share can simply represent greater efficiency or product quality on the part of the alleged monopolist”). These arguments can be obscure in two ways. First, it may not be explained how prices in excess of marginal cost may be sustained in the presence of competition since firms with marginal cost below price have an incentive to increase output. In supposing that they will refrain, it seems that noncompetitive behavior is contemplated. Second, the meaning of competition is sometimes unclear or the term may be used in unconventional ways, as suggested by the first point. A straightforward case involving price in excess of marginal cost in long-run equilibrium involves product differentiation, where firms incur fixed costs (which limits the extent of entry) and cannot expand output and steal rivals’ customers simply by pricing infinitesimally below what others charge, due to consumers having at least some nontrivial preference for different producers’ products. See, e.g., JEAN TIROLE, *THE THEORY OF INDUSTRIAL ORGANIZATION*, ch. 7 (1988). This case is often termed monopolistic competition, a form of imperfect competition, precisely because it does entail each firm possessing some market power (defined in terms of the Lerner index). Further refinement and resolution of these issues is unnecessary for present purposes.

¹²⁴ Interestingly, this argument about the definition and measurement of market power is generally advanced in the monopolization context without comment on its implications elsewhere, notably, for horizontal mergers. Under this approach, two parties could merge to monopoly and concede that prices would rise, say, 20%, and claim that the government has not made out a *prima facie* case. The reasoning would be that market power is required for liability and, since market power is defined not as price elevation above marginal cost or what is ordinarily understood to be a competitive price, but above whatever level is necessary to recover fixed costs, incentivize innovation, and so forth, market power has not even been shown to exist unless the monopoly price is demonstrated to be unnecessary in this regard. (Implicitly, the firms are claiming that price was below the appropriate level before the merger.) Of commentators who advance these views about the definition or meaning of market power and concerns about fixed costs, innovation, and so forth, I am unaware of any who then consistently deploy them in this fashion to criticize standard analysis of horizontal mergers.

volves a powerful patent that allows its owner to reap very substantial rewards on account of the ability to price significantly in excess of marginal cost for an extended period of time.¹²⁵ As defined here, this case would involve the possession of significant market power. To be sure, the firm's ability to earn this stream of profits may well be socially desirable in light of the *ex ante* stimulus to innovation provided by the prospect of such gain, which is the central justification for awarding patents. Market power attributable to the legitimate exercise of a valid patent should not be the basis for liability.

In cases like this, it seems better to say that there is, indeed, market power, but no liability, rather than to deem market power to be absent. The reason is that, if we do the latter, we then would not be able to hold liable abusive anticompetitive behavior associated with the patent. Suppose, for example, that the patent was improperly obtained, that the patent holder engaged in exclusionary practices to deter the entry of rivals or to drive some out of the market, or that holders of multiple competing patents wish to merge in order to elevate prices to a much greater extent. It seems that those deeming market power to be nonexistent, by definition, would be forced to exonerate — to not even scrutinize — such acts.¹²⁶

Instead, it appears that such analysts would change their view on market power. But at that point, we have a largely semantic game,

¹²⁵ Views closest to those here appear in Schmalensee's commentary on qualms expressed by Landes and Posner, *supra* note 13, at 956–57, regarding the finding of market power in certain settings. See Schmalensee, *supra* note 14, at 1795–96; *id.* at 1808 n.74 (“If one followed Landes and Posner's recommendation, most ‘clean-handed’ monopolists could thus rebut the inference that they had *any* market power. Evidence of the sources of market power is much more appropriately used to judge the legality of the acquisition and exercise of that power. It would cause considerable confusion if ‘market power’ were to be redefined (as Landes and Posner implicitly suggest) so that ‘clean-handed’ firms could never possess it.”).

¹²⁶ Consider another example outside the monopolization context. One commentator would limit market power to prices in excess of levels set by any process, not necessarily competitive as conventionally understood, as long as it is legal; for example, he (like others) holds that price elevation through oligopolistic coordination is legal as long as no secret meetings and the like are involved. See Baker, *supra* note 2, at 143 n.53. That definition creates problems for the application of a rule of reason analysis to, say, trade association behavior that is alleged to facilitate oligopolistic coordination. If such coordination is legal and is occurring, then the elevated price is the baseline, there is no further elevation relative to that (by definition), and there is thus no market power. As a consequence, if the rule of reason inquiry has a market power prerequisite (so we do not have to scrutinize all manner of cooperative activity), then the trade association would be immune. The response would be to say that, in such a case, the practice should be challenged and, if invalid, then the original premise that the oligopolistic price elevation was perfectly legal no longer holds, so we can find market power after all. As in the example in the text, this manner of thinking and communicating seems needlessly convoluted. What is more surprising is that one would insist on expressing ideas this way rather than in the simpler and more straightforward fashion advocated in the text and practiced broadly.

and one that can be confusing.¹²⁷ This seems to be especially so when a market power requirement is deployed as a screen, leading to the cessation of investigations or dismissal of cases when it is sufficiently low, without having to bother with analysis of the challenged acts.¹²⁸

In any event, these concerns about how best to define market power are unrelated to the question taken up in this Article. Whether using these alternative or conventional approaches, the market definition / market share paradigm is equally unhelpful. Standard analysis of market definition is not particularly attuned to fixed costs, recouping the costs of innovation, and the like, and certainly not in any distinctive manner. More directly, each criticism stands independently of this debate over definitions: formula (4) that relates market share and other factors to market power still is not meaningful beyond homogeneous goods markets, comparative statements about market shares still need some concept of a standard reference market to be intelligible, choosing the best market still requires an estimate of market power (however defined) as a prerequisite, and the focus on particular cross-elasticities rather than on the elasticity of demand is still a mistake for the reasons given — regardless of whether one is looking, for example, at percentage rates of change in response to price movements in the short run or in the long run.

C. *Is Market Definition Ever Needed?*

In light of the demonstration in Parts II–V that the market definition / market share paradigm is, as a matter of principle, incapable of helping to make valid inferences about market power, one can hope (and it seems logically entailed) that market definition is not really necessary after all. This section briefly mentions some of the main ways

¹²⁷ Taking the definitions literally, two internally consistent results are possible, and the formal statements do not provide any basis for selection: (1) There is no market power; hence, there is no violation due to abuse (since it is not counted as an abuse in the absence of market power), and since we have just concluded that there is no finding of abuse, our original statement of no market power is correct (since we deem it nonexistent in the absence of abuse). (2) There is market power; hence, there is a violation due to abuse (which we can deem to exist, having met the market power requirement), and since we have just found an abuse, our original statement that there is market power (which required the simultaneous existence of abuse) is valid. This tension can easily be seen in the preceding note, since Jonathan Baker defines the benchmark price in that case as that reached by “sellers [who] engage in oligopoly conduct without reaching an agreement *or otherwise violating the antitrust laws*,” *id.* (emphasis added), since the question motivating the market power inquiry is whether the antitrust laws have been violated.

¹²⁸ If one believes that it is existing practice that is muddled and, in particular, that courts often fail to insist on demonstration of abuse or are too generous in finding it, then the seemingly convoluted approach may nevertheless have pragmatic appeal. In that event, it might be helpful to introduce some new, supplemental terminology, perhaps something like “provisional market power” — that is, market power, as defined by the Lerner index, but deemed tentative until there is a finding that an exclusionary practice (or other anticompetitive behavior) has occurred.

that market power is relevant under competition law and confirms that market definition really is unnecessary as a matter of policy, leaving questions about legal doctrine to section E. Monopolization is considered first and horizontal mergers second,¹²⁹ followed by some more general observations.

Suppose that a monopolistic act is directed at alternative actual or potential suppliers of the same product produced by a defendant firm alleged to have monopoly power and to be abusing it. If the dominant firm model is applicable, expression (4) provides a means of determining how much market power the firm has. In addition, it provides a way to assess the extent to which the act, if successful as alleged, would augment market power and thus induce a profit-maximizing firm to increase price. In this situation, the exclusionary act is imagined to reduce rivals' supply, which augments the firm's market share and may also reduce rivals' supply elasticity. Expression (4) directly translates these effects into influences on our desired market power measure. That is, one can apply the formula to both the scenario with and that without the act and measure the difference in market power. As explained in Part II, use of this formula assumes that one sticks with the homogeneous goods market — however high or low may be the firm's share, the market elasticity of demand, and rivals' supply elasticity. Indeed, redefinition undermines our ability to employ the formula (that is, unless we essentially undo the redefinition).

Next, suppose that the alleged monopolistic act is directed at a demand substitute. In that case, the effects can include an increase in share and a reduction in the market elasticity of demand, and we can again use expression (4) — and, accordingly, we would not want to redefine the market. How can we determine the effect on the market elasticity of demand if, say, the practice is directed at one or a particular subset of possible substitutes? One would employ essentially the method outlined in section V.C that addresses the use of cross-elasticities to assess mergers in differentiated product industries — again, without ever redefining markets.¹³⁰

Market definition also currently plays a central but unnecessary and affirmatively misleading role in evaluating horizontal mergers.¹³¹ Consider first unilateral effects — the merging firms' incentives to

¹²⁹ The analysis in this brief section is not comprehensive even as to the applications it considers. However, it covers many of the most important settings, and others can largely be assimilated to one or another of them. For example, many horizontal restraints examined under the rule of reason, where market power inquiries may be employed, involve fears about facilitating collusion, so the market power analysis used in the case of coordinated effects in horizontal mergers would roughly be applicable.

¹³⁰ See Kaplow & Shapiro, *supra* note 12, at 1185–86.

¹³¹ For comments specifically on the approach to market definition in the U.S. Merger Guidelines, see notes 99 and 138.

raise prices, taking rivals' prices as given (that is, assuming no coordination)¹³² — which have been the main focus of enforcement agencies for quite some time.¹³³ The appropriate method of analysis for differentiated products mergers is sketched in section V.C — and it involves no market definition inquiry. For unilateral effects in homogeneous goods markets, using the Cournot model that is often employed, one would simply stick with the homogeneous goods market (which is the domain of standard Cournot analysis). Application of that model makes use of a variant of expression (4) and, in any case, does not involve market redefinition.¹³⁴

Now consider possible coordinated effects due to horizontal mergers.¹³⁵ Typically, one is considering oligopolistic price elevation via interdependent pricing among firms that produce goods that are homogeneous (or nearly so), for coordination is much more difficult when differentiation is significant. The market power question concerns how much the firms could elevate price if, indeed, the merger enables successful coordination. To measure this, we need to know market power in the homogeneous goods industry, so redefining markets is beside the point.¹³⁶ To be sure, the presence of close substitutes may imply a higher market elasticity of demand (as discussed at length in Part V), but that market elasticity is what it is (however we may wish to determine it in a particular case), and it too feeds directly into the formulas presented in subsection II.B.1, which, again, are applicable in the homogeneous goods market. Therefore, in each scenario, market redefinition is unnecessary and can only misdirect the analysis.

A few broader remarks are in order. As mentioned at many points, much analysis is properly conducted in the homogeneous goods market. That is a “market,” but one that does not involve any market definition analysis — asking whether it or some broader market is best — in order to use it. Instead, one uses this market because there exist developed tools for making market power inferences from basic information about such a market. Redefining the market undermines our

¹³² See, e.g., U.S. MERGER GUIDELINES, *supra* note 5, § 6.

¹³³ See, e.g., SIGRID STROUX, US AND EC OLIGOPOLY CONTROL 187–88 (2004).

¹³⁴ See, e.g., Kaplow & Shapiro, *supra* note 12, at 1084–85 (presenting the standard analysis indicating that a firm's markup equals the firm's market share divided by the elasticity of market demand and also the formula indicating that the industry-wide average, output-weighted margin equals the HHI divided by the elasticity of market demand; for an illustration and qualification, see note 99).

¹³⁵ See, e.g., U.S. MERGER GUIDELINES, *supra* note 5, § 7.

¹³⁶ See, e.g., Werden, *supra* note 41, at 532. If the firms we suppose to be coordinating cannot raise price significantly even if their coordination is perfect, there is no point in inquiring into broader markets. This inquiry would be interesting only if we were supposing that cross-market coordination might occur. The success of such coordination, if achieved, can be determined much as in the case in which we measure the unilateral effects of a merger between two firms producing differentiated products.

ability to conduct such analysis properly. These points are the central conclusions of Part II.

Also, the present analysis does not rule out that some thinking in terms of market metaphors might be helpful, for example, to guide enforcement agencies' setting of priorities.¹³⁷ If agencies wish to examine markets that are ripe for price fixing, they would naturally focus on those that seem most worrisome, which would be markets in which coordination is more likely to be feasible and in which resulting price elevations, conditional on success, would be largest. As mentioned, feasibility is greatest in homogeneous goods markets. Therefore, it would be natural to focus primarily on such markets and, moreover, to pay particular attention to those with lower market elasticities of demand. In this exercise, market redefinition is not useful.¹³⁸

Yet another, stronger use of market definition is as a quick screen, whether, again, by enforcement agencies allocating scarce resources, or by adjudicators, such as in dismissing cases that almost surely are baseless. The problem is that we can only engage in abbreviated analysis using market definition — which requires choosing one definition over others — if we overcome the obstacles revealed throughout this Article, which are argued to be insurmountable. Consider just the point in Part IV that one cannot choose the best market without already having in hand an estimate of market power that is at least as informative as one's market share inference in the chosen (best) market. This logic applies equally to truncated inquiries based on limited inputs that are undertaken for screening purposes.

Some elaboration of this final point may be helpful. Suppose that the broader market definition is obviously better and that this conclusion is clear enough that we do not need extensive data gathering and

¹³⁷ A number of economists have suggested to me that the market definition metaphor may also summarize decisions about how broadly to collect data on possible substitute products when empirically estimating demand elasticities or for purposes of calibrating merger simulations. Such boundary decisions, however, reflect costs of data collection and empirical hunches on how likely a (moderately to substantially distant) substitute is to influence the estimation of other parameters, a qualitatively different calculus from that involved in defining markets for purposes of making market power inferences. Additionally, some find the metaphor useful to help suggest relevant paths of inquiry at the outset of a market power investigation. This usage — again, quite different from that examined in this Article — may be fairly benign, although pitfalls have been mentioned, including that one may be led to focus on particular cross-elasticities rather than on the demand elasticity.

¹³⁸ For markets in which even successful oligopolists could not significantly elevate price, one might keep broadening the boundaries, as under the Merger Guidelines, in order to see which broad markets would involve significant price elevations conditional on success. (Indeed, the hypothetical monopolist SSNIP test, as a purely logical matter, seems much more tailored to answering this question, *see* Kaplow & Shapiro, *supra* note 12, at 1172, than to those it purports to address, a criticism developed via numerical example in note 99.) However, since coordination would be infeasible in most such broad markets, this method of prioritizing investigations does not seem very sensible.

expert analysis. Then, it is necessarily true that we are confident from what we can see that, in the narrow, homogeneous goods market, there is negligible market power. (That is, after all, what it means to state that the broader market is clearly superior, at least using the criteria for the best market definition that are examined in section IV.A.) For example, if we were considering a horizontal merger in the market for orange juice, to reasonably conclude that the market should also include grapefruit juice, or all fruit juices, or raw oranges, or anything else, we would necessarily have decided that the substitutes in question greatly limit the ability of the merged orange juice firm to raise price nontrivially — that is, that the market elasticity of demand in the narrow, homogeneous goods market is high enough to negate significant market power. But if we have already reached that conclusion, we are done.¹³⁹ To take another example, shares may be very low even in the narrow, homogeneous goods market, in which case there typically would not be significant market power. In either case, if it is indeed possible to screen using such immediate empirical judgments,¹⁴⁰ we do not need to (and, logically, should not want to) consider redefining markets in order to do so. It may be that judges would feel uncomfortable directly announcing conclusions about market power but more comfortable redefining the market, a point considered next.

D. *The Allure of Market Definition*

Courts as well as analysts and enforcement agencies are strongly attracted to the market definition / market share paradigm. Preceding analysis indicates that, as a matter of logic and fairly basic economics, this fondness is misplaced. Section E examines the extent to which insistence on the market definition process is rooted in legal doctrine, whereas the present section considers briefly the notion that the allure is due to the (mis)impression that choosing markets is less difficult than examining market power directly.

¹³⁹ More precisely, as outlined in section IV.A, to have decided that the market power inference from the market share in the broader market is more accurate (or more likely to lead to the right answer) than the market power inference from the market share in the narrower market is to assume the conclusion about market power.

¹⁴⁰ There is a real danger of overlooking important instances of market power. For example, in mergers in differentiated product industries, it is easy to reject narrow market definitions as obviously incorrect, but more careful analysis of some such mergers indicates that looks can be deceiving. The most familiar example is *FTC v. Staples, Inc.*, 970 F. Supp. 1066 (D.D.C. 1997). (Even if one disagrees with the government's econometric analysis in that case — a question on which no opinion is offered here — the point is that a conclusion many would have seen as obvious is not.) Additional errors can arise in cases with highly inelastic supply, which may exist with oil refining and electricity generation, for example. In such industries, low shares can convey significant market power (assuming that the market elasticity of demand is not very high), a conclusion that is apparent from expression (4).

We ordinarily suppose that making yes/no decisions or choosing one item from a small set is easier than coming up with a quantitative estimate. Likewise, if one has to rationalize one's choice (such as in a judicial opinion or an agency announcement), the task of explaining why *A* is better than *B* seems easier than defending an estimate of 7.8. Moreover, myriad legal doctrines require dichotomous choices rather than numerical pronouncements (damage awards and discretionary fines being obvious exceptions), which tends to make courts, lawyers, and legal academics less comfortable with rules requiring quantification. This familiar phenomenon helps to explain the appeal of the market definition / market share paradigm¹⁴¹ and perhaps also a reluctance to scrutinize imperfections that may require abandonment of this seemingly simpler method.

In light of the criticisms in preceding Parts, this undoubtedly common impression has to be an illusion. Part IV makes this point especially apparent. As explained, one cannot choose which market definition is best, even from a list limited to two choices, without first forming an estimate of the extent of market power — the very thing that a simple, binary decision is hoped to render unnecessary.

This reality may be elusive because there are many other settings in which we make choices without first formulating quantitative estimates, particularly when cases are not close. For example, one might look at a tree and see plainly that its height is much closer to that of a sapling than of a giant redwood. By contrast, there is no way to see (or feel or otherwise directly sense) the magnitude of a firm's market power or the degree to which it differs from the inference one would make from a market share in some market. No aspect of the analysis is sensory; "markets" as the term is used in this context are pure abstractions.¹⁴² Consider, for example, the illustration at the end of the preceding section involving whether the orange juice market is a relevant market. There is nothing one can view (that is, other than data, reports of consumer behavior, and so forth) to see in the literal sense whether one particular market definition is "closer" to the truth than another. Among other things, one would first have to generate an image of the truth, which is precisely the point of this argument.

There is another reason that quantitative statements are resisted: if a court found that a dominant firm had the ability to elevate price

¹⁴¹ See, e.g., Kaplow, *supra* note 22, at 1832; AREEDA, KAPLOW & EDLIN, *supra* note 14, at 492.

¹⁴² See *supra* note 75. One might add that they also may be hypothetical abstractions, concerning, for example, behavior that would occur if a merger were allowed to go forward (making it a bit like asking, at the time a seed is planted, whether the height of a never-before-seen type of tree will be closer to that of one versus another existing tree, a question one could answer only by using available science to predict the height of the prospective tree).

20%, it may then feel compelled to articulate various market power thresholds — here, the amount of power that constitutes “monopoly power” or “dominance” — in quantitative terms as well. Yet this too has been resisted. Most commonly, the two questions — how much market power exists in a particular case and how much market power suffices — are conflated, with market shares being used as a combination metric.¹⁴³ Further reflection on this method suggests that, rather than answering both questions in a satisfactory manner, neither is illuminated. The point here is that courts’ and others’ reluctance to make quantitative statements on the legal policy question regarding the requisite level of market power (or increase in market power) is reinforced by eschewing direct answers to the question of how much market power is present — which avoidance is aided by adhering to the market definition / market share paradigm.¹⁴⁴

E. The Law on Market Definition

Although this Article is concerned with what makes sense as a conceptual matter, this section briefly considers the extent to which rejection of the market definition process can be reconciled with existing law. Of course, standard practice, at least on the surface, relies heavily on the market definition / market share paradigm, and countless cases have indicated with varying degrees of firmness that market definition is required, as emphasized at the outset. The analysis in this section is against that backdrop. Since this state of affairs is so familiar, the discussion here will err on the side of boosterism in order to give sharper exposure to a different outlook and provoke further thought on the subject. The emergent bottom line is that it seems plausible that existing practice could adapt substantially if agencies and adjudicators were receptive. Also important, it seems at a minimum that room can be made for the incorporation of alternative, more direct means of as-

¹⁴³ See AREEDA, KAPLOW & EDLIN, *supra* note 14, at 492–93, 498–99. This problem is elaborated in Louis Kaplow, *Market Share Thresholds: On the Conflation of Empirical Assessments and Legal Policy Judgments*, 7 J. COMPETITION L. & ECON. (forthcoming 2011).

¹⁴⁴ Consider a further implication of courts’ reluctance to offer quantitative statements about anything other than market share when joined with a further queasiness about being too specific about what market shares suffice. In one hypothetical monopolization case, suppose that one market definition involves a dominant firm share of 90% and another definition results in a share of 55%. Since the court may be uncertain about the right ultimate legal answer if it goes with the latter, it may have some inclination to select the definition yielding the 90% share in order not to have to confront the issue. Likewise, in another hypothetical case, in which one definition yields a share of 55% and another a share of 20%, the latter choice would be more comfortable for the same reason. The murkiness of the market definition process also makes it fairly easy for an adjudicator to reach any decision that might be desired on other grounds (for example, if the claim that a practice is abusive seems frivolous, but the analysis that would support that feeling is time-consuming and complex, one can often select a broad market definition that renders the other inquiry moot).

sessing market power that are increasingly being developed, even if the market definition process is not rejected across the board. To develop these points, this section will first consider a fairly immediate way to subordinate market definition even under doctrine that is assumed to require it and then will examine aspects of legal doctrine (with a U.S. focus) that bear on whether market definition can be discarded altogether.

A straightforward temporizing solution, assuming that market definition is formally required, is to make it explicitly the final step in the analysis. That is, one would first use whatever evidence and modes of inference enable one to obtain a best estimate of market power, without any use of the market definition / market share paradigm, and then use that information to choose a market that yields the right legal conclusion, as elaborated in subsection IV.A.2. As explained, this final step is theoretically pointless. But if the law is thought to demand it, it can be done.

This result-oriented technique should be easy to reconcile with a formal market definition requirement. First, it does in the end define a market. Second, as Part IV shows, this sort of method is really the only logically coherent way to do so if one wants to avoid error. Thus, unless there is both an inflexible market definition requirement and a legal insistence that it be met in some other, illogical fashion, the method works.¹⁴⁵ Put another way, if courts or other decisionmakers were indeed convinced by the analysis presented above, it seems difficult to understand how they could be persuaded to adopt a market definition that they fully appreciated would lead to the wrong answer. Relatedly, it is hard to see how they could find expert testimony compelling when it was based on a mode of analysis that they believed to be incoherent.¹⁴⁶

¹⁴⁵ Even if other approaches were nevertheless allowed, it is hard to see how enforcement agencies and private parties would be prohibited from offering and making central to their case high-quality direct evidence of market power, including methods that economists have been developing in recent decades and will continue to formulate.

¹⁴⁶ The problem concerning the use of experts, which is widespread in market definition inquiries, is severe. In light of this Article's analysis, what could an expert say when purporting to offer an opinion on market definition (unless taking this results-oriented out)? For example, if asked whether more conventional analysis of market definition was rooted in accepted economic understanding, an honest answer (other than "no, not at all") would be difficult to come by. As mentioned in subsection II.B.2, the construct really does not exist in industrial organization economics. As a consequence, an argument could be made in U.S. courts that conventional expert opinions on market definition should not be admissible. Rule 702 of the Federal Rules of Evidence — as amended in light of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), and subsequent cases — requires that an expert's "testimony is the product of reliable principles and methods." Query what those principles and methods could be with regard to market definition in light of the concept's nonexistence in the field of expertise and the foregoing analysis in the text.

Moreover, it appears likely that this sort of approach is already employed, perhaps even frequently, without actors necessarily being explicit about the practice. The method seems apparent in the *Staples* merger decision,¹⁴⁷ in which there was direct evidence of market power and the court presented its market definition choice only after assessing that evidence and coming very close to (if not reaching) the statement of a legal conclusion.¹⁴⁸ If more direct inquiries into market power were undertaken more often and decisionmakers were more explicit that market definition was little more than a restatement of their conclusion, one might expect the practice to spread further and that any requirement for conclusory articulations of relevant markets would eventually wither.

Consider next the extent to which the law demands that relevant markets be defined. As mentioned, the discussion here will focus on U.S. doctrine for concreteness, although it should be noted that laws elsewhere are often similar in relevant respects and many other jurisdictions' governing statutes do not explicitly require market definition.¹⁴⁹ The law's purpose, case law, and statutory requirements will be noted in turn.

First, and perhaps most distinctive to the United States (but showing increasing signs of acceptance elsewhere), modern competition law and policy affirmatively embrace basic economic understandings, both for assessing behavior in particular cases and when formulating legal rules. The U.S. Supreme Court in *Sylvania's* overruling of *Schwinn* emphasized economic substance over legal form: "Nor is there even an assertion in [*Schwinn*] that the competitive impact of vertical restrictions is significantly affected by the form of the transaction. . . . [E]ven the leading critic of vertical restrictions concedes that *Schwinn's* distinction between sale and nonsale transactions is essentially unrelated

¹⁴⁷ *FTC v. Staples, Inc.*, 970 F. Supp. 1066 (D.D.C. 1997).

¹⁴⁸ *Id.* at 1075–81; see also James A. Keyte & Neal R. Stoll, *Markets? We Don't Need No Stinking Markets! The FTC and Market Definition*, 49 ANTITRUST BULL. 593 (2004) (arguing that the FTC has been moving away from market definition in merger and nonmerger cases); Robert H. Lande & James Langenfeld, *From Surrogates to Stories: The Evolution of Federal Merger Policy*, ANTITRUST, Spring 1997, at 5, 6 ("[T]here have been many recent attempts by the Federal Trade Commission and the Department of Justice to shift the focus of investigations away from market definition . . . [toward] try[ing] directly to predict what will happen to future industry prices."); Kühn, *supra* note 80, at 316 (suggesting that the European Commission "appears to be open to the general trend of moving toward a practice of joint determination of market definition and market power instead of insisting on a mechanical two-step exercise of market definition and market share assessment").

¹⁴⁹ There are exceptions. See ten Kate & Niels, *supra* note 41, at 297 ("Some more recently installed competition regimes, such as Mexico in 1992, have even incorporated the term 'relevant market' explicitly in the official language of their competition statutes, thereby turning it from an analytical tool in the determination of market power into a statutory prerequisite to decisions in competition cases.").

to any relevant economic impact.”¹⁵⁰ In *Matsushita*, the Court was specifically concerned about allegations that made no “economic sense,”¹⁵¹ notably a situation in which defendants “had no rational economic motive to conspire”¹⁵² in the fashion that a plaintiff suggested, not about claims that are wholly in accord with economic rationality but fail to satisfy formalistic legal requirements arising from the Court’s interpretations of the statute. More recently, *Leegin*’s¹⁵³ overruling of *Dr. Miles*¹⁵⁴ was defended on the ground that the latter “Court justified its decision based on ‘formalistic’ legal doctrine rather than ‘demonstrable economic effect.’”¹⁵⁵ Against this background, it seems difficult to sustain a requirement and an interpretation thereof that seems entirely formalistic and makes no economic sense. This conclusion is further supported by the Court’s embrace of an economic concept of market power, toward which the market definition inquiry purports to be directed.¹⁵⁶

The case law, as noted, appears to contain an overwhelming endorsement of the market definition / market share paradigm, as evidenced by its widespread use as well as numerous statements of differing degrees of strength in various contexts about the need to define a relevant market.¹⁵⁷ Even so, there are many reservations, both explicit

¹⁵⁰ *Cont’l T.V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36, 54, 56 (1977) (overruling *United States v. Arnold, Schwinn & Co.*, 388 U.S. 365 (1967)).

¹⁵¹ *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574, 587 (1986).

¹⁵² *Id.* at 596.

¹⁵³ *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 127 S. Ct. 2705 (2007).

¹⁵⁴ *Dr. Miles Med. Co. v. John D. Park & Sons Co.*, 220 U.S. 373 (1911).

¹⁵⁵ *Leegin*, 127 S. Ct. at 2714 (quoting *Sylvania*, 433 U.S. at 58–59). The *Leegin* Court further emphasized that antitrust principles on vertical restraints were to be formulated by reference to “differences in economic effect” so that “it is necessary to examine . . . the economic effects of vertical agreements to fix minimum resale prices” to determine what legal rule should apply. *Id.*; see *id.* at 2722.

¹⁵⁶ See sources cited *supra* note 13.

¹⁵⁷ See, e.g., *Spectrum Sports, Inc. v. McQuillan*, 506 U.S. 447, 459 (1993) (“The concern that § 2 might be applied so as to further anticompetitive ends is plainly not met by inquiring only whether the defendant has engaged in ‘unfair’ or ‘predatory’ tactics. Such conduct may be sufficient to prove the necessary intent to monopolize, which is something more than an intent to compete vigorously, but demonstrating the dangerous probability of monopolization in an attempt case also requires inquiry into the relevant product and geographic market and the defendant’s economic power in that market.”). It is important to keep in mind, however, that the Court in *Spectrum* was rejecting an interpretation under which intent or conduct alone was asserted to be sufficient to establish a dangerous probability of success, without any separate market power inquiry; no attention was directed to whether means of establishing market power directly might suffice in lieu of the indirect inference drawn using the market definition process. See also *DG Competition Article 82*, *supra* note 1, ¶ 11 (“The concept of dominance contained in Article 82 of the Treaty relates to a position of economic strength on a market. In the application of Article 82 it is therefore necessary to define a relevant market.”); *Developments — and Divergence — in Merger Enforcement*, ANTITRUST, Fall 2008, at 9, 13, 19 (in panel discussion on merger enforcement, European lawyers indicate a greater focus on market definition and market share by the European Commission than by U.S. enforcement agencies).

and implicit. One notable point is that, pretty much from the beginning, the Supreme Court insisted that market shares should be viewed as only prima facie indicators, and that other evidence could suggest nonstandard inferences or call for a revision of the market definition.¹⁵⁸ An underlying thrust of such statements and numerous other pronouncements in the cases is that the market definition process is highly imperfect, provides only a window into market power, and is subject to correction if not displacement in light of evidence that persuasively bears on actual market power. Indeed, the Supreme Court and numerous lower courts have explicitly endorsed proof of market power (or actual anticompetitive effects) by other, more direct means when such are available.¹⁵⁹ Some commentators similarly view the courts as treating the market definition process as a means rather than

¹⁵⁸ The opinion most associated with this view is *United States v. General Dynamics Corp.*, 415 U.S. 486 (1974), which in turn cites *Brown Shoe* for the proposition that “statistics concerning market share and concentration, while of great significance, [are] not conclusive indicators of anticompetitive effects.” *Id.* at 498; see also Landes & Posner, *supra* note 13, at 979 (citing various Supreme Court cases in support of “the recent trend toward regarding market share statistics in merger cases as providing merely presumptive evidence of market power, which can be rebutted by bringing in other factors”). The view that market shares require interpretation is not confined to the United States. See, e.g., *DG Competition Article 82*, *supra* note 1, ¶ 30 (“In any event, the Commission interprets market shares in the light of likely market conditions, for instance, whether the market is highly dynamic in character and whether the market structure is unstable due to innovation or growth.”); *id.* ¶ 32 (“The strength of any indication based on market share depends on the facts of each individual case. Market share is only a proxy for market power, which is the decisive factor. It is therefore necessary to extend the dominance analysis beyond market shares, especially when taking into account the difficulty of defining relevant markets in Article 82 cases . . .”).

¹⁵⁹ See, e.g., *FTC v. Ind. Fed’n of Dentists*, 476 U.S. 447, 460–61 (1986) (“Since the purpose of the inquiries into market definition and market power is to determine whether an arrangement has the potential for genuine adverse effects on competition, ‘proof of actual detrimental effects, such as a reduction of output,’ can obviate the need for an inquiry into market power, which is but a ‘surrogate for detrimental effects.’” (quoting 7 PHILLIP AREEDA, *ANTITRUST LAW* ¶ 1511, at 429 (1986))); *Toys “R” Us, Inc. v. FTC*, 221 F.3d 928, 937 (7th Cir. 2000) (“[Toys ‘R’ Us] seems to think that anticompetitive effects in a market cannot be shown unless the plaintiff, or here the Commission, first proves that it has a large market share. This, however, has things backwards. As we have explained elsewhere, the share a firm has in a properly defined relevant market is only a way of estimating market power, which is the ultimate consideration. . . . The Supreme Court has made it clear that there are two ways of proving market power. One is through direct evidence of anticompetitive effects.” (citing *Ind. Fed’n of Dentists*, 476 U.S. at 460–61)); Werden, *supra* note 13, at 380–81 (“While market share has long been the staple of market power analysis, three courts of appeals have held: ‘Market share is just a way of estimating market power, which is the ultimate consideration. When there are better ways to estimate market power, the court should use them.’” (quoting *Ball Mem’l Hosp., Inc. v. Mut. Hosp. Ins., Inc.*, 784 F.2d 1325, 1336 (7th Cir. 1986), an opinion by Judge Easterbrook, and citing in support an opinion by now-Justice Thomas, joined by now-Justice Ginsburg); *ANTITRUST LAW DEVELOPMENTS*, *supra* note 1, at 229, 347; Eric L. Cramer & Daniel Berger, *The Superiority of Direct Proof of Monopoly Power and Anticompetitive Effects in Antitrust Cases Involving Delayed Entry of Generic Drugs*, 39 U.S.F. L. REV. 81, 100 (2004).

an end in itself.¹⁶⁰ Courts do not tend to ignore probative evidence or logical argument about market power when such is offered, asserting that their hands are tied by the market definition regimen.¹⁶¹

U.S. statutes are sometimes taken to impose a market definition requirement, at least in certain contexts. Notably, Clayton Act section 7 prohibits mergers “where in any line of commerce or in any activity affecting commerce in any section of the country, the effect of such acquisition may be substantially to lessen competition or to tend to create a monopoly.”¹⁶² Some suggest that *Brown Shoe*, nearly a half century ago, interpreted this language to require market definition.¹⁶³ The statutory language, however, is manifestly amenable to another

¹⁶⁰ See, e.g., LAWRENCE A. SULLIVAN & WARREN S. GRIMES, *THE LAW OF ANTITRUST* 74 (2d ed. 2006) (“2.6d. A Showing of Anticompetitive Effect May Obviate the Need for Market Definition [section heading]. Antitrust focuses on abusive exercise of market power. The exercise of defining the market is merely a tool for determining if that has occurred; it should not be an end in itself. In *Eastman Kodak Co. v. Image Technical Services, Inc.*, 504 U.S. 451, 466–67, 477 (1992), the Supreme Court stressed that ‘legal presumptions that rest on formalistic distinctions rather than actual market realities are generally disfavored in antitrust law.’ Confronted with Kodak’s arguments that aftermarkets were not appropriate for measuring market power, the Court wrote: ‘It is clearly reasonable to infer that Kodak has market power to raise prices and drive out competition in aftermarkets, since respondents offer direct evidence that Kodak did so.’ Disputes about market definition, then, are of little consequence in the face of actual evidence of anticompetitive effects.”); see also AREEDA, HOVENKAMP & SOLOW, *supra* note 2, at 232–33 (although noting other reasons for the use of market definition, stating: “Finding the relevant market and its structure is typically not a goal in itself but a mechanism for considering the plausibility of antitrust claims that the defendants’ business conduct will create, enlarge, or prolong market power. . . . While the inferences to be drawn from market shares are limited, there may be no practical alternative when the data are insufficient to find the firm’s elasticities of supply and demand or its residual demand curves. If known, these numbers would measure market power directly, without any need for market definition. Because they so often lack such data, antitrust courts traditionally define a market and examine the firms’ market shares.” (footnotes omitted)).

¹⁶¹ Recall that, even if a market definition is required, it makes little difference if discussion of market power inferences need not be influenced much or at all by the market definition that is chosen; that is, the requirement may be treated as a mere formality.

¹⁶² 15 U.S.C. § 18 (2006). Importantly, mergers may also be attacked under Sherman Act section 1’s prohibition on “[e]very contract, combination . . . , or conspiracy . . . in restraint of trade or commerce,” *id.* § 1, so even if a doctrinal impediment exists under the Clayton Act, its significance is unclear. Note further that arguments similar to that in the text on Clayton Act section 7 can be applied to Sherman Act section 2’s language referring to monopolization of “any part of the trade or commerce among the several States, or with foreign nations.” *Id.* § 2. The responses in the text are applicable, although it should also be noted that the inference from the statutory language is even weaker with the Sherman Act (in part because the clause seems to refer to Congress’s authority to regulate interstate and foreign commerce), and the Supreme Court has for over a century failed to interpret the Act literally. See, e.g., *Standard Oil Co. v. United States*, 221 U.S. 1, 63–65 (1911) (rejecting the view that “every contract . . . in restraint of trade” applies to every contract that in some fashion restrains trade; instead the determination of which restraints are prohibited is governed by a rule of reason).

¹⁶³ See, e.g., AREEDA, HOVENKAMP & SOLOW, *supra* note 2, at 233 & n.4 (citing *Brown Shoe* as among the Supreme Court cases endorsing the need to define markets to satisfy Clayton Act section 7’s language, relying on the language quoted in the text to follow, without mentioning the opposite available interpretation).

interpretation, namely that the statute was intended to reach all anti-competitive mergers, wherever they may occur and whatever sorts of commerce they may affect. Indeed, the *Brown Shoe* Court itself seemed to embrace just this meaning when it stated: "Because § 7 . . . prohibits any merger which may substantially lessen competition 'in any line of commerce' (emphasis supplied [by the Court]), it is necessary to examine the effects of a merger in each such economically significant submarket If [any competition-lessening] probability is found to exist, the merger is proscribed."¹⁶⁴ Furthermore, if one considers a direct approach to market power that forsakes market definition, it is easy to meet any such requirement. Consider again the example of the orange juice merger. If direct inquiry indicates that significant price elevation is likely, then one would conclude that the line of commerce consisting of orange juice is affected in the requisite manner. If not, there is no anticompetitive effect to begin with, so the question of where it exists is moot.¹⁶⁵

In summary, it appears that existing law may not embody a significant constraint on the use of coherent, economically sound analysis of market power. If this view is excessively Panglossian, then the

¹⁶⁴ *Brown Shoe Co. v. United States*, 370 U.S. 294, 325 (1962). Moreover, earlier in the opinion, when discussing the legislative history of the statute, the Court stated that "Congress neither adopted nor rejected specifically any particular tests for measuring the relevant markets, either as defined in terms of product or in terms of geographic locus of competition, within which the anti-competitive effects of a merger were to be judged." *Id.* at 320-21. In discussing horizontal aspects of the merger, the Court elaborated:

Section 7 of the Clayton Act, prior to its amendment, focused upon this aspect of horizontal combinations by proscribing acquisitions which might result in a lessening of competition between the acquiring and the acquired companies. The 1950 amendments made plain Congress' intent that the validity of such combinations was to be gauged on a broader scale: their effect on competition generally in an economically significant market. Thus, again, the proper definition of the market is a "necessary predicate" to an examination of the competition that may be affected by the horizontal aspects of the merger.

Id. at 335 (footnote omitted) (quoting *United States v. E.I. du Pont de Nemours & Co.*, 353 U.S. 586, 593 (1957)). This language suggests that the motivation for defining markets was to encompass anticompetitive effects that might lie beyond those confined to the two merging parties rather than to limit the law's application to some subset of demonstrable anticompetitive effects.

¹⁶⁵ One might add the general points that formalistic interpretations of the antitrust law have long been rejected (in addition to the prior discussion in the text); that earlier decisions, including *Brown Shoe* in particular, are no longer followed along many dimensions; and that, for those inclined toward legislative fidelity in this realm, the legislative history of the 1950 enactment creating the current version of section 7 is largely at odds with the central thrust of modern merger law more broadly (and also fails to endorse economic markets understood in terms of horizontal competitors). See, e.g., Derek C. Bok, *Section 7 of the Clayton Act and the Merging of Law and Economics*, 74 HARV. L. REV. 226, 236-37 (1960) ("To anyone used to the preoccupation of professors and administrators with the economic consequences of monopoly power, the curious aspect of the debates is the paucity of remarks having to do with the effects of concentration on prices, innovation, distribution, and efficiency. To be sure, there were allusions to the need for preserving competition. But competition appeared to possess a strong socio-political connotation which centered on the virtues of the small entrepreneur to an extent seldom duplicated in economic literature.")

question becomes one of reform rather than interpretation. In any event, enforcement agencies through policy statements, guidelines, and internal operations may have substantial leeway to adjust existing practice.¹⁶⁶

VII. CONCLUSION

The market definition / market share paradigm plays a prominent role in competition law regimes. Its central justification is that it offers a useful means of making inferences about market power, indeed one that is easier or more reliable than other means of market power determination. Upon analysis, however, it appears that this widely accepted view is always false: it is difficult to render coherent the process of inferring market power from market shares in redefined markets, and, even if this obstacle is overcome, there is no way to define relevant markets in the first instance that does not presume the conclusion, specifically, that one already knows the answer to the market power question that the market definition process is designed to illuminate.

First, as Part II explains, market share is relevant to inferring market power in homogeneous goods markets, following a standard economic formula that relates market power to a dominant firm's or a group of firms' market share, the market elasticity of demand, and rivals' elasticity of supply. However, this method is valid only in the homogeneous goods market. Once markets are redefined to include imperfect substitutes, as is standard practice, one needs a new mechanism for translating market shares into market power. But it turns out that the only correct way to do so requires undoing the market redefinition, that is, returning to the homogeneous goods market. This result is not problematic, however, in that proper inferences can be made in that market even when, say, there are many close substitutes for the original product in question. Thus, there is never any occasion for choosing among markets.

Second, as Part III shows, the notion of making market power inferences from market shares presupposes some sort of benchmark for market share interpretation, which prior analysts have not usually appreciated. Although there exist ways to construct a standard reference market to serve as the needed reference point, the task has not actually

¹⁶⁶ Indeed, much of current market definition practice in U.S. courts is attributable to the Horizontal Merger Guidelines. See Hillary Greene, *Guideline Institutionalization: The Role of Merger Guidelines in Antitrust Discourse*, 48 WM. & MARY L. REV. 771 (2006). They have been reformed in 2010 in a manner that reduces the emphasis on traditional market definition, see sources cited *supra* note 5, and could be relaxed further.

been undertaken, and this lacuna highlights a further deficiency in the standard paradigm.

Most problematic, however, is the criticism presented in Part IV: Under any plausible criterion for deeming one market definition to be better than another — the central question in any market definition analysis — the only way of applying the criterion presupposes that one has already formulated a best estimate of market power. However, since the only purpose of the market definition inquiry is to aid in making inferences about market power, the entire procedure is revealed to be pointless. In particular, choosing the relevant market adds no information or analysis that can in any way improve on the market power estimate with which one begins. Worse, the most natural criterion, which also seems closest to what most analysts have in mind — choosing that market which yields the most accurate measure of market power — actually discards information and thus sometimes leads to mistaken conclusions. This defect can be avoided by substituting a patently circular method under which the market definition process is transparently superfluous.

In addition to these inherent defects in the market definition / market share paradigm, Part V notes the further problem that it leads us to focus on particular substitutes, and their cross-elasticities with the product in question, rather than on the overall elasticity of demand, which is what is directly relevant to proper market power inferences. Accordingly, the standard approach uses a subset of available information and employs it in a manner that reduces the reliability of conclusions about market power. Ironically, this feature involves additional effort and requires information that tends to be more difficult to obtain. Accordingly, the market definition approach has another significant defect, whereas these added challenges simply do not arise if one abstains from engaging in market definition. Sometimes knowledge about certain cross-elasticities is useful — notably, in evaluating mergers between firms producing highly differentiated products — but the proper way to employ this information does not involve defining a relevant market.

Part VI addresses a number of additional topics, including the allure of market definition and whether the law requires market definition even if it makes no sense. Regarding the first of these subjects, it is suggested that either/or choices seem simpler than quantitative pronouncements, which may appear to be more difficult for an advocate or decisionmaker to defend. However, since the former decisions have been shown to be impossible to make coherently without first performing the latter, this attraction is illusory.

The law in various jurisdictions routinely employs and is often seen to demand use of the market definition / market share paradigm. Yet it is hardly clear that existing requirements impose much if any limit on innovation in the methods by which market power is established.

First, the best available information can be used to assess market power, which is logically required to determine which market definition is best (by any plausible criteria) in any event, and pronouncement of a market definition can be rendered a largely conclusory final step. Second, current understandings of the purpose of competition law (which emphasize economic substance over formalism), much case law, and an assessment of statutory requirements are all plausibly consistent with subordinating or even abandoning the conventional market definition process. Whether these channels will be increasingly pursued seems to depend more on whether academic commentators, enforcement agencies, experts, practicing lawyers, and courts come to view the market definition / market share paradigm as not merely imperfect but fundamentally defective.

At its core, however, this Article is not about the best interpretation of existing law. Nor is it about how, in practice, market power might best be measured in various industries or for different legal inquiries under competition law.¹⁶⁷ Rather, the claim is conceptual: the market definition / market share paradigm is incoherent. Among other reasons, there exists no way to employ it — which requires a determination of which market definition is best according to some plausible criterion — without first determining the extent of market power as best one can. Therefore, however hard it may be to measure market power in a particular case, and whatever means may be most suitable despite their limitations, there is no getting around the need to formulate a best estimate. Nothing in the market definition process can help. Relatedly, if legal authorities are to insist on market definition, they are in a sense commanding something impossible and, as generally conducted, needlessly misleading — or, if the process is properly corrected, devoid of content and thus merely vacuous.

¹⁶⁷ As explained in Part IV, the argument does not depend on the quality of one's best estimate of market power or the types of evidence from which it is derived.