HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

MARKET DEFINITION

Louis Kaplow

Discussion Paper No. 745

05/2013

Harvard Law School
Cambridge, MA 02138

Market Definition

Louis Kaplow*

Abstract

Market definition has long held a central place in competition law. This entry surveys recent analytical work that has called the market definition paradigm into question on a number of fronts: whether the process is feasible, whether market share threshold tests are coherent, whether the hypothetical monopolist test in merger guidelines is counterproductive, and whether and when the frequent focus on cross-elasticities is useful.

Forthcoming, in THE OXFORD HANDBOOK OF INTERNATIONAL ANTITRUST ECONOMICS (Roger D. Blair & D. Daniel Sokol eds., 2013)

The economics of market definition is a paradox – or, more precisely, an oxymoron. On one hand, market power is a central element of most competition rules throughout the world, and market definition is the most widely used method of assessing market power, taken by some to be mandatory in certain settings. On the other hand, the concept of market definition does not exist in industrial organization economics (Fisher 1987, 2008; Kaplow and Shapiro 2007; ten Kate and Niels 2009), and recent work suggests that it cannot be rationalized in a coherent, useful manner (Kaplow 2010, 2011a, 2011b).

Accordingly, this chapter's focus is conceptual.[1] Section 1 makes explicit what is necessary if market definition is to be rendered a useful tool in illuminating market power and explains how the predicates are lacking or circular. Section 2 explores market share threshold tests, revealing how they conflate empirical questions and legal policy judgments. Section 3 specifically examines merger guidelines and the hypothetical monopolist test, showing how the method is inappropriate in principle and misleading in practice even under the best of conditions. Section 4 addresses problems arising from the common focus on specific substitutes and their cross-elasticities, which practice is induced by the felt need to define a market.

1. On the (In)Coherence of the Market Definition Process

Under most competition regimes, market power or changes therein play a central role.[2] Exclusionary practices – e.g., abuse of dominance in the European Union and monopolization in the United States – require a showing of significant market power as well as anticompetitive acts. Horizontal mergers are condemned when they sufficiently augment market power. And many horizontal arrangements – short of naked price fixing and related behavior – are carefully scrutinized and ultimately prohibited only if nontrivial market power is present.

Taking as given the critical importance of market power, and the fairly widespread understanding that it refers to the ability to profitably elevate price, the question remains how best to measure it. The most common method in competition law involves defining the so-called relevant market and then making market power inferences from market shares in that market. This section presents two critical flaws that make the market definition / redefinition process useless or worse (Kaplow 2010, 2011a).

---

[1]There exists substantial empirical work on the measurement of market power. However, given that market definition is not a conceptually coherent economic concept, there cannot really exist empirical work on market definition per se (and that which exists, not surprisingly, most often is really about market power itself). Some economists have suggested to me that market definition is perhaps a useful metaphor for guiding data collection, such as in determining which related products to include. This use of the notion, however, is qualitatively different from the use in competition law that is the focus here. Moreover, the analytical connection is not very close; as the analysis in section 1 implies, it would be difficult to use an incoherent means to guide empirical investigation, much less one that needs the output of such analysis as an input in applying the method.

[2]As the text indicates, sometimes the market power inquiry is addressed to levels and other times to changes (although changes themselves can be expressed in terms of two levels, before and after, say, a merger). For convenience, the language of levels will most often be employed, although section 3 on horizontal mergers focuses on changes. For further discussion, see Kaplow and Shapiro (2007).

1.1. Market Redefinition and Market Share

To begin, it is necessary to ask how market definition and market shares relate to market power, which is defined as the degree to which price (P) may profitably be elevated above a competitive level, where price equals marginal cost (MC). In a direct sense, it is not. However, measuring market power directly requires assessing marginal cost, which often is difficult in light of the need to allocate joint costs and to determine which costs are fixed versus variable in the pertinent time frame. Still, the difficulty of measuring a parameter does not ordinarily lead economists to measure something altogether different.

The answer has to be found in models that in which one can derive a relationship between market share and market power. The most familiar is one in which there is a dominant firm facing a competitive group of fringe firms in a market for a homogeneous good. (A variant involving quantity competition in a homogeneous goods markets is examined in subsection 3.1.) The Lerner Index (L) in this case is determined by the formula (Stigler 1940; Landes and Posner 1981; Kaplow and Shapiro 2007):

$$L = \frac{P - MC}{P} = \frac{1}{|\varepsilon_f|} = \frac{S}{|\varepsilon_d| + (1-S)\varepsilon_r}.$$

In this formula, the firm's elasticity of demand is $\varepsilon_f = (dQ/dP)(P/Q)$, where Q denotes the quantity of the firm's output; the market elasticity of demand is $\varepsilon_d = (dX/dP)(P/X)$, where X denotes total market demand; and the rivals' collective elasticity of supply is $\varepsilon_r = (dY/dP)(P/Y)$, where Y denotes rivals' total supply. Finally, S is the dominant firm's market share, and hence $1-S$ indicates rivals' aggregate share.

The first equality defines the Lerner index: the fraction of the price that is in excess of marginal cost. The next derives from the dominant firm's profit-maximizing calculus: a higher price raises profits on retained sales but sacrifices profits on forgone sales. The optimal elevation is given by the inverse of the firm's elasticity of demand, which indicates the rate at which the firm loses sales in all forms. The final equality decomposes the firm's demand elasticity into the sum of the market elasticity and fringe firms' supply elasticity, the latter weighted by their share. This summation reflects that sales may be lost to other products or to other suppliers of the same product. The intuitive relevance of market share to market power can be seen in this rightmost term: the S in the numerator is due to the fact that a given price increase (caused by a quantity reduction) is profitable to the dominant firm in proportion to its share of industry sales, and the $1-S$ in the denominator indicates that a higher dominant firm share makes the rivals' share lower and hence reduces the impact of a given supply response (the elasticity measuring that response in percentage terms).

Using the dominant firm's share in the homogeneous products market, we can therefore determine its market power if we can ascertain the market elasticity of demand and the fringe's elasticity of supply (or, with some techniques, Baker and Bresnahan 1988, we might simply eschew all this and measure the firm's demand elasticity directly). It should be emphasized that this is so regardless of how many substitutes exist for the firm's product or how strong those substitutes are. In other words, this formula gives us the answer we seek without having to undertake any of the standard market definition analysis. (This statement, in terms of product market definition, has analogues for

geographic market definition, which might be viewed as pertaining to the rivals' supply response.)

Having concluded that market redefinition is unnecessary, suppose next that we do it anyway. That is, suppose that there are some substitutes deemed to be sufficiently close that we should broaden the market to include them. If this step is undertaken, we then need a formula – a new formula – to infer market power in this now-heterogeneous goods market. But what is that formula?

The short answer is that there is none. The above formula is valid only for markets with homogeneous goods. It has long been supposed that perhaps one could apply an analogue to that formula for the broadened market, making suitable interpretations of the elasticities in the rightmost term – the market share being fairly straightforward, and lower, supposing that the dominant firm does not also produce the now-included substitute products. This strategy is pursued in Landes and Posner (1981), with criticisms advanced, for example, in Schmalensee (1982) and Kaplow (1982).

The problem is that these common suggestions amount to little more than hand-waiving. For the reinterpreted elasticities to be grounded, there needs to be a model that generates a formula that enables us to perform the necessary translation. As shown in Kaplow (2010), there exists a unique way to reinterpret these elasticities, which involves scaling them down to precisely offset the effects of the lower S in the formula. Put more bluntly, the only correct method involves undoing the market redefinition. This can be seen, for example, by taking the case in which the rivals' supply elasticity is zero. In that event, if the redefinition reduces the share by, say, 72%, one would simply deem the market elasticity of demand in the combined market to be 72% less than the that in the original, homogeneous goods market. Obviously, this fudge factor gives us the correct answer, and any other adjustment will not. But if this method is the only valid way to interpret elasticities in the combined market, there obviously is no point to ever redefining markets.

1.2. Criterion for Market Definition

To choose the best market definition presupposes a criterion for ranking alternatives. For concreteness, suppose that in some case the choice has come down to two definitions, which we will call *Narrow* and *Broad*. (More generally, whatever method is used to determine which of these two is superior would be used to eliminate the other alternatives.) Now, whichever of these two market definitions are chosen, we will then have to make a market power inference from the share in the chosen market. But we have just learned that such is not really possible unless we have stayed with the homogeneous goods market.

For present purposes, set aside that obstacle and suppose that we do have some means of making a market power inference conditional on the market definition selected. Presumably, we infer a higher level of market power from the larger share in *Narrow* than we infer from the lower share in *Broad*. Moreover, to make the choice of interest, we will suppose that the truth lies somewhere between these two inferences. That is, if we choose *Narrow*, we expect that we will overstate market power at least somewhat, and if we choose *Broad*, we will understate power. Given that some error is inevitable in this choice, which one is best?

4

As discussed in Kaplow (2010), the literature is essentially silent on the matter of the criterion, a remarkable observation in light of routine practice by economists and adjudicators who have purported to define thousands of markets in the last half century. The most plausible criterion that seems implicit in prior discussion is that the best market definition is that which involves the smallest error – that is, the smallest gap between the market power we will infer conditional on the chosen market and the truth.[3] And, since we do not actually know the truth, the benchmark presumably is our best estimate (or, in many settings, guesstimate might better depict the situation).

A moment's reflection on this criterion and the methodology necessary to implement it reveals an overwhelming problem. We cannot determine which market definition involves a smaller error without first measuring the error from each possible choice. To measure the error from either market definition involves determining the (absolute value of the) difference between the market power we would infer conditional on choosing that definition and our best estimate of market power.[4] Hence, the first step in applying the criterion requires the formulation of our best estimate of market power. But if that is so, why are we not finished at that point?

After all, the purpose of the market definition exercise is to enable an inference about market power. Yet, if our method of choosing the market definition presupposes that we already have a best estimate of market power, there is no further work to be done. Now, that best estimate may sometimes be a very good one and other times it may be rather feeble. In either case, our best estimate is still our best estimate.

The market definition process, however, is worse than pointless. It was noted just above that some error is inevitable in our choice of markets. But that error is not inevitable – it is entirely avoided – if we refuse to choose one of the two markets and just stick with our initial best estimate. After all, if we determined that *Narrow* was the best market definition – which is to say that the market power inference in that market overstates market power (relative to our best estimate) by less than the market power inference in *Broad* understated market power – we are still, knowingly, overstating market power. Likewise, if *Broad* is the best market definition, we know we are understating market power. Both choices involve error, in addition to the extra effort required to complete the market definition task.

This error in market power inferences will sometimes translate into error in legal outcomes. For example, when we choose *Narrow* as best and thus overstate market power, we will sometimes prohibit a practice where the result would be exoneration if we had instead employed our best estimate. Likewise, even when *Broad* is superior, we will sometimes allow practices that should have been condemned.

There does exist a solution of sorts: one can make market definition the conclusion of the legal inquiry rather than an intermediate step. That is, one can use one's best estimate of market power to determine whether it makes sense, say, to disallow a merger or penalize the use of some alleged exclusionary practice. Once that is decided,

---

[3]The implicit use of a linear, symmetric loss function is merely for expositional convenience.

[4]Throughout this discussion, the basis for the best estimate (guesstimate) is not addressed. Note that, depending on the context (e.g., initial screening versus adjudication) and available data, different techniques may be appropriate, and in all setting it may be relevant to consider a variety of information, including not only formal econometric analysis but also, for example, impressions of pertinent industry participants.

one could, after the fact, choose the market definition that ratifies the decision. For example, if the decision is to condemn, one would choose *Narrow*.

This patently circular method is superfluous, properly understood. But it does have important virtues. First, it does get the best answer, which the existing method may not. Second, for those agencies, economists, lawyers, or judges operating in a jurisdiction that is understood to require market definition,[5] this tactic provides an out – a way to avoid error while formally adhering to the requirement. And we suspect that this way of approaching the problem is at least implicit in some actual practice. For example, some view the government's strategy and the court's opinion in the *Staples* merger case in this manner.[6] Finally, if pushed to define a market – a theoretically incoherent enterprise – there is not a better way to proceed.

In sum, the market redefinition process entails unnecessary work and produces outcomes that are inferior to those arising when it is eschewed entirely. This finding is, in a sense, convenient in light of the earlier point that there does not exist any valid way to make market power inferences in redefined markets anyhow – that is, unless one always undoes market redefinitions and reverts to the homogeneous goods market.[7]

2. Market Share Thresholds

Some competition law statutes, many government guidelines and policy statements, prominent court decisions, and legal treatises articulate market power requirements in terms of market shares. For example, for abuse of dominance in the European Union or monopolization in the United States, it might be stated that the dominant firm ordinarily must have a market share of at least 50%. And merger guidelines typically offer various market share thresholds – often in terms of HHIs, the sum of the squares of market shares – for when challenges have varying degrees of likelihood, including the case of essentially no likelihood (safe harbors).

This standard practice is highly problematic. Conventional wisdom emphasizes that, for any such statements to be meaningful, one must look at the market shares in the relevant (best) market. However, the foregoing analysis shows this requirement to be impossible to satisfy in an economically sensible way. Moreover, even in what is regarded to be a properly defined market, Landes and Posner (1981), among others, have

---

[5]Kaplow (2010) points out that the basis for viewing market definition as legally mandatory is far weaker in the United States than is commonly supposed. For example, courts (including the Supreme Court) have explicitly stated that market definition is a means to an end (which is determining market power), that direct means may be employed, and that when market definition is indeed used, the market shares must be interpreted in light of relevant economic considerations. Nor is it clear that market definition is compelled by the merger statute, Clayton Act §7, either on its face or as interpreted by the Supreme Court. In addition, mainstream legal commentators seem increasingly willing to forgo market definition, at least for mergers in markets with differentiated products (e.g., Areeda and Hovenkamp 2009, pp. 84–88).

[6]FTC v. Staples, Inc., 970 F. Supp. 1066 (D.D.C. 1997).

[7]This logic applies as well to attempts to use market definition as a quick screen. For example, one still needs to define a market, and such presupposes some market power guesstimate, which might be employed directly. It is sometimes suggested that a fairly broad market definition, in which resulting shares are trivial, is obviously correct. But this idea is tantamount to stating that the market demand or fringe supply elasticity in the narrow market is obviously extremely high, which directly implies that market power is low.

shown how a given share can imply a very wide range of market power levels because the elasticities in the above formula are not limited to single values or even a confined set of values.

Kaplow (2011b) elaborates on the serious difficulties created for competition law. Everyone recognizes that market shares in a given market setting need to be interpreted, and that the guidance entailed by market share threshold tests should be understood as presumptions or rules of thumb – not rigid, sharp decision rules. The difficulty is that this acknowledgement only tells us what not to do; it fails to provide an affirmative prescription. This gap is much worse than meets the eye, for it is quite unclear what in principle is the sign or direction of the proper adjustments, much less their magnitude.

For example, consider a 50% presumptive required share for abuse of dominance. In a given case, we might know that the actual necessary share may be notably higher, perhaps 60% or 70%, whereas there might exist other cases where the requisite market power is implied by a share of 45% or 40%. Statements to this effect are ubiquitous, but it is extremely difficult to make them operational, for they all presume a benchmark that remains unspecified and is quite mysterious.

To make sense of the preceding example, key hurdles must be overcome. First and foremost, one would need to know how much market power is taken to be conveyed by the 50% market share contained in the presumptive threshold.[8] But the market power tests are not stated in terms of market power levels, and instead usually only in terms of market shares. We might impute some translation between market power and market share – what Kaplow (2010, 2011b) refers to a standard reference market translation table – but the basis for doing so in practice is unclear. Perhaps the promulgators of the threshold had in mind the textbook "market." But there is no such thing. Perhaps they meant the typical or average situation. But where do we suppose such figures could have come from? They would have had to specify a sample of markets (whatever that means), then engage in market definition since all agree that the shares are only meaningful in so-called relevant markets (but section 1 explains that such is infeasible), and finally determine the market power associated with different shares in each such market, even though in the actual markets only one situation prevailed.

Only if these obstacles are surmounted can one know how to proceed in a given case. After all, if a firm under scrutiny argues that its 55% share is insufficient, because it conveys atypically low market power for such a share, we can only give meaning to its claim and assess its validity if we know what is the typical market power conveyed by a 55% share – and the 50% share in our presumptive threshold. Note further that, as part of this effort, we would need to determine how much power this firm has, at which point (echoing section 1), one should naturally wonder why we are not finished – that is, once we know its actual market power, why should we be asking what market power is typically conveyed by a 55% share, a 50% share, or any other share?

The present problem can usefully be illustrated, as in Kaplow (2011b), by considering Judge Hand's famous pronouncement in the *Alcoa* case that ninety percent "is enough to constitute a monopoly; it is doubtful whether sixty or sixty-four percent

---

[8]The text oversimplifies, for example, by supposing that the threshold would vary only with the market power associated with given market shares in different industries and not with the practices under scrutiny or other factors. In a sense, the illustration presents the easiest case for market share thresholds to work.

would be enough; and certainly thirty-three per cent is not."[9] Consider three of the infinity of possible meanings of this dictum in terms of market power:

Market Share and Market Power in *Alcoa*

Market Share	Market Power: Low	Market Power: Medium	Market Power: High
33%	1%	5%	25%
60/64%	3%	15%	50%
90%	10%	30%	75%

Starting with the "Medium" column, Hand might have thought that a market share of 33% would have implied that Alcoa could elevate price only 5% above a competitive level, which is certainly insufficient for the offense of monopolization; that a share of 60/64% would have implied a 15% elevation, which probably is not enough; and that a share of 90% implied a 30% elevation, which certainly is enough. In this case, the implicit market power legal threshold that rationalizes Hand's opinion is, say, in the neighborhood of a 20% minimum elevation.

But, looking at the "Low" column, Hand might have envisioned much smaller levels of market power being associated with the pertinent shares, with his implicit legal threshold perhaps closer to a 5% elevation. Or, examining the "High" column, he might have understood far greater market power levels to be involved with the stated shares, with the implicit legal threshold at 60% or so.

Consider what this ambiguity means for *Alcoa* as a precedent. In a subsequent case, we know we are to interpret market shares in context, and that a given market share may mean more or less than usual. Suppose that the share is 55%, as above. Under *Alcoa*, who wins on this issue? The short answer is that we have no idea. Now, add that further analysis reveals the share to be associated, in the present case, with a 20% price elevation. Then, if Hand had in mind the "Low" column, there is far more than enough market power; if he meant "High," there is not nearly enough; and if he meant "Medium," it is a close call.

This indeterminacy – and we have only considered three of the possible interpretations – raises insurmountable problems of application, whether for lawyers giving advice to businesses, agencies deciding when they have a plausible case, or adjudicators seeking to follow the precedent. This predicament seems insoluble, for there is no way to determine what Hand meant. One possibility would be to reanalyze the aluminum industry of over a half century ago to figure out the actual market power implications of these shares – no easy task and, in any case, whatever we might determine using modern methods has no necessary connection to what Hand was thinking, and the latter is relevant to interpreting the precedent. Likewise for other means of construction.

The foregoing, unfortunately, is not merely a nit-pick at Judge Hand. Just about any market share threshold in any statute, guideline, opinion, or text is subject to the same problems. (A partial exception is the hypothetical monopolist test in various

---

[9]United States v. Aluminum Co. of Am. (*Alcoa*), 148 F.2d 416, 424 (2d Cir. 1945).

8

merger guidelines, the subject of section 3.) Indeed, this was the problem at the outset in the example of a 50% presumptive market share threshold for abuse of dominance. By reference to the above table, such a threshold might implicitly refer to a minimum price elevation of roughly 2%, 10%, or 40% – or any other figure since those three columns hardly represent the only possibilities. Market share threshold tests are, on reflection, entirely uninformative about the implicit market power requirement.

   Moreover, they pose the aforementioned difficulty of how to determine whether the inference of market power in a given case involving, say, a 55% share should be adjusted at all in light of the circumstances, in which direction, and by how much. It seems we would have to know both how much actual market power exists in the present case and also how much market power is implicitly required. But once we knew all that, there would be no point in adjusting the 55% share upward or downward to then compare it with the 50% share in the test, for we would already know whether the market power in the case under consideration met the market power legal threshold.

   Just as market definition and redefinition has been a subject of extensive discussion and implementation for decades, yet we really cannot make economic sense of the notion, so too market shares and market share threshold tests have been at the center of competition law rules, debates, and applications in countless cases, but without any coherent basis. Regarding market shares, the problem can in part be viewed as one involving the possibility of communications. A market share threshold test is supposed to state to the world something about how much market power is required, but it does not. Frequent arguments that market shares in one or another case imply more or less market power than meets the eye assumes that there is some level of market power ordinarily associated with a given market share, which is not so. More fundamentally, it assumes that – however determined, even if by arbitrary convention – there is some degree of power that both speaker and listener have in mind, but there is not. Terms like "high" and "low" are inherently relative, and all market share discourse that is ultimately linked to market power presupposes a common benchmark, a translation table of sorts. Those who pronounce, criticize, and apply the law seem to be speaking the same language, all the while failing to communicate.[10]

   Another useful way to view the problem is to recognize that a focus on market shares conflates two qualitatively distinct questions (Kaplow 2011b): How much market power exists in a given case? And how much market power is necessary to meet the legal test? The first is an empirical question, the second a matter of policy. Moreover, market shares are not in the proper units to answer either question. The amount of market power in a given case refers to the degree of profitable price elevation. Stating that a firm has a 50% market share does not answer that question. Likewise, a market power legal threshold might indicate, say for horizontal mergers, how much of a

---

   [10]To consider further the degree of this failure, consider administering a survey to economists with significant engagement in competition policy or to members of competition agencies or courts. It might ask: "How much market power (measured by the extent to which a dominant firm profitably elevates price above marginal cost) is typically associated with a market share of 50%? Conversely, they could be asked: "What market share do you think is typically possessed by a dominant firm that can profitably elevate price by 20%?" One can ponder how widely such answers would vary, if those surveyed were willing to answer at all. Or one could ask them instead how they think others in the target groups would respond to these two questions. Again, contemplate the predicted mean and variance of the responses.

predicted price increase is necessary to trigger a challenge. Again, stating that the HHI must be at least 2500 and the increase due to the merger at least 200 does not answer the question. Obviously, it is impossible to respond to two distinct questions, one empirical and another about optimal policy, with a single answer. Market shares, however, answer neither one.

Applied to Hand's statement in *Alcoa*, we can ask whether his famous pronouncement meant to say something about how much market power existed in that case – which he had to determine to apply the legal test – or about how much market power is required in a monopolization case – which he needed to state in order to know whether the test was satisfied, and which must be taken to be his meaning if the case is to serve as a precedent. As just explained, and as illustrated above, he in fact answered neither. From his assertion, we have no idea how much market power Hand thought Alcoa possessed and we have no clue how much market power he deemed to be legally necessary.

3. Merger Guidelines and the Hypothetical Monopolist Test

Agency assessment of horizontal mergers is one of the most important competition law domains, and market power inquires are the key determinant of most decisions whether to block proposed mergers. Many jurisdictions have merger guidelines, often following the blueprint originally established in the United States involving application of the hypothetical monopolist test for market definition purposes. E.g., U.S. Horizontal Merger Guidelines (2010), EU Guidelines on the Assessment of Horizontal Mergers (2004), EU Commission Notice on the Definition of Relevant Market (1997), EU DG Competition Discussion Paper (2005).

In brief and rough terms, the method is as follows (with attention to product market definition for concreteness): One begins with the homogeneous goods market and asks whether a hypothetical monopolist in that market could profitably elevate price, say, 5%, above current levels. If so, that is the relevant market. If not, one adds the next group of substitutes and asks whether a hypothetical monopolist in that redefined, broader market could profitably elevate price 5%. If so, that is the relevant market. If not, one continues until the test is satisfied.

Then, various guidelines typically examine the post-merger HHI in the resulting market as well as the increment to the HHI due to the merger in order to determine safe harbors, ranges of likely challenge, and so forth. For example, under the U.S. Guidelines, a merger that raises the HHI by more than 200 points and results in an HHI above 2500 is presumed to be likely to enhance market power sufficiently to warrant a challenge.

Subject to various qualifications (such as how many substitutes to add at each stage, and in which order), the hypothetical monopolist test provides a determinate market definition. But we already know from subsection 1.2 that it must be problematic. If it were the best market definition, it would still be inferior to direct use of a best estimate of market power. And since it is not the best – for example, it does not ask which definition is closest to the truth – it is likely inferior.

Moreover, we should understand that use of the resulting market shares is problematic. First, whenever the test does not stop with the homogeneous goods market, there is a conceptual problem with making inferences in the broader market. Second, all

10

of the HHI thresholds are in the class of market share tests, which do not provide a basis for application in given cases since they are not denominated in market power terms.

In addition, we should be wary of the hypothetical monopolist test on another ground: Why are we applying a hypothetical *monopolist* test to a *merger*? Unless we are considering a merger to monopoly, the test is patently posing the wrong question.

We will now see how all these problems arise in concrete settings. To do so, it is useful to consider the three types of anticompetitive effects that have received the most emphasis in merger guidelines and commentary: unilateral effects in homogeneous goods markets, unilateral effects in differentiated products markets, and coordinated effects (Kaplow 2011a).

3.1. *Unilateral Effects: Homogeneous Goods*

Beginning with unilateral effects with homogeneous goods, the immediate question is why one should ever go beyond the first stage of the hypothetical monopolist test. If it fails in round one, we expand the market, at which point it is no longer one with homogeneous goods, and the pertinent model is inapplicable. To elaborate, one of the few contexts – aside from that with a dominant firm and competitive fringe, examined in subsection 1.1 – in which there is a model and resulting formula that uses market shares is the Cournot quantity-competition model with homogeneous goods. This formula indicates that the industry-wide average, output-weighted margin equals $\text{HHI}/|\varepsilon_d|$ (where the HHI is represented in ten-thousandths, so that its range is from 0 to 1). See, for example, Ordover, Sykes, and Willig (1982) and Kaplow and Shapiro (2007) – and we will ignore the important qualifications in Farrell and Shapiro (1990) because they are not pertinent for present purposes.

Now, if we really have Cournot quantity competition, this formula tells us that we can compute the price elevation as a function of the market shares (the HHI) *in the initial, homogeneous goods industry*. That is, if we stick with the original, homogeneous goods market, we can compute our price elevation: specifically, we can compare the elevation using the pre- and post-merger HHIs. But if we follow the hypothetical monopolist test when it tells us to redefine (broaden) the market, application of our formula is no longer valid and there is no other formula to use in its place. Hence, the standard merger guidelines method is the wrong one in this setting.

It is also the case that the standard method does a poor job even of generating a consistent ordering of possible mergers: that is, whether or not it indicates that a challenge is appropriate is quite poorly correlated with the predicted price elevation. To see this, consider under the U.S. Guidelines a merger in which the 5% hypothetical monopolist test is barely satisfied for the homogeneous goods market. Suppose further that the merger raises the HHI from 2300 to 2501, in which case it is deemed presumptively to pose a threat, as mentioned. The pre-merger elevation would roughly equal 0.23 divided by the market elasticity of demand, whereas a hypothetical monopolist, which we have supposed can raise price 5% over the prevailing level, has a Lerner index equal to 1.0 divided by that elasticity. These relationships are consistent with an elasticity of approximately 15, a pre-merger elevation of about 1.49%, and a post-merger elevation of approximately 1.62%, for a merger-induced elevation of 0.13%—that is, under two tenths of one percent

By contrast, if the hypothetical monopolist in the homogeneous goods market can only raise price 4.9%, we are instructed to expand the market. Now, if the first set of substitutes that we now include involves far more revenue than does the original market, then even a merger to monopoly in the original homogeneous goods market would not trigger scrutiny. But we had just stipulated that such a merger would raise price 4.9%. Let us now combine these two results: the Guidelines method would presumptively condemn the first merger, which raises price by 0.13%, and exonerate the second, which raises price by 4.9%, which is to say, over thirty-five times more than does the first merger.

This example provides two lessons. First, the hypothetical monopolist test is a poor means of attempting to assess the price effects of mergers. Other considerations, such as entry or efficiencies, may alter the outcomes. But to see if the test fundamentally makes sense, it is appropriate to consider a basic case in which such factors are held constant – even better, for simplicity, are irrelevant – and ask whether the Guidelines' hypothetical monopolist test works at all well in even the simplest setting, where no complications are present. Unfortunately, it does not.

Second, various merger guidelines' use of HHI criteria illustrate the point in section 2 that it is a mistake to denote market power tests using market share thresholds. This mismatch helps to explain how a given set of HHI ranges can generate such inconsistent outcomes, condemning a merger raising price by 0.13% while allowing one that raises price by 4.9%. Interestingly, merger guidelines provide elaborate discussions of agencies' methodologies in assessing price effects but do not ordinarily indicate what price effects would be sufficient to condemn a merger (Hausman and Sidak 2007). As a thought experiment, suppose that in a challenged merger, an agency's experts and the merging parties' experts simultaneously filed reports, and it turned out that both predicted precisely the same price increase – say, 3.1%, or perhaps 1.4%, or 0.2%. Who would win in each instance? Merger guidelines are silent. Indeed, if they are followed conscientiously, one might argue that direct conclusions on price effects are not decisive. After all, these guidelines purport to call for market definition using the hypothetical monopolist test and assessment based on HHIs in the market that is thereby selected.

The hypothetical monopolist test and the various market share thresholds indeed perform very poorly in homogeneous goods industries with quantity competition. But this problem is totally avoidable because the measure needed to apply the first round of the hypothetical monopolist test – the market elasticity of demand for the good in that market – is sufficient to directly assess the effect of the merger on price.[11] Contemplating market redefinition is entirely unnecessary and only leads the analysis astray.

3.2. Unilateral Effects: Differentiated Products

Now consider unilateral effects when products are differentiated. In this particular setting, it has become fairly accepted that market definition is not very useful, as developed in the literature on critical loss analysis and upward pricing pressure (e.g.,

---

[11]As mentioned, there are complications, which should be taken into account in practice but are orthogonal to the foundational problems with the hypothetical monopolist test.

Baker and Bresnahan 1985; Harris and Simons 1989; O'Brien and Wickelgren 2003; Farrell and Shapiro 2010). To assess a merger of two firms that, let us suppose, each produce a single product, we can predict price elevation by knowing the firms' margins on the two products and the diversion ratio between them. Simply put, after the merger, there is a stronger incentive to raise the price on each good because part of the lost sales are to the other product now controlled by the same entity. Because the cost of raising price is lessened, the motive to raise price increases.[12]

Importantly, determination of the relevant factors – the diversion ratios and the margins on the two products – does not require market definition. Accordingly, the hypothetical monopolist test and the guidelines' various HHI thresholds are beside the point. One could attempt to employ them, but they hardly would provide a reliable guide.

There is one way in which to avoid the problem: apply the hypothetical monopolist test to an initial market with only the two firms' products and no others. And then stop. The first round of the test indicates how much a hypothetical monopolist would raise price. If it is the hypothetical "monopolist" consisting of the two merged firms and no others, that is our answer. Of course, to apply this variation of the test is to inquire directly into how much the merger will elevate price. Not also that it is important to stop, contrary to the dictates of the hypothetical monopolist test. If one instead expands the market, the HHIs there, matched against the guidelines' thresholds, will not be very useful.

3.3. Coordinated Effects

Finally, consider the possibility that a merger will enhance the likelihood of coordinated price elevation. In this event, we likewise want to stick with the homogeneous goods market. Price coordination usually involves products that are fungible, or nearly so – as suggested by theory (coordination becomes far more complex as the dimensionality of the problem increases) and by evidence (most prosecuted price-fixing cases involve homogeneous intermediate goods: see Connor 2007, pp. 136–53; Harrington 2006, pp. 98–102; Hay and Kelley 1974, pp. 29–38). Therefore, to predict the price effects of the merger, we need to ascertain the degree to which the merger makes coordination more likely and how much successful coordination would increase price. The hypothetical monopolist test is at most applicable to the latter question, but again no redefinition is involved. In round one, we can ask how much a hypothetical monopolist of the homogeneous goods industry would elevate price above current levels. Once we have the answer, we are done, for our question is precisely how much would coordination – the many firms acting as a hypothetical monopolist of that market – raise price. We still need to weight our result in terms of the merger's increment to the probability and degree of success, but that is another matter. Again, observe that knowledge of how much a hypothetical monopolist of some broader, heterogeneous goods market could elevate price, and what are the HHIs before and after the merger in that market, is beside the point.[13]

---

[12]The analysis implicitly assumes that the products are substitutes.

[13]One might occasionally want to consider coordination across markets, if the added complexity might be surmountable. In such instances, however, we would wish to analyze the predicted price elevation achievable

Although the proper analysis in these three scenarios differs, some conclusions are common. First, redefining markets is inappropriate in all the cases. Second, the iterative hypothetical monopolist test in particular is incorrect in every instance. Third, the guidelines' HHI thresholds are not useful either.

This analysis of the hypothetical monopolist test and their standard use in merger guidelines serves two purposes. Most obviously, it suggests that current methods, if actually followed, are quite problematic and, relatedly, that if guidelines reforms deemphasize market definition, that move is sensible at least in principle. Second, the discussion reinforces the conclusions in the first two sections. It has long been understood that the market definition process is imperfect, yet most have found it useful in light of the difficulty of the alternatives. Its complete failure to help – and its consistent tendency to lead us astray – in the realm of horizontal mergers, where the method has received the most attention and refinement, signals that the broader conceptual problems examined in the first two sections are indeed insuperable.

4. Elasticities versus Cross-Elasticities

In addition to the aforementioned defects with the market definition approach to making market power inferences, it is also useful to consider a problem that has been long understood (e.g., Kaplow 1982, Schmalensee 1982, Simons and Williams 1993, Werden 1998) but whose severity is not fully appreciated. When it is thought necessary to define a market, boundaries must be drawn; hence, particular products are either in or out, an all-or-nothing choice.[14] Furthermore, it is common to determine whether a particular substitute is sufficiently close so as to be included by examining its cross-elasticity (or, in the language of many legal materials, interchangeability) with the central product of concern.

Substitutes' cross-elasticities do not, however, appear in the market power formula in subsection 1.1. Their relevance is due to their bearing on the magnitude of the market elasticity of demand, the first term in the denominator on the right side of that expression. The precise relationship is as follows:

$$|\varepsilon_d| \equiv |\varepsilon_{11}| = 1 + \sum_{i=2}^{N} \frac{R_i}{R_1} \varepsilon_{i1}.$$

On the left side we have the market elasticity of demand, which for notational clarity in this context is helpfully restated as the elasticity of demand for the initial product (designated as product 1) with respect to a change in its own price. This own-price elasticity, in turn, equals 1 plus the revenue-weighted sum of the cross-elasticities of demand with all of the other N products.[15] The revenue weight for a product i (R_i) is the

---

by those groupings of firms for which coordination was determined to be feasible – not by whatever grouping happened to be selected by the hypothetical monopolist test.

[14]Again, analogous reasoning is applicable to geographic market definition.

[15]The explanation for the 1-plus term is that if, say, the demand for none of the other products changed when the price of product 1 rises, then the elasticity of demand for that product would be 1, not 0: since the same revenue continues to be spent on product 1, a one percent increase in price implies a one percent fall in quantity (literally so for infinitesimal changes).

ratio of expenditures on that product to expenditures on our initial product, and the cross-elasticity (ε_{i1}) considered here is the percentage change in the quantity of product i as the price of product 1 is increased.

Although substitutes' cross-elasticities are relevant to the market elasticity of demand in our original market, there are a number of reasons that we should not ordinarily focus on cross-elasticities (Kaplow 2010). First, information on the overall demand elasticity is what ultimately matters, it subsumes the cross-elasticities for all other products, and such information is often as easy or easier to obtain and as or more reliable than information on particular cross-elasticities.

Second, focus on specific cross-elasticities is incomplete in that really the cross-elasticities of all other products are relevant. Consideration of only the closest substitutes could omit a great deal. One might, for example, in one case include a single, close substitute and in another case omit a dozen more distant substitutes even if the latter collectively are more powerful. For example, if revenue weights were equal and the cross-elasticities of the dozen were each one-fourth that of the single, close substitute, then the dozen together would have three times the power of the single, close substitute. In some instances there may not be a problem if the inclusion of close substitutes is employed merely to form a lower bound on the market elasticity of demand, which if sufficiently large would rule out anticompetitive concerns.[16]

Third, inclusion implies overstatement of the restraining force of the substitutes because even close substitutes generally are not perfect substitutes. This point raises the need to interpret market shares in the redefined market, which subsection 1.1 indicates is possible only by undoing the redefinition and returning to the homogeneous goods market.

Fourth, the importance of the revenue weights is often implicitly ignored. Cross-elasticities indicate percentage changes in demand for a substitute, but the importance of a given percentage change depends on the base: hence the revenue weights. For example, a substitute with a cross-elasticity of 5 but a revenue weight of 0.1 has only half the impact of a substitute with a cross-elasticity of 1 but a revenue weight of 1.0. One type of mistake is that we might be led to exclude a product that has a fairly low cross-elasticity when its large revenue weight indicates that it imposes a strong restraint on price increases of the product in question. On the other hand, when we include products with large revenue weights, shares drop dramatically, which can lead to substantial understatement of market power, as illustrated by the second part of the unilateral effects example in subsection 3.1. Moreover, the magnitude of the resulting errors can be huge: the revenue weights in the example differ by a factor of ten, but when one contemplates the range of goods and services in an economy, these weights can differ by many orders of magnitude.

All these difficulties are avoided by focusing on the market elasticity of demand in the original market rather than attending to the cross-elasticities of demand for various particular substitutes. Observe that there is a close relationship between the market definition approach and the mistaken focus on cross-elasticities. If one thinks it necessary to define a market, and if that process is understood as a determination of which goods are inside the market and which are outside, then it seems necessary to

---

[16]If some of the omitted products were strong complements, one would not obtain a lower bound.

inquire whether or not each plausible substitute is sufficiently close, which in turn leads to an inquiry into its cross-elasticity. In contrast, if the direct question of market power is kept front and center in the inquiry, then one is led to ask about the market elasticity of demand rather than to focus on particular cross-elasticities.

The foregoing does not, however, imply that cross-elasticities are always irrelevant. Initially, there is the point at the outset that the market elasticity is an aggregate of the cross-elasticities, and sometimes estimating some key cross-elasticities – and applying the revenue weights – will enable a useful lower bound estimate of the overall demand elasticity. Additionally, in certain settings particular cross-elasticities are useful. In predicting the price effects of differentiated products mergers, as discussed in subsection 3.2, one needs to know the diversion ratio between the merging firms' products, which in turn is indicated by the (revenue-weighted) cross-elasticities (e.g., Farrell and Shapiro 2010). Relatedly, if an alleged exclusionary practice is directed at particular substitutes, its potential impact is similarly determined (Kaplow and Shapiro 2007). Therefore, cross-elasticities sometimes are useful, but none of these valid applications involve market definition.

5. Conclusion

The market definition process is incoherent. In a redefined market, there is no valid way to infer market power from the market shares therein – that is, unless one reverses the market redefinition and reverts to the homogeneous goods market. And there is no sensible way to deem one market definition superior to another without saying which definition entails less expected error, and error can only be defined by reference to some best estimate of market power. Hence, one needs a best estimate of market power to define a market, whereas the whole point of the market definition exercise is to facilitate an inference of market power. Moreover, whatever inference one draws from the market shares in the market that is chosen differs from the best estimate that was needed to choose the market and hence can only lead to worse legal outcomes.

The market definition approach is also associated with stating market power tests in terms of market share thresholds. This practice as well is conceptually flawed. Since even in a so-called relevant market any given market share can be associated with greatly differing levels of market power, market share thresholds constitute indeterminate legal tests.

The hypothetical monopolist test for market definition and the HHI thresholds in merger guidelines reflect these deficiencies. In none of the basic settings – unilateral effects with homogeneous goods, unilateral effects with differentiated products, and coordinated effects – does the hypothetical monopolist test produce the desired result. In the first and third case, one best predicts price effects of mergers by sticking with the homogeneous goods market, regardless of what the hypothetical monopolist test dictates. For differentiated products, one focuses on the diversion between the products in question, which involves no market definition at all (unless one stipulates the market to be that consisting of the products covered by the merger, again without regard to the outcome of the hypothetical monopolist test). This test, and the guidelines' HHI ranges, fail to indicate which mergers involve the greatest price effects even in the most basic settings.

The market definition process tends to produce a focus on particular substitutes and their cross-elasticities with the product in question, which emphasis is a by-product of the deemed need to determine which products are inside and which are outside the relevant market. But it is the overall market elasticity of demand that determines market power; particular products provide only a partial picture; inclusion of inevitably imperfect substitutes can lead to the understatement of market power; and ignoring the widely varying revenue weights can produce highly misleading market power estimates. Cross-elasticities are sometimes important to consider, but not to define markets.

This conceptual examination of the economics of market definition reveals its bankruptcy. The implication is that growing research in recent decades on direct means of inferring market power (e.g., Baker and Bresnahan 1988, 1992; Werden and Froeb 2008; Whinston 2006) should receive increasing attention and further development.

References

Areeda, P. E., & Hovenkamp, H. (2009). *Antitrust Law*, vol. 4, 3rd ed. Austin, TX: Wolters Kluwer.

Baker, J. B., & Bresnahan, T. F. (1985). The gains from merger or collusion in product-differentiated industries. *Journal of Industrial Economics, 33,* 427–444.

Baker, J. B., & Bresnahan, T. F. (1988). Estimating the residual demand curve facing a single firm. *International Journal of Industrial Organization, 6,* 283–300.

Baker, J. B., & Bresnahan, T. F. (1992). Empirical methods of identifying and measuring market power. *Antitrust Law Journal, 61,* 3–16.

Connor, J. M. (2007). Price-fixing overcharges: legal and economic evidence. *Research in Law and Economics, 22*, 59–153.

European Union. (1997). *Commission notice on the definition of relevant market for purposes of community competition law*. O.J. (C 372).

European Union. (2004). *Guidelines on the assessment of horizontal mergers under the council regulation on the control of concentrations between undertakings*. O.J. (C 31) 5–18.

European Union. (2005). *DG Competition discussion paper on the application of Article 82 of the treaty to exclusionary abuses*.

Farrell, J., & Shapiro, C. (1990). Horizontal mergers: an equilibrium analysis. *American Economic Review, 80*, 107–126.

Farrell, J., & Shapiro, C. (2010). Antitrust evaluation of horizontal mergers: an economic alternative to market definition. *The B.E. Journal of Theoretical Economics, 10*(1) (Policies and Perspectives), Article 9.

Fisher, F. M. (1987). Horizontal mergers: triage and treatment. *Journal of Economic Perspectives, 1*(2), 23–40.

Fisher, F. M. (2008). Economic analysis and "bright-line" tests. *Journal of Competition Law & Economics, 4,* 129–153.

Harrington, J. E., Jr. (2006). How do cartels operate? *Foundations & Trends in Microeconomics, 2*(1).

Harris, B. C., & Simons, J. J. (1989). Focusing market definition: how much substitution is necessary? *Research in Law and Economics, 12*, 207–226.

Hausman, J. A. & Sidak, J. G. (2007). Evaluating market power using competitive benchmark prices instead of the Herfindahl-Hirschman index. *Antitrust Law Journal, 74,* 387–407.

Hay, G. A., & Kelley, D. (1974). An empirical survey of price-fixing conspiracies. *Journal of Law and Economics, 17*, 13–38.

Kaplow, L. (1982). The accuracy of traditional market power analysis and a direct adjustment alternative. *Harvard Law Review, 95*, 1817–1848.

Kaplow, L. (2010). Why (ever) define markets? *Harvard Law Review, 124*, 437–517.

Kaplow, L. (2011a). Market definition and the merger guidelines. *Review of Industrial Organization, 39,* 107–125.

Kaplow, L. (2011b). Market share thresholds: on the conflation of empirical assessments and legal policy judgments. *Journal of Competition Law & Economics, 7*, 243–276.

Kaplow, L., & Shapiro, C. (2007). Antitrust. In Polinsky, A. M., & Shavell, S., eds., *Handbook of Law and Economics 2,* 1073–1225. Amsterdam: Elsevier.

Landes, W. M., & Posner, R. A. (1981). Market power in antitrust cases. *Harvard Law Review, 94,* 937–996.

O'Brien, D. P., & Wickelgren, A. L. (2003). A critical analysis of critical loss analysis. *Antitrust Law Journal, 71*, 161–184.

Ordover, J. A., Sykes, A. O., & Willig, R. D. (1982). Herfindahl concentration, rivalry, and mergers. *Harvard Law Review, 95*, 1857–1874.

Schmalensee, R. (1982). Another look at market power. *Harvard Law Review, 95*, 1789–1816.

Simons, J. J., & M. A. Williams. (1993). The renaissance of market definition. *Antitrust Bulletin 38,* 799–857.

Stigler, G. J. (1940). Notes on the theory of duopoly. *Journal of Political Economy, 48*, 521–541.

ten Kate, A., & Niels, G. (2009). The relevant market: a concept still in search of a definition. *Journal of Competition Law & Economics, 5,* 297–333.

U.S. Department of Justice & Federal Trade Commission. (2010). *Horizontal merger guidelines*.

Werden, G. J. (1998). Demand elasticities in antitrust analysis. *Antitrust Law Journal, 66,* 363–414.

Werden, G. J., & Froeb, L. M. (2008). Unilateral competitive effects of horizontal mergers. In Buccirossi, P., ed., *Handbook of Antitrust Economics*, 43–104. Cambridge, MA: MIT Press.

Whinston, M. D. (2006). *Lectures on Antitrust Economics*. Cambridge, MA: MIT Press.