Errors in Antitrust Enforcement Matter More than You Think

*Dennis W. Carlton**

Abstract: This Article will trace out the implications of how preexisting market distortions affect the magnitude of the harm from making an error in antitrust enforcement. This Article will apply well-accepted logic from cost-benefit analysis to show that pro-competitive mergers—or indeed any pro-competitive activities—produce benefits that could be substantially underestimated if one were to ignore the presence of preexisting distortions. Conversely, anti-competitive activities produce harms that could be substantially underestimated by an analysis that ignores preexisting distortions. These observations are especially relevant today as recent research has shown that the level of market power at the product level has increased, perhaps substantially, since the 1980s. On the buying side, there has been a growing recognition of the importance of monopsony power, especially in the labor market. These observations suggest that an analysis of mergers or other conduct must consider the costs and benefits of actions in a world with preexisting distortions. The conclusion is that antitrust matters more than one might otherwise think and therefore there are large costs to abandoning or de-emphasizing economic principles, as some have suggested, since the use of economics should lead to fewer errors.

* David McDaniel Keller Professor of Economics Emeritus, University of Chicago; Research Associate, National Bureau of Economic Research (NBER); and Senior Managing Director, Compass Lexecon. I thank Joseph Farrell, Ken Heyer, Louis Kaplow, Daniel O'Brien, Gregory Pelnar, Sam Peltzman, Steven Salop, Monica Stephens, Theresa Sullivan, and Ralph Winter for helpful comments.

Introduction

At least since Professor Oliver Williamson's famous American Economic Review 1968 article, economists have used cost-benefit analysis to measure gains and losses from various types of competitive conduct.[1] Williamson used an example of a merger to illustrate tradeoffs between the benefits from merger-specific efficiencies and the harms from anti-competitive conduct arising from the merger.[2] His analysis compares the ex-post world to an initial world with no market power.[3] It is common in cost-benefit analysis to measure the costs and benefits of an action in comparison with the existing world, which typically has preexisting market distortions (i.e., a gap between price and marginal cost).[4] For example, these market distortions can arise from the presence of market power or the existence of taxes.[5] It is well-known that taking into account preexisting distortions when analyzing firm conduct can alter the calculation of additional costs or benefits from some action such as a merger.[6] Williamson was well-aware that his analysis would need to be adjusted if the initial world already had existing distortions from market power.[7] A few subsequent authors have pursued how an initial antitrust analysis would change in the presence of existing distortions.[8]

This Article traces out the implications of how preexisting market distortions affect the magnitude of the harm from making an error in antitrust enforcement. This Article applies well-accepted logic from cost-benefit analysis to show that pro-competitive mergers—or indeed any pro-competitive activities—produce benefits that could be substantially underestimated if one were to ignore the presence of preexisting distortions. Conversely, an anti-competitive activity produces harm that

1 Oliver E. Williamson, *Economies as an Antitrust Defense: The Welfare Tradeoffs*, 58 AM. ECON. REV. 18, 25 (1968).

2 *Id.* at 26–28.

3 *Id.* at 25–26.

4 *See* Arnold C. Harberger, *Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay*, 9 J. ECON. LITERATURE 785, 790 fig.2 & n.5 (1971).

5 *Id.* at n.5; *see also* Williamson, *supra* note 1, at 28; Oliver E. Williamson, *Economies as an Antitrust Defense Revisited*, 125 U. PA. L. REV. 699, 711–12 (1977) (listing examples of qualifications to consider in a "naïve model").

6 Harberger, *supra* note 4, at 790; *see* R. G. Lipsey & Kelvin Lancaster, *The General Theory of Second Best*, 24 REV. ECON. STUDS. 11, 31 (1956).

7 *See* Williamson, *supra* note 1, at 712.

8 *See, e.g.*, Thomas W. Ross & Ralph A. Winter, *The Efficiency Defense in Merger Law: Economic Foundations and Recent Canadian Developments*, 72 ANTITRUST L.J. 471, 495–96 (2005) (examining how Canadian authorities reached an incorrect conclusion because they failed to account for existing distortions when evaluating a merger in the propane industry).

could be substantially underestimated by an analysis that ignores preexisting distortions. These observations are especially relevant today as recent research has shown that the level of market power at the product level has increased, perhaps substantially, since the 1980s.[9] On the buying side, there has been a growing recognition of the importance of monopsony power, especially in the labor market.[10] These observations suggest that an analysis of mergers or other conduct must consider the costs and benefits of actions in a world with preexisting distortions. The conclusion is that antitrust matters more than one might otherwise think and therefore there are large economic costs to abandoning or de-emphasizing economic principles, as some have suggested,[11] since the use of economics should lead to fewer such errors.

Part I shows how the usual cost-benefit analysis changes in the presence of preexisting distortions such as market power or taxes. This analysis is then applied to the calculation of optimal damages, using a seminal article by Professor William Landes. Part II explains several policy insights that emerge from such an analysis. Finally, Part III discusses some refinements in using this cost-benefit analysis for antitrust. These refinements extend the analysis to include non-price variables such as investment. The conclusion explains that de-emphasizing antitrust economics would be a mistake. Government agencies with their well-trained economic staffs should be able to apply the more complicated analyses described here. In contrast, courts do not necessarily have that capability and therefore, for them, a good rule of thumb is to ask whether total industry output will rise or fall as a result of the conduct under analysis.

[9] Although the price-cost margin does appear to have increased, the magnitude of the increase is a matter of contention. *See* Dennis W. Carlton, *Some Observations on Claims That Rising Market Power Is Responsible for U.S. Economy Ills and That Lax Antitrust Is the Villain*, COMPETITION POL'Y INT'L ANTITRUST CHRON., Aug. 2020, at 1, 6–7; *see also* Lucia S. Foster, John C. Haltiwanger & Cody Tuttle, *Rising Markups or Changing Technology?* 4–5 (Nat'l Bureau of Econ. Rsch., Working Paper No. 30491, 2022) (applying different empirical methods to find lower estimates of markups than prior papers).

[10] *See generally* Orley Ashenfelter, David Card, Henry Farber & Michael R. Ransom, *Monopsony in the Labor Market: New Empirical Results and New Public Policies*, 57 J. HUM. RES. 3–4 (Supp. 2022) (reviewing literature on monopsony and labor markets).

[11] *See, e.g.*, Lina M. Khan, *The End of Antitrust History Revisited*, 133 HARV. L. REV. 1655, 1682 (2020) (reviewing TIM WU, THE CURSE OF BIGNESS: ANTITRUST IN THE GILDED AGE (2018)) (concluding that a reformed antitrust framework should include values such as justice and fairness).

I. Altering the Standard Analysis to Account for Preexisting Distortions

A distortion between price and marginal cost can arise from market power or taxes.[12] This Article focuses on these two sources of distortions although there can be many others, especially arising from various regulations. It should be straightforward to see how the analysis below can be modified to take account of any existing distortion. The first section focuses on distortions that arise in the product market directly affected by the conduct under examination. The second section discusses distortions in other markets that are related to the one directly affected by the conduct under examination.

A. *Preexisting Distortion in the Market Where the Conduct Takes Place*

Suppose that an industry is in a competitive equilibrium at price, p0, and quantity, q0, where p0=c=marginal cost. Then, suppose there is a merger (to keep it simple, suppose there are only two firms pre-merger) or some other single firm conduct (again for simplicity in this case assume there is only one firm) that causes the equilibrium to shift to price, p1, and quantity, q1, as shown in Figure 1. Price has risen and quantity has fallen compared to the preexisting equilibrium, and so consumers are clearly worse off. The harm to consumers (lost consumer surplus) is given by C + D + E in Figure 1. The gain to the supplying firms is C + D, so the deadweight loss to society (lost total surplus) is E. Ordinarily, E is so much smaller than C + D for small price increases that it is often ignored, though ignoring it can be misleading for large price changes. Technically, for small changes in price and quantity, E is considered "second order" compared to C + D because, unlike C + D, it is a triangle whose sides are Δp (p1 - p0) and Δq (q0 - q1), while C + D is a rectangle whose sides are Δp (p1 - p0) and q1. Hence, C + D Is larger than E.[13] (Roughly speaking, the multiplication of Δp and Δq is considered "second order" since the multiplication of two tiny numbers is very tiny.) Generally, in antitrust evaluation, regardless of whether one is concerned with the change in consumer welfare (C + D + E) or total welfare (E), one should want to prohibit actions that lower consumer or total welfare. In practice, it is rare to come up with different answers depending on which surplus criteria is used.[14] That is

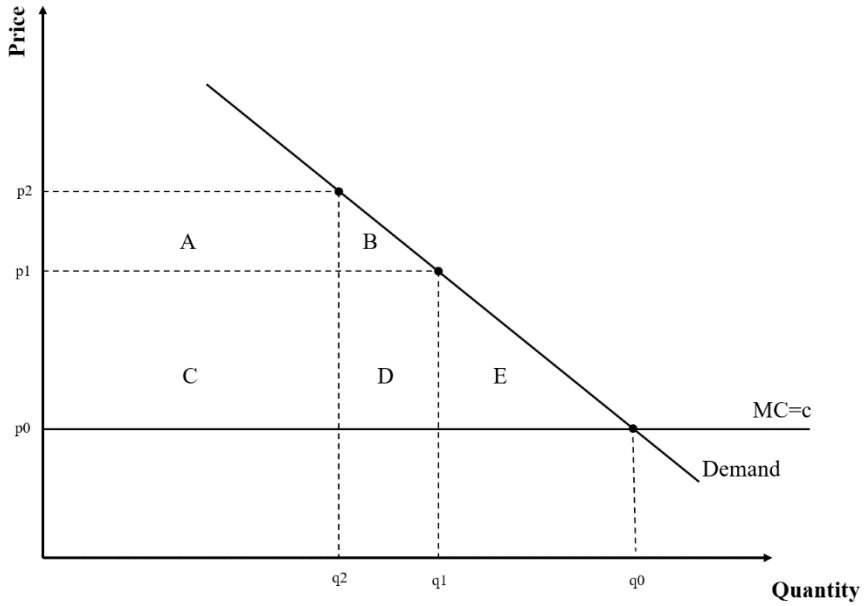
[12] Harberger, *supra* note 4, at 794.

[13] The area of E = .5 × Δp × Δq while the area of C + D = Δp × q1. The ratio of the area of E to the area of C + D is, therefore, .5 × Δq/q1, which is small when Δq is small.

[14] Professors Areeda and Hovenkamp argue that "few if any decisions have turned on the difference" between the total versus consumer welfare standard. Phillip E. Areeda & Herbert

why the recent debate about whether total or consumer surplus is the "better" measure is not important for policy.

Figure 1: Effects of Conduct in the Market in Which it Occurs



Now, in Figure 1, suppose that there is a preexisting distortion, either because of market power or taxes. Recall that a distortion means that there is a gap between the price a consumer pays and the marginal cost a supplier incurs.[15] In this situation, instead of starting at the equilibrium (p0, q0) in Figure 1, we start at (p1, q1). Also suppose that the distortion, measured by the difference between p1 and p0, is not small enough to be ignored. For example, instead of a few percentage point difference, suppose the difference between p1 and marginal cost is 60%, which some have indicated is a typical price-cost margin,[16] or suppose there is a sales tax of 10%, which is roughly what a sales tax is in some localities.[17] With

Hovenkamp, ANTITRUST LAW: AN ANALYSIS OF ANTITRUST PRINCIPLES AND THEIR APPLICATION ∫ 114b (5th ed. 2022). Richard Posner argues that the profits of the firm will be dissipated by unproductive activity and, thus, should be considered a loss to society when calculating total surplus. *See* Richard A. Posner, *Theories of Economic Regulation*, 5 BELL J. ECON. & MGMT. SCI. 335, 339 (1974).

15 *See* Harberger, *supra* note 4, at 794.

16 *See* Jan De Loecker, Jan Eeckhout & Gabriel Unger, *The Rise of Market Power and the Macroeconomic Implications*, 135 Q.J. ECON. 561, 575 (2020).

17 *See* Janelle Fritts, *State and Local Sales Tax Rates, 2023*, TAX FOUND. (Feb. 7, 2023), https://perma.cc/VT3C-EAHZ ("The five states with the highest average combined state and local

this preexisting distortion in place, imagine that there is some conduct that results in the price rising from p1 to p2 and quantity falling from q1 to q2. Note that the preexisting distortion can be of the same order of magnitude or even larger than many anti-competitive price effects. That is, the preexisting distortion (the difference between p1 and c) can be as large or larger than the increase in price from p1 to p2.

The harm to consumers from the additional price increase related to the conduct at issue is A + B. Assuming the distortion is due to market power of the supplying firms, the net gain to the supplying firms is A - D, and the net incremental harm to society from the conduct is B + D. The reason D now appears as part of the harm to society is that some consumers are driven off by the higher price p2, reducing quantity from q1 to q2; consumers' valuation of each of those lost units between q1 and q2 can be read off the demand curve and exceeds the marginal cost to society for each of the lost units (which equals c). Thus, when quantity declines, area D represents the loss to society—and to firms—from the conduct. If the preexisting distortion is instead due to taxes, then the loss to consumers is still A + B, but the gain to firms is now A and the loss to the government (from lower tax revenues), and therefore to society, is D. Again, the total incremental harm to society from the conduct is B + D.

So far, the conduct always leads to a reduction in output and therefore is always anti-competitive. If we are interested in either consumer or societal welfare, then that conduct should be prevented. Most conduct in antitrust cases, such as merger activity or exclusionary behavior, also has the potential to improve efficiency.[18] If the improved efficiency lowers marginal costs sufficiently so that prices wind up falling, then from both a consumer and total surplus viewpoint, the conduct is desirable. It is of course possible that there can be a reduction in, say, fixed costs and a reduction in output such that a merger increases total surplus but decreases consumer surplus.[19] As alluded to earlier, these cases in which consumers are harmed but total surplus rises are extremely rare. Hence, a good, practical rule of thumb is that the conduct is desirable if it lowers price and expands total output.

To illustrate how much difference accounting for area D can make, let us return to Williamson's 1968 article. Williamson showed that

sales tax rates are Louisiana (9.550 percent), Tennessee (9.548 percent), Arkansas (9.46 percent), Alabama (9.25 percent), and Oklahoma (8.98 percent).").

[18] *See, e.g.*, DENNIS W. CARLTON & JEFFREY M. PERLOFF, MODERN INDUSTRIAL ORGANIZATION 20–25 (4th ed. 2005). As one example of an efficiency, suppose that two manufacturing firms merge and by doing so are able to achieve economies in scale in manufacturing with the result that the average cost of producing an item falls.

[19] For example, if we start at (p0, q0) and a merger allows prices to rise to p1 but fixed costs fall by more than E, the merger will harm consumers but benefit society.

efficiencies are important to take account of in evaluating a merger and showed that even a merger that raised price could produce a net benefit to society, using total surplus.[20] I have already discussed that such a case is rare in antitrust in which consumers are harmed but total surplus rises. Still, I can use Williamson's example to illustrate the importance of area D. Williamson's analysis calculates for a variety of different assumptions on price change and demand elasticity, the required percentage decline in unit costs needed to offset the price increase and increase total surplus.[21] He shows, for example, that if the demand elasticity is -2 and the percent increase in price is 5%, then for total surplus to rise, as long as costs fall by at least 0.25%, total welfare increases, in the absence of initial distortions.[22] The point of his numerical example is to emphasize how even tiny cost efficiencies can lead a merger to increase total surplus even if prices rise. But if the initial equilibrium has a preexisting distortion of say 10% (so price is above marginal cost by 10%), then if one recalculates Williamson's required efficiency to offset a 5% price increase, the required efficiency jumps to about 1.3%—an approximate fivefold increase above what Williamson calculated. If one were to assume a distortion of not 10% but say 50%, then the required efficiency is around 8%—an increase of 32-fold above what Williamson calculated. So, area D can matter a lot.

It is also possible to derive optimal damages for an antitrust violation. Landes explained how to do this in a 1983 seminal article for the case when there are no preexisting distortions.[23] Landes showed that to deter the firm(s) from anti-competitive activity, a penalty based on the net harm to others is necessary.[24] Thus, if the preexisting distortion at (p1, q1) is caused by market power, the optimal penalty for anti-competitive conduct that moves us from (p1, q1) to (p2, q2) is A + B because that is the harm to consumers.[25] However, if the distortion is caused by taxes, the optimal penalty is A + B + D because that is the harm to consumers and the government.[26] As already mentioned, area D can be substantial and of the same order of magnitude as area A, depending on the facts of the market at issue.

---

20 Williamson, *supra* note 1, at 22–23.

21 *Id*. at 23.

22 *Id.*

23 *See generally* William M. Landes, *Optimal Sanctions for Antitrust Violations*, 50 U. CHI. L. REV. 652 (1983) (explaining how to derive optimal damages for an antitrust violation).

24 *Id.* at 674.

25 The firm's increased profits from moving from (p1, q1) to (p2, q2) is A - D, so if the firm pays a penalty of A + B, its net profit is A - D - A - B or –(D + B). That is obviously negative, and only if there is an efficiency at least equal to (D + B) would it be efficient to society for the conduct to occur.

26 The firm's increased profits from moving from (p1, q1) to (p2, q2) is equal to A and, thus, after paying the penalty of A + B + D, its net profit is –(D + B) as in the prior footnote.

B. *Preexisting Distortions in a Related Market*

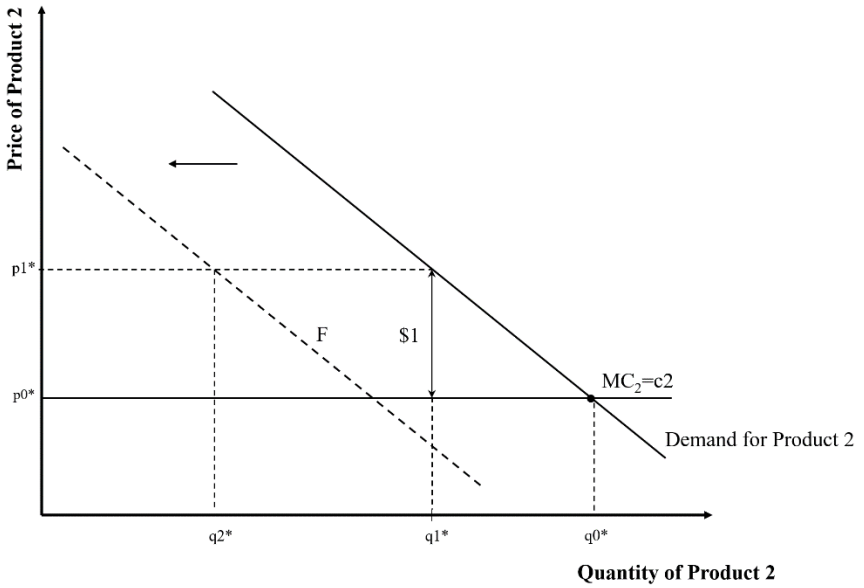
Now consider the effects of conduct in one market on a related market. Suppose that there is some product 2, not owned by the firms that produce product 1, and whose consumption depends on the price of product 1. That means the demand curve for product 2 will depend upon the price of product 1 so that the demand curve for product 2 will shift in response to a change in the price of product 1. Suppose that product 2 is a complement to product 1 so that its demand curve shifts to the right (demand increases) as the price of product 1 falls and its demand curve shifts to the left (demand decreases) as the price of product 1 rises. If the market for product 2 is initially undistorted (the price of product 2 equals its marginal cost), then there is no need to do anything to account for shifts in the demand curve for product 2 in evaluating the welfare consequences of conduct in the market for product 1.

But suppose the market for product 2 is distorted in its initial equilibrium either because of market power in product 2 or taxation of product 2. The welfare consequences of conduct in market 1 will need to account for what happens in market 2. For example, suppose there is a distortion such that price exceeds marginal cost by $1 at the initial equilibrium for product 2. See Figure 2. Then, if the price of product 1 rises from p1 to p2, that causes the entire demand curve for product 2 to shift to the left and consumption of product 2 falls from q1* to q2*. That causes an additional harm in product market 2, roughly equal to $1 times (q1* - q2*), or area F in Figure 2. The way to think about this harm is that even holding the price of product 2 unchanged, a decrease in the quantity of product 2 consumed by, say, one unit, deprives society of a net benefit of $1, the gap between what the marginal consumer values product 2 at and the cost to society of producing the product. That harm falls on either the tax authority (if the distortion in product market 2 is from a tax) or on product 2 firms (if the distortion in product market 2 is from market power). When a related market such as that of product 2 is initially distorted in addition to the market for product 1, the total harm to society from an anti-competitive act in market 1, that raises the price of product 1 from p1 to p2, will be B + D + F. Notice that area F is not second order and, therefore, F cannot be omitted from a harm calculation without risking serious error.[27] This is why public finance economists, when calculating the cost of imposing a tax on a product, recognize the

[27] Economists generally prefer using total surplus rather than consumer surplus for cost-benefit analysis. Whatever surplus measure (consumer or total) one chooses to use, however, area F should be added to the harm even if the harm represented by F is to producers in product market 2 since those producers have nothing to do with the act in product market 1.

importance of taxes in related markets.[28] Obviously, taking F into account would further increase the required cost savings, calculated above, to offset a price increasing merger.

Figure 2: Effects of Conduct in a Related Market



The fact that product markets 1 and 2 are each distorted in their initial equilibria means there is an externality because actions in one market affect consumer and social welfare in another market. This means there is a coordination problem that is not being internalized by independent firms, an issue recognized long ago by Professor Augustin Cournot.[29] This provides a rationale for integration to enable a firm to internalize these interactions between the demand for related products. Integration to address this complements issue is akin to vertical integration to eliminate double marginalization.[30]

28 *See* Harberger, *supra* note 4, at 789–90.

29 *See* AUGUSTIN COURNOT, RESEARCHES INTO THE MATHEMATICAL PRINCIPLES OF THE THEORY OF WEALTH 99–116 (The MacMillan Company ed., Nathaniel T. Bacon trans., 1927) (1838) (analyzing the mutual relations of producers).

30 For a discussion of the elimination of double marginalization, see CARLTON & PERLOFF, *supra* note 18, at 439–45. If product 2 were a substitute instead of a complement for product 1, then the demand curve for product 2 would shift out to the right in Figure 2 as the price of product 1 rises. If product 2 is an initially distorted market because of, say, market power, then the expansion of output in product 2 as a result of the price rise in product 1 will lead to a reduction in deadweight loss (i.e.,

It is straightforward to apply Landes' insights to derive the optimal damages for an antitrust violation in product market 1 when there is a preexisting distortion in product market 1. Optimal damages equal the net harm to others, in this case, A + B + F, if the initial distortion in product market 1 is from market power, and A + B + D + F, if the initial distortion is instead due to taxes.

II. Practical Implications for Policy

At a time when there are complaints about economists and complicated quantitative economic models playing too large a role in antitrust compared with lawyers relying on documents, it is appropriate to ask what insights flow from the recognition that most markets are initially distorted by either market power, taxes, or other factors. Isn't that just an additional complication that should be ignored? It is important to emphasize that unlike the many recent, complicated structural economic models used in horizontal and vertical merger analysis, cost-benefit analysis has been used for over fifty years around the world in varied settings.[31] This does not mean that cost-benefit analysis does not rely on assumptions but rather that, compared to recent models in industrial organization, it has withstood the test of time. That indicates that its insights should not be dismissed without further analysis. There are at least several major insights that flow from the cost-benefit analysis presented in Part I.

Insight 1: Antitrust matters more than one might think.

Pro-competitive mergers generate more benefits than one might think if one were to ignore preexisting distortions. For example, suppose we are initially at (p2, q2) in Figure 1. There is a proposed merger that will create efficiencies (not shown in Figure 1) of amount G that will cause the equilibrium in Figure 1 to move from (p2, q2) to (p1, q1) and the equilibrium in Figure 2 to move from (p1*, q2*) to (p1*, q1*). The gain to society is B + D + F + G.[32] If one failed to recognize the importance of existing distortions, one would estimate the gain to society as only B + G. Failure to recognize the presence of preexisting distortions would lead one to underestimate the gain to society by omitting D + F. Those areas can be

area F becomes a subtraction from the harm otherwise created in market 1). There is a reduction because an output expansion of product 2 increases total surplus.

[31] *See generally, e.g.*, Richard O. Zerbe & Tyler Scott, *A Primer for Understanding Benefit-Cost Analysis*, ACTIONABLE INTEL. FOR SOC. POL'Y (2015), https://perma.cc/VXS6-2Z75 (observing the use of benefit-cost analysis by the United States government through presidential executive orders dating back to the Reagan administration in the 1980s).

[32] To avoid excessive repetition, I will use society's welfare in these calculations, but the reader can use consumer surplus (with or without taxation) and find that similar conclusions follow.

large relative to B + G especially since the preexisting distortions giving rise to D and F are of the same order of magnitude (i.e., not second order) as the change in price, p2 - p1, and can easily be at least as large as the change in price. For example, a 5% price decline sounds great, but it easily can be the case that the preexisting distortion in Figure 1, p1 - p0,—or in Figure 2, p1* - p0*—can be at least as large as 5%, leading to even greater benefits than what would be measured from the initial price decline if one ignored preexisting distortions. Preexisting mark-ups or sales taxes can easily be of that magnitude or more.[33] Therefore, the required efficiency to generate a pro-competitive merger, one that benefits consumers or society, is lower—and, depending on the facts, perhaps much lower—than what an analyst would estimate if that analyst ignored the preexisting distortions. *The bottom line is that aggressive antitrust policy that stops or chills efficiency-generating mergers or conduct can deprive the economy of significant benefits.*

It is often difficult for courts to evaluate claimed efficiencies from either a merger or exclusionary conduct. There is no simple way around the need to evaluate the reliability of that evidence, but any claim, including from antitrust agencies, that proposed efficiencies (or proposed improvements in the alignment of incentives) can be achieved through contract should not be blindly accepted,[34] especially when such contracts have not emerged in the pre-transaction setting. For example, in a merger case, to claim that a contract could achieve the efficiencies claimed by the proposed merger begs the question of why a contract has not done so already. Although one must always be on guard for pretextual efficiencies, one must similarly not be too quick to accept the mantra that any efficiencies can be achieved by contract. Such an argument ignores the insights initially made by Professor Ronald Coase[35] and the enormous subsequent literature.[36] Indeed, if one assumes that contracts can reproduce any proposed merger efficiency, why allow mergers at all?

Proponents of aggressive antitrust policy can point to the converse implication that flows from the cost-benefit analysis. The cost-benefit analysis implies that the harm from anti-competitive conduct is more

33 *See, e.g.*, Fritts, *supra* note 17.

34 *See, e.g.*, Opinion of the Commission, Illumina, Inc., FTC Docket No. 9401, at 81 (Apr. 3, 2023) (asserting that "contracting with other capable firms" was an adequate substitute to merger). I served as an expert for Illumina.

35 *See* Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386, 391–92 (1937).

36 *See, e.g.*, OLIVER E. WILLIAMSON, MARKETS AND HIERARCHIES: ANALYSIS AND ANTITRUST IMPLICATIONS 238–45 (1975) (noting that contracts do not remove the incentive to cheat); Bengt Holström & John Roberts, *The Boundaries of the Firm Revisited*, 12 J. ECON. PERSP., no. 4, Fall 1998, at 74 (noting the "hold-up" problem in which two parties cannot resolve their transaction through contract).

than one would otherwise think if one ignored preexisting distortions. Go back to Figure 1 and suppose that there is a merger with no efficiencies that moves the equilibrium from (p1, q1) to (p2, q2) and in Figure 2 from (p1*, q1*) to (p1*, q2*). The harm to society from the conduct, as explained in the prior section, is B + D + F. If one ignored the presence of preexisting distortions, one would have measured the harm as B. Again, areas D + F can be as large or larger than B so one could greatly underestimate the harm. *The bottom line is that aggressive antitrust policy that stops or chills harmful activity can prevent the imposition of large costs on the economy.*

So, how should one judge efficiencies to determine whether a merger, on balance, is good or bad for consumers or society? Business documents, analysts' discussions, and testimony from business people are relevant to the efficiency evaluation of evidence.[37] Perhaps the best role an economist can play in the evaluation is analyzing a track record, if there is one, of the firm achieving proposed efficiencies from similar past conduct, trends in the industry, and any evidence of distortions that has not been corrected by contract or would be difficult to achieve through contract. Although there is no easy answer, ignoring or claiming that efficiencies are irrelevant to the legal decision of whether there is an antitrust harm is akin to sticking one's head in the sand and risking the imposition of large costs on the economy.

Insight 2: In using its discretion to bring cases and trying to prioritize such cases, government agencies should consider, among other factors, the preexisting distortions in a market and related markets.

Economists are taught that resources, even for a government agency, are limited. Cases that prevent harm or establish important precedents that will prevent future harms should take account of existing distortions. For example, there are several industries that face restrictions on competition either through state licensing restrictions or state legislation.[38] Many states have franchise laws that limit entry. Many states have laws on distribution of, for example, alcoholic beverages.[39] Pro-competitive conduct in those markets could potentially yield large benefits.

The prior section showed how large underestimates of competitive harm or benefit are possible when one ignores preexisting distortions. During any merger or conduct investigation, one must consider the significant preexisting distortions, otherwise one risks making a large error.

---

[37] U.S. DEP'T OF JUST. & FED. TRADE COMM'N, MERGER GUIDELINES § 2.2 (rev. ed. 2010).

[38] *See, e.g.*, Aaron Edlin & Rebecca Haw, *Cartels by Another Name: Should Licensed Occupations Face Antitrust Scrutiny?*, 162 U. PA. L. REV. 1093, 1104–07 (2014).

[39] *Control State Directory and Info*, NAT'L ALCOHOL BEVERAGE CONTROL ASS'N, https://perma.cc/8L9P-AY7F.

Insight 3: In using cost-benefit analysis, looking at the aggregate harm to either society or consumers is a more sensible antitrust policy in contrast to a policy that attempts to weigh the effects on each affected group.

One might wonder how this insight emerges from the prior analysis, but it is embedded in the understanding of consumer surplus. There are at least two ways to illustrate the point. The analytic way is to think of a demand curve that slopes downward with each person demanding one unit of the good. Suppose Joe is willing to pay $11 for one unit of the good, Tom is willing to pay $10, and Sally is willing to pay $9. Suppose the price is initially $8 and it rises as the result of some conduct to $11. Joe loses $3 of consumer surplus, Tom loses $2, while Sally loses $1 for a total loss of consumer surplus of $6. Economists typically add together the lost surpluses.[40] One could argue that Sally is more deserving than Tom and her loss should count a lot more than Tom's. Although theoretically possible to do, such calculations in antitrust deliberations are likely to lead to an operational nightmare. Subjective valuations of individuals' welfare are bound to differ among analysts and policymakers and would destroy all objectivity and predictability in antitrust decisions. It is preferable that the desire to help Sally over Tom would lead to other public policies better suited to addressing such concerns than to twist antitrust policy into a mechanism for the social change desired by the current policymakers.

The second way of making the same point is to consider that virtually every merger has some winners and losers among consumers. If Bob is a regular customer of firm 2 and firm 1 buys firm 2, firm 1 may conform firm 2's business practices to align with firm 1's. But as a long-term customer of firm 2, Bob might not like that. Or, as another example, consider an airline merger. If, as a result of a merger, a plane is repositioned from route AB to route BC because more people will now fly on the plane, that seems desirable overall. But, of course, people on route AB would possibly be harmed. Should that be enough to stop the merger? One would suspect most people would say no.[41]

This principle of treating all consumers the same in estimating an aggregate effect also applies to other "related" markets. Many mergers or other conduct can affect a related market, such as the labor market, even if the merger or other action itself does not alter market power in the labor market. When that related market is itself subject to initial distortions, the cost-benefit analysis of the conduct must take into account the effect in the related market. In light of recent research indicating that labor

40 *See, e.g.*, Harberger, *supra* note 4, at 785.

41 As a general matter, based on my experience in many airline mergers investigated by the Department of Justice ("DOJ"), this view coincides—or at least has coincided—with that of the DOJ. *See e.g.*, Ken Heyer, Carl Shapiro & Jeffrey Wilder, *The Year in Review: Economics at the Antitrust Division, 2008-2009*, 35 REV. INDUS. ORG. 349 (2009).

markets often have market power on the buying side (e.g., some monopsony power), it follows that such effects should be accounted for.[42] Moreover, if the merger or other conduct does directly increase or decrease market power in the labor market, then that should also be taken into account. The principles of cost-benefit analysis would suggest not keeping track of each group harmed separately, but instead adding together the harm or benefit of each group to come up with an aggregate number.

Insight 4: Courts have a difficult enough job trying to measure pro- and anti-competitive conduct. We should not increase their burden to evaluate related markets unless the related markets are so intrinsically intertwined with the market being examined that ignoring the interaction would not be economically sensible.

It is a heavy burden for courts to listen to and evaluate complicated economic analysis. For that reason, courts should consider existing market distortions in related markets only when such effects are so intrinsically linked to the conduct being examined that ignoring them would not be economically sensible.[43] I already discussed a few examples of this under insight 3. Another important example is in the context of so-called two-sided markets. An example of a two-sided market is credit cards in which a consumer with a card uses the card to pay a merchant.[44] The merchant pays a merchant fee to the credit card company and the credit card company often pays a reward to the consumer for using its card.[45]

In a recent antitrust case, *Ohio v. American Express Co.*,[46] the Supreme Court dismissed a government challenge to American Express rules that forbade merchants from encouraging consumers to use other credit cards that had lower merchant fees.[47] Although I have been highly critical of the Supreme Court's decision, my criticism had to do with the technical details of market definition and how the Court used that to dismiss what appeared to be anti-competitive conduct without any attempt to pay attention to the evidence.[48] Regardless of whether the market was defined

42 Ashenfelter et al., *supra* note 10, at 1.

43 Both the DOJ and Federal Trade Commission have many highly qualified economists who are capable of evaluating complicated economic studies and performing their own studies.

44 Jean-Charles Rochet & Jean Tirole, *Two-Sided Markets: A Progress Report*, 37 RAND J. ECON. 645, 645 (2006). The use of the term "two-sided market" in the economic literature is not necessarily meant to imply anything about relevant antitrust market definition.

45 *Id.* at 646.

46 138 S. Ct. 2274 (2018). I have served as an expert adverse to credit card companies.

47 *Id.* at 2289–90.

48 Dennis W. Carlton & Ralph A. Winter (2018), *Vertical Most-Favored-Nation Restraints and Credit Card No-Surcharge Rules*, 61 J.L. & ECON., 215, 240–41 (2019); Dennis W. Carlton, *The*

as one-sided or two-sided, an economic analysis required a full understanding of what would happen to both sides of the market because of their close interaction.[49] The error the Court made is instructive. The Court said that in a two-sided transaction market, the price paid by buyers (merchants) plus the price paid by consumers (the rewards received) is the full price, and that is all that matters.[50] That logic would be correct if the two sides were used in fixed proportions *and* the full price was all that mattered to understanding how the market operated. But the essence of a two-sided market is that the full price is *not* a sufficient price to know how the two sides operate.[51] Exactly how much each side pays matters a great deal to market operation even if the full price is unchanged. That is enough to expose the logical error in the Court's reasoning, as my *Journal of Law and Economics* article and Justice Stephen Breyer's dissent explain.[52] However, the two sides of the market are so intertwined that it would be misleading to analyze how conduct affected only one side of this two-sided market.

III. Refinements

There are at least two refinements to antitrust thinking that follow from cost-benefit analysis. The first is that welfare depends on more than price and output. That is, the demand and marginal cost curves in Figures 1 and 2 take as given lots of economic factors such as product quality and investment decisions in capital and a career. Those factors matter a great deal for future welfare and can be influenced by the conduct under analysis. The second is that sometimes the creation of market power can lead to a more efficient outcome. That observation creates a slippery slope for antitrust enforcement. Let us discuss these two refinements in more detail.

Regarding the first refinement, this Article has so far focused on a typical antitrust issue: the effect of conduct on price and output. But there are many other dimensions, in addition to price, such as quality and investments, that conduct can affect with consequences for consumer and

Anticompetitive Effects of Vertical Most-Favored-Nation Restraints and the Error of Amex, 2019 COLUM. BUS. L. REV. 93, 104; *see also* Ohio v. Am. Express Co., 138 S. Ct. 2274, 2303–05 (Breyer, J., dissenting).

49 Carlton & Winter, *supra* note 48, at 234–35.

50 *Am. Express*, 138 S. Ct. at 2287.

51 Rochet & Tirole, *supra* note 44, at 665.

52 Carlton & Winter, *supra* note 48, at 241–42; *Am. Express,* 138 S. Ct. at 2297 (Breyer, J., dissenting).

societal welfare.[53] These dimensions are often more difficult to evaluate than price for two reasons. First, it is often unclear whether all consumers value quality dimensions in the same way. For example, Dennis might prefer higher fares and more frequent flights of airplanes while Jane might prefer lower fares and less frequent flights. Theoretically, one can deal with this issue when evaluating conduct, though empirically it could be difficult.

The second more complicated issue associated with the first refinement has to do with investments. The future of many industries depends on investing in capital and introducing new products. But the effect on welfare of introducing those new products can be hard to measure. A similar issue arises when considering long-run career decisions. Many career decisions, once made, are hard to reverse. But the decisions of future workers will be affected by current and future expectations about working conditions. In the short to medium run, there may be little effect on labor supply if wages fall. But in the long run, there could be a big effect. For example, many doctors have complained that the current concentration of insurance companies has reduced rates of compensation.[54] Those concerns may not affect the supply of doctors in the short run, but can certainly affect the number of doctors over time. Obviously, long-run considerations such as this are difficult to model. One might expect that doctors will try to form groups or entities to negotiate with a concentrated insurance industry, even though sometimes government agencies have intervened to prevent this.[55] There will likely continue to be organizational changes in the health care industry to deal with these situations, though there are no guarantees that the response to the existing regulations will result in an efficient provision of health care.

Consider now the second refinement, what can be labeled as the "slippery slope" issue. Imagine that some industry faces a group of input suppliers that are oligopolists and charge a price above marginal cost. It is well-known that such an industry structure is not an efficient one because the industry is facing an input price that exceeds the input's marginal cost. Suppose that the industry engages in a massive horizontal consolidation with the sole purpose of being able to bargain better with input supplying firms. It is quite possible that the outcome will produce lower input prices

[53] Elyse Dorsey, *Antitrust in Retrograde: The Consumer Welfare Standard, Socio Political Goals, and the Future of Enforcement*, *in* The Global Antitrust Institute Report on the Digital Economy 109, 133 (Joshua D. Wright & Douglas H. Ginsburg eds., 2020), https://perma.cc/UZ6M-K3UJ.

[54] *See, e.g.*, Reed Abelson, *Doctors Accuse UnitedHealthcare of Stifling Competition*, N.Y. TIMES (Apr. 1, 2021), https://perma.cc/2QMN-7QQX.

[55] *See* U.S. DEP'T OF JUST. & FED TRADE COMM'N, IMPROVING HEALTH CARE: A DOSE OF COMPETITION 18 (2004), https://perma.cc/V6HE-CP54.

(at the margin at least) and that could lead the industry to expand output. That, of course, could be socially desirable and a government agency might allow such an outcome. But one might worry that what is really going on is a recognition that competition does not work in the industry. To remedy that, the government agency allows organizational changes to create a more concentrated industry with hopefully efficient bargaining. Maybe such an outcome could occur. If one really believes bargaining is always efficient and transaction costs are low, however, then the Coase theorem, as Professor Harold Demsetz pointed out long ago, tells us there is no need for antitrust at all![56] The market will always lead to an efficient solution through bargaining.[57] But bargaining is not costless and, especially with asymmetric information, we cannot always reach the efficient outcome.

So, we are stuck in some sense. We can allow a less competitive structure to induce a more efficient bargain, but that is a slippery slope in which antitrust—a set of rules premised on competition—is being used to create a concentrated market structure that eliminates lots of competition. If antitrust is used in this way, an analysis should be conducted as to whether antitrust is sufficient to guide the market to a more efficient outcome or whether some additional (or less) regulation is needed, recognizing full well the powerful lessons from U.S. history on the very high costs of regulation in stifling innovation and perpetuating inefficiencies.[58]

Conclusion

This Article shows that in the presence of preexisting distortions in both directly affected and related markets, the failure to do a proper antitrust analysis has the potential to create mistakes much larger than one might think. Deemphasizing economic analysis would therefore be unwise since such a deemphasis is likely to lead to errors in enforcement that this article has shown are likely to be larger than expected based on analyses that ignore these preexisting distortions. It is true that the complications of cost-benefit analysis can be hard to account for, but government agencies are well-suited to do so with their staffs of well-trained economists. Courts, on the other hand, are less suited to apply complicated cost-benefit analyses, and mostly should ignore these complications, with some exceptions. And while admittedly sometimes

56 Harold Demsetz, *The Cost of Transacting*, 82 Q.J. ECON. 33, 33–34 (1968).

57 *Id.*

58 *See, e.g.*, CARLTON & PERLOFF, *supra* note 18, at 682–721.

hard to apply, a good rule of thumb for courts to follow is to allow conduct that will expand total industry output.