7

Competition and Innovation
Did Arrow Hit the Bull's Eye?

Carl Shapiro

> The only ground for arguing that monopoly may create superior incentives to invent is that appropriability may be greater under monopoly than under competition. Whatever differences may exist in this direction must, of course, still be offset against the monopolist's disincentive created by his preinvention monopoly profits.
> —Arrow (1962, 622)

> As soon as we go into details and inquire into the individual items in which progress was most conspicuous, the trail leads not to the doors of those firms that work under conditions of comparatively free competition but precisely to the doors of the large concerns . . . and a shocking suspicion dawns upon us that big business may have had more to do with creating that standard of life than with keeping it down.
> —Schumpeter (1942, 82)

7.1 Introduction

The fiftieth anniversary of the publication of NBER *Rate and Direction of Inventive Activity* volume is an opportune time to revisit what is arguably the most important question in the field of industrial organization: what organization of business activity best promotes innovation?

Carl Shapiro is the Transamerica Professor of Business Strategy at the Haas School of Business and professor of economics at the University of California at Berkeley.

The views presented here should not be attributed to any other person or organization. The author thanks Jonathan Baker, Joe Farrell, Richard Gilbert, Ken Heyer, Michael Whinston, Tor Winston, and participants in the NBER 50th Anniversary Conference on the Rate and Direction of Inventive Activity for very helpful conversations and comments on earlier drafts and presentations of this chapter.

Needless to say, this question has received intense attention by economists and other social scientists, especially since the middle of last century, when the critical importance of innovation to economic growth became more widely appreciated.[1] Hence, I wade into this topic with considerable trepidation. So, let me state at the outset that this essay is intended to be somewhat speculative: an audacious attempt to distill lessons from the huge and complex literature on competition and innovation that are simple and robust enough to inform competition policy.

My ambitious task is made somewhat more manageable because I confine myself to one specific question: how can *competition policy* best promote innovation? I do not attempt to address broader questions regarding innovation policy or competitive strategy. Within the realm of competition policy, I focus on the assessment of proposed mergers. Even in this more limited area, I am not the first to attempt to distill robust principles suitable for competition policy. To the contrary, I follow closely in the footsteps of Baker (2007), Gilbert (2006), and Katz and Shelanski (2005 and 2007), and borrow unabashedly from their work. Baker (2007) is closest in spirit to this chapter: he identifies four principles relating competition and innovation and argues strongly that antitrust fosters innovation.[2]

Before putting forward my central thesis—hypothesis, really—let us review the bidding.

Arrow (1962) famously argued that a monopolist's incentive to innovate is less than that of a competitive firm, due to the monopolist's financial interest in the status quo. This fundamental idea comports with common sense: a firm earning substantial profits has an interest in protecting the status quo and is thus less likely to be the instigator of disruptive new technology. In Arrow's words: "The preinvention monopoly power acts as a strong disincentive to further innovation."[3] Consciously oversimplifying, the Arrow position can be summarized by this principle:

Arrow: "Product market competition spurs innovation."

1. I make no attempt to survey the huge theoretical and empirical literature that explores the relationship between competition and innovation, and I apologize in advance to those whose important contributions are not explicitly cited here. I rely heavily on Gilbert (2006) and Cohen (2010). See also Sutton (1998) and (2007). Aghion and Griffith (2005) and Aghion and Howitt (2009) discuss the relationship between competition and economic growth.

2. Baker's four principles are: (1) competition among firms seeking to develop the same new product or process encourages innovation; (2) competition among firms producing an existing product encourages them to find ways to lower their costs or improve their products; (3) firms that expect to face more product market competition after innovating have less incentive to invest in R&D; and; (4) a firm will have an extra incentive to innovate if doing so discourages its rivals from investing in R&D.

3. Arrow (1962, 620). Put differently, the secure monopolist's incentive to achieve a process innovation is less than that of a competitive firm because the monopolist with lower costs will merely replace itself, while the competitive firm will (by assumption) take over the market, in which it previously earned no economic profits. Tirole (1997, 392), dubbed this the "replacement effect."

Schumpeter (1942), by contrast, even more famously emphasized that a great deal of innovation is attributable to large firms operating in oligopolistic markets, not to small firms operating in atomistic markets.

> The firm of the type that is compatible with perfect competition is in many cases inferior in internal, especially technological, efficiency. (Schumpeter 1942, 106)

While he was no fan of entrenched monopolies, Schumpeter argued that larger firms have greater incentives and ability to invest in R&D.[4] He dismissed perfect competition as the ideal market structure, stressing the importance of *temporary* market power as a reward to successful innovation:

> A system—any system, economic or other—that at *every* point in time fully utilizes its possibilities to the best advantage may yet in the long run be inferior to a system that does so at *no* given point in time, because the latter's failure to do so may be a condition for the level or speed of long-run performance. (Schumpeter 1942, 83)

Consciously oversimplifying, the Schumpeter position can be summarized in this principle:

> Schumpeter: "The prospect of market power and large scale spurs innovation."

Let the battle be joined. Arrow versus Schumpeter, in the superheavyweight class.

Wait a minute. Are the Arrow and Schumpeter positions really incompatible? This chapter advances the claim that they are *not,* at least so far as competition policy is concerned.

What do we actually *need* to know about the relationship between competition and innovation for the purposes of competition policy? For merger enforcement, we need a framework to evaluate the effects of a proposed merger on innovation. In practice, the relevant mergers are those between two of a small number of firms who are best placed to innovate in a given area. For other areas of antitrust enforcement, we typically seek to evaluate the impact on innovation of a specific business practice, such as the package licensing of a group of patents or the decision to keep an interface proprietary rather than open. For these purposes, I argue here that we do not need a universal theory of the relationship between competition and innovation. I also argue that the Arrow and Schumpeter perspectives are fully compatible and mutually reinforcing.

Consciously oversimplifying yet again, I offer three guiding principles. These are stand-alone, ceteris paribus principles, but they work in con-

4. Schumpeter also argued that large established firms operating in oligopolistic markets are better able to finance R&D than are small firms operating in atomistic markets. In the light of today's highly developed capital markets, including venture capital markets, this argument has much less salience today.

cert, weaving together and integrating the Arrow and the Schumpeter perspectives:

> Contestability: "The prospect of gaining or protecting profitable sales by providing greater value to customers spurs innovation."

The Contestability principle focuses on the extent to which a firm can gain profitable sales from its rivals by offering greater value to customers. Sales are contestable in the relevant sense if profitable sales shift toward the successful innovator. This in turn depends on the nature of ex post product market competition. If market shares are sticky, for example, because consumers have strong brand preferences or high switching costs, relatively few sales are contestable and innovation incentives will be muted.

The Arrow effect fits well with the Contestability principle: for a given level of ex post sales, a firm with few ex ante sales has more to gain from innovation. Put differently, a firm that will make substantial sales even if it does not innovate (such as Arrow's incumbent monopolist, which faces no threat) has muted incentives to innovate.

The Schumpeter effect also fits well with the Contestability principle: companies making major innovations often are rewarded with large market shares, leading to high ex post market concentration. Conversely, a small firm that will not be able to grow much, even if it successfully innovates, has lower incentives to invest in R&D than a larger firm.

> Appropriability: "Increased appropriability spurs innovation."

The Appropriability principle operates at the level of the firm. Greater appropriability by one firm can reduce appropriability by other firms and thus retard *their* innovation.

The Appropriability principle focuses on the extent to which a successful innovator can capture the social benefits resulting from its innovation.[5] In practice, appropriability depends heavily on the extent to which a firm can protect the competitive advantage associated with its innovation. If imitation is rapid, so a firm that successfully innovates is unable to differentiate its products or achieve a significant cost advantage over its rivals, ex post profit margins will be low and innovation incentives will be muted. With rapid and effective imitation, contestability can be of limited relevance, since an innovating firm will not be able to offer superior value to customers.

The Schumpeter effect fits well with the Appropriability principle: one cannot expect substantial innovation (from commercial firms, at least) if rapid imitation causes ex post competition to be so severe that even a successful innovator earns little profit.

5. The social contribution of a firm that develops a new product before others do so independently only reflects the value of the earlier development, not the total benefits associated with the new product. See Shapiro (2007) for a more extensive discussion of appropriability in the context of multiple independent invention.

Synergies: "Combining complementary assets enhances innovation capabilities and thus spurs innovation."

The Synergies principle emphasizes that firms typically cannot innovate in isolation. The quest for synergies is especially important in industries where value is created by systems that incorporate multiple components, as in the information and communications technology sector. The Synergies principle is directly relevant for competition policy since procompetitive mergers and business practices allow for the more efficient combination of complementary assets.

The Contestability and Appropriability principles relate to the *incentive* to innovate. The Synergies principle relates to the *ability* to innovate. *None* of these principles relates directly to product market concentration.

This chapter advances the hypothesis that the Contestability principle, the Appropriability principle, and the Synergies principle are sufficiently robust to guide competition policy. I sketch out the argument that these three principles provide the conceptual and empirical basis for a rebuttable presumption that a merger between two of a very few firms who are important, direct R&D rivals in a given area is likely to retard innovation in the area. Furthermore, I suggest, somewhat tentatively, that we have a pretty good understanding of the circumstances under which that presumption is rebutted and innovation is furthered by allowing two important, direct R&D rivals to merge. I also suggest that these three principles can usefully guide competition policy in other areas.

Perhaps you already are convinced that innovation is generally spurred by competition as reflected by the intuitive notions of contestability, appropriability, and synergies. If so, you may want to stop right here, or skip to the later discussion where I apply these principles to competition policy. But as someone actively involved in antitrust enforcement, it appears to me that a rather different, and misleading, "complexity proposition" has taken root and threatens to become the conventional wisdom, namely:

Complexity 1: "The relationship between competition and innovation is so complex and delicate that there should be no presumption that the elimination of product market or R&D rivals will diminish innovation."

A version of this complexity proposition specific to mergers has also gained some currency:

Complexity 2: "The relationship between competition and innovation is so complex and delicate that there should be no presumption that a merger between two of a very few firms conducting R&D in a given area is likely to diminish innovation."

These propositions echo various more general statements from the literature on competition and innovation, where it has become de rigueur to

emphasize that "competition" has ambiguous effects on innovation. For example, Gilbert (2006) states that the incentives to innovate

> [D]epend upon many factors, including: the characteristics of the invention, the strength of intellectual property protection, the extent of competition before and after innovation, barriers to entry in production and R&D, and the dynamics of R&D. *Economic theory does not offer a prediction about the effects of competition on innovation that is robust to all of these different market and technological conditions.* Instead, there are many predictions and one reason why empirical studies have not generated clear conclusions about the relationship between competition and innovation is a failure of many of these studies to account for different market and technological conditions. (Gilbert 2006, 162, emphasis supplied)

In a similar vein, Motta (2004) writes:

> Both theoretical and empirical research on the link between market structure and innovation is not conclusive, even though a "middle ground" environment, where there exists some competition but also high enough market power coming from the innovative activities, might be the most conducive to R&D output. (Motta 2004, 54)

Davis (2003) is an example of the type of message that is reaching antitrust practitioners. He states that there is a "consensus or near-consensus" that "the relation of market structure to market conduct and performance in innovation is far more problematic than in the case of price competition" (695–96).

Certainly, the overall cross-sectional relationship between firm size or market structure and innovation is complex. Just think of all the variations we often see in the real world.

On the Arrow side of the ledger, that is, in praise of innovation by firms without a strong incumbency position, we have the following:

- Disruptive entrants are a potent force. They can shake up a market, bringing enormous value to customers. The mere threat of disruptive entry can stir inefficient incumbent firms from their slumber.
- Firms without a significant incumbency position may have a freer hand to innovate because they are not tied to an installed base of customers. Christenson (1997) provides an insightful and influential study along these lines.
- Firms with strong incumbency positions often resist innovations that threaten those positions. Such resistance can even take the form of exclusionary conduct that violates the antitrust laws.
- Start-up firms often play the role of disruptive entrants, introducing new products or processes.
- Firms with suitable capabilities entering from adjacent markets often play the role of disruptive entrants.

On the Schumpeterian side of the ledger, that is, in praise of innovation by large firms with an established incumbency position, we have the following:

- Some highly concentrated markets exhibit rapid innovation, and some atomistic markets seem rather stuck in their ways. One suspects that these differences are not simply the result of differences in technology opportunity.
- Larger firms often are closer to the cutting edge in technology than their smaller rivals.
- Larger firms can have greater incentives to achieve process improvements because they can apply these improvements to a larger volume of production. In contrast, a smaller firm that cannot grow significantly, even if it successfully innovates, and cannot license out its innovation, has a lower incentive to lower its costs.
- Large firms often acquire innovative start-up firms, or enter into other arrangements such as licenses or joint ventures with them, thereby accelerating the adoption and diffusion of those firms' inventions.

On top of all this, we know that appropriability matters a great deal for innovation incentives.

So, let me be clear: nothing in this chapter should be read to question the proposition that the overall relationship between product market structure and innovation is complex. The relationship between firm size and innovation is also complex. General theoretical or empirical findings about these relationships remain elusive, in part because a firm's innovation incentives depend upon the *difference* between its pre-innovation and post-innovation size. This difference depends upon the ex ante market structure and reflects the ex post market structure.

But we are not totally at sea. Yes, the world is complex, but my aim here is to suggest some general lessons for competition policy when evaluating innovation effects. Even stating these lessons requires that we be quite careful in defining our terms. Implementing them requires that one be willing and able to distinguish different settings based on a few key, observable characteristics. This approach is similar to the one advocated by Gilbert (2006), who writes:

> The many different predictions of theoretical models of R&D lead some to conclude that there is no coherent theory of the relationship between market structure and investment in innovation. That is not quite correct. The models have clear predictions, although they differ in important ways that can be related to market and technological characteristics. It is not that we don't have a model of market structure and R&D, but rather that we have many models and it is important to know which model is appropriate for each market context. (Gilbert 2006, 164–65)

The argument developed here is that competition policy can be usefully and substantially guided by the Contestability principle, the Appropriability principle, and the Synergies principle. Let me illustrate how that could work, by way of a real-world example.

In 2003 and 2004, the Federal Trade Commission (FTC) reviewed the merger between Genzyme and Novazyme, the only two firms pursuing enzyme replacement therapies to treat Pompe disease, a rare genetic disorder. The FTC Chairman Timothy Muris, explaining the Commission's decision not to challenge the merger, explicitly relied on the proposition that "economic theory and empirical investigations have not established a general causal relationship between innovation and competition."[6] This statement, taken alone, is unobjectionable. As noted before and discussed more later, much of the theoretical and empirical literature on the relationship between market structure and innovation emphasizes complexity while seeking to explain how different factors affect that relationship, recognizing that both market structure and innovation are endogenous. Nonetheless, I argue here that we *do* know enough to warrant a presumption that a merger between the only two firms pursing a specific line of research to serve a particular need is likely to diminish innovation rivalry, absent a showing that the merger will increase appropriability or generate R&D synergies that will enhance the incentive or ability of the merged firm to innovate.

Applying the Contestability, Appropriability, and Synergies principles might well have led to a different outcome in the Genzyme/Novazyme merger. Since these two companies were the only ones with research programs for enzyme replacement therapies to treat Pompe disease, the merger eliminated R&D rivalry—and thus reduced contestability—in that area. Successful innovation in this case clearly offered the prospect of gaining significant, profitable sales: the first innovator would establish a new market, and the second innovator could capture profitable sales from the first if its treatment was sufficiently superior. Invoking a presumption that a merger between the only two R&D rivals in a given area reduces competition, the merger would have been challenged absent a showing that it substantially increased appropriability or led to significant innovation synergies to offset the reduced inventive to innovate resulting from the merger. See section 7.5.2 for an extended discussion of this case.

The Genzyme and Novazyme merger is just one (prominent) example of how the "complexity perspective" on competition and innovation has taken root. As Katz and Shelanski (2007) note, some observers "argue that innovation provides a rationale for a more permissive merger policy. One argument advanced in support of this line of reasoning appeals to what is

6. Statement of Chairman Timothy J. Muris in the Matter of Genzyme Corporation/Novazyme Pharmaceuticals Inc., January 13, 2004, at http://www.ftc.gov/opa/2004/01/genzyme.shtm, citing FTC (1996) vol. I, chapter 7, at 16.

known as 'Schumpeterian competition,' in which temporary monopolists successively displace one another through innovation."[7] While not going as far as Chairman Muris, Katz and Shelanski are sufficiently swayed by these arguments to write: "In brief, we recommend that merger review proceed on a more fact-intensive, case-by-case basis where innovation is at stake, with a presumption that a merger's effects on innovation are neutral except in the case of merger to monopoly, where there would be a rebuttable presumption of harm" (6). While merger analysis tends to be highly fact-intensive, whether or not innovation effects are at issue, the standard proposed by Katz and Shelanski appears to be markedly more lenient than the one antitrust law usually applies to horizontal mergers, where there is a rebuttable presumption of harm from a merger that substantially increases concentration and leads to a highly concentrated market.[8]

Here I question whether such a lenient standard is appropriate for evaluating the impact of mergers on innovation. Yet I do not want to direct too much attention to presumptions and burdens of proof, which are more the stuff of lawyers than economists. Nor do I want to overstate the differences between my approach and that of Katz and Shelanski.[9] The key operative question is whether one can obtain reasonable accuracy in merger enforcement, in cases involving innovation, by focusing the inquiry on (1) the extent of future rivalry between the two merging firms, including consideration of the innovative abilities, efforts, and incentives of other firms, and (2) any merger-specific efficiencies that will enhance the incentive or ability of the merged firm to engage in innovation. Part (1) here asks whether the merger significantly reduces contestability; if so, part (2) asks whether the merger nonetheless enhances innovation by increasing appropriability or enabling merger-specific synergies. See section 7.5 later in the chapter.

Likewise, in evaluating the impact of specific conduct by a dominant firm on innovation, the operative question for competition policy is not whether large firms innovate more than small ones, or whether concentrated market structures are associated with more or less innovation than atomistic market structures. After all, competition policy, at least as practiced in the United States today, is not about engineering market structures or the size distri-

7. Katz and Shelanski (2007, 4, footnote omitted).

8. The strength of the "structural presumption" in antitrust law has declined in recent decades, but not nearly to the point where only mergers to monopoly are presumed to substantially lessen competition. See Baker and Shapiro (2008). The 2010 Horizontal Merger Guidelines issued by the Department of Justice and the Federal Trade Commission state in Section 5.3: "Mergers resulting in highly concentrated markets [HHI greater than 2,500] that involve an increase in the HHI of more than 200 points will be presumed to be likely to enhance market power. The presumption may be rebutted by persuasive evidence showing that the merger is unlikely to enhance market power."

9. See the later discussion of the FTC's 2009 challenge to the proposed merger between Thoratec and Heartware. Shelanski was Deputy Director of the Bureau of Economics at the time of that challenge. See also the discussion of the 2010 Horizontal Merger Guidelines. Shelanski was closely involved in developing these new guidelines (as was this author).

bution of firms. The operative question is whether the specific conduct at issue that allegedly excludes a rival, such as a refusal to open up an interface, will benefit customers by spurring innovation or harm them by retarding innovation (e.g., by excluding an innovative rival or reducing the competitive pressure placed on the dominant firm). See section 7.6 later in the chapter.

Section 7.2 shows that the emerging conventional wisdom—that there is no reliable relationship between competition and innovation—results in part from the peculiar and unhelpful way that the notion of "more competition" has been defined in the industrial organization and endogenous growth literatures. Section 7.3 gives a brief summary of the relevant empirical literature, which strongly supports the general proposition that greater competition spurs innovation, broadly defined. Section 7.4 discusses the Contestability, Appropriability, and Synergies principles and argues that they are sufficiently robust to guide competition policy. Sections 7.5 and 7.6 apply these three principles to merger enforcement policy and dominant firm conduct, respectively. Section 7.7 concludes.

7.2 Competition and Innovation: What Went Wrong?

Much of the literature on the relationship between competition and innovation has, unfortunately, given policymakers a clouded and distorted picture of what we really know about this relationship. As a result, the literature has not been as helpful to practitioners as it could be. Worse yet, the way in which the literature has been summarized and translated for policymakers is leaving a misleading impression, especially for nonspecialists. The problem stems in large part from the way the term "competition" has been used in that literature.

7.2.1 Equating "More Competition" with "Less Product Differentiation"

In the theoretical industrial organization literature on competition and innovation, "more competition" frequently is modeled as "less product differentiation." If the products offered by the various competing firms are close substitutes, price competition is more intense. So, "less product differentiation" is not an unreasonable way to define "more competitive pressure," at least in a static oligopoly setting. However, this has resulted in numerous statements in the literature that can be misleading and unhelpful for the purpose of competition policy, especially merger enforcement. In particular, while merger enforcement can directly affect the number of independent firms competing in an industry, it does not directly affect the extent of product differentiation among the products offered by those firms.

The danger can be illustrated by the discussion in Aghion and Griffith (2005). They begin in chapter 1, "A Divorce Between Theory and Empirics," with what they label as the "dominant theories of the early 1990s."

These are static models of product differentiation in which an increase in product market competition is modeled as a reduction in the extent of product differentiation, such as lower transportation costs in a model of spatial differentiation. Innovation is then measured by the equilibrium number of firms in the market, where entry involves a fixed cost. With weaker product differentiation, price/cost margins are smaller, and fewer products are supplied in the free entry equilibrium. This simple and uncontroversial proposition about product *variety* is characterized as a "Schumpeterian effect of product market competition" (11). Aghion and Griffith go on to state, "we again obtain an unambiguously negative Schumpeterian effect of product market competition on innovation" (12).[10]

I am not disputing the results in these simple models of product differentiation. Nor am I disputing that innovation incentives are low if successful innovation merely places a firm in a market where its product is only slightly differentiated from other products and where the firm has no cost advantage. What I am disputing is that such a proposition is helpful for competition policy, or innovation policy more generally. Meaningful product innovation involves the development of new products that are superior to, or at least significantly distinct from, existing products. Meaningful process innovation involves the development or adoption of production processes (broadly defined to include business methods) that significantly reduce costs. These static models of oligopoly do not involve anything I recognize or credit as innovation. They may help us understand how many brands of toothpaste will be introduced, but they cannot help us understand how firms choose to invest to develop new products that are markedly superior to current offerings. These models were never designed to study rivalry to develop new and improved products or processes. The effect of changing a parameter measuring the degree of differentiation among products is just not directly relevant to competition policy.[11]

For Aghion and Griffith, this discussion is merely a launching pad, and I do not mean to suggest that they base any of their conclusions or policy recommendations on these simple static oligopoly models. Indeed, they immediately go on to note two important and powerful forces missing from these models: "the interplay between rent dissipation and preemption incentives, and the differences between vertical (i.e., quality improving) and horizontal innovations" (13). Nonetheless, their framing of the issues is indicative of how the conversation has developed, and how research findings are trans-

10. Similarly, Aghion et al. (2005), summarizing the "main existing theories of competition and innovation," states: "The leading IO models of product differentiation and monopolization . . . deliver the prediction that more intense product market competition reduces postentry rents, and therefore reduces the equilibrium number of entrants" (710, footnoted omitted).

11. Baker (2007) puts this nicely: "antitrust is not a general-purpose competition intensifier. Rather, antitrust intervention can be focused on industry setting and categories of behavior where enforcement can promote innovation" (589, footnote omitted).

lated and conveyed to policymakers. They summarize the "early theoretical and empirical literatures" as follows: "theory pointed to a detrimental effect of competition on innovation and growth, while the empirical literature instead suggested that more competitive market structures are associated with greater innovative output, an idea that had much support in policy circles" (3–4).

These passages from Aghion and Griffith (2005) accurately reflect a strand of the theoretical literature that equates the concept of "more competition" with "less product differentiation." For much more detail on these models, see Boone (2000) and (2001), Aghion, et al. (2001), and Sacco and Schmutzler (2011). Vives (2008), "Innovation and Competitive Pressure," surveys and synthesizes this literature.[12] Schmutzler (2010) uses a generalized "competition parameter." By definition, increases in this parameter lead to lower equilibrium profit margins and a greater sensitivity of a firm's equilibrium output to that firm's cost level. Schmutzler explores the relationship between "more competition," as defined by increases in this parameter, and the level of R&D investment. While there is nothing inherently incorrect or misleading about modeling "more competitive pressure" as "less product differentiation," defining "more competition" this way can lead to statements about competition and innovation that are unhelpful or even misleading for merger enforcement policy.

In particular, the statement that "more competition discourages innovation" can be misused or misunderstood in the context of competition policy, or innovation policy more broadly. The statement, "innovation incentives are low if ex post competition is so intense that even successful innovators cannot earn profits sufficient to allow a reasonable risk-adjusted rate of return on their R&D costs" strikes me as more defensible and far more accurate, if less pithy. I doubt these conditions are common, except perhaps when appropriability is low, in which case the root problem is one of low appropriability, not excessive competition. But at least this far more precise statement is not misleading.

Clarity and precision in defining "competition" can reduce perceived complexity regarding the impact of competition on innovation.

7.2.2 Equating "More Competition" with "More Imitation"

The endogenous growth literature also explores the relationship between competition and innovation, albeit from a different perspective. See Aghion and Griffith (2005) and Aghion and Howitt (2009).[13] The paper by Aghion

12. In an oligopoly model with restricted entry, Vives also studies the relationship between the number of firms and innovation. This measure of competition is more relevant to merger enforcement policy, as discussed in section 7.5.

13. In Aghion and Howitt (2009), see especially chapter 12, "Fostering Competition and Entry," and the references therein. For a recent survey on this literature, see Scopelliti (2010).

et al. (2005), "Competition and Innovation: An Inverted-U Relationship," has been especially influential. The model used by Aghion et al. (2005) is far better for considering innovation than are the static oligopoly models just discussed, because it is a dynamic model in which firms invest to develop new and superior products.

However, as I now explain, this strand of literature typically equates "more competition" with "more imitation." This has led to the unfortunate sound bite, typically paired with a reference to Schumpeter, that "greater competition discourages innovation." Aghion et al. (2005) write:

> [I]ncreased product market competition discourages innovation by reducing postentry rents. This prediction is shared by most existing models of endogeneous growth. . . . where an increase in product market competition, or the rate of imitation, has a negative effect on productivity growth by reducing the monopoly rents that reward new innovation. (Aghion et al. 2005, 711, footnote omitted)

The standard growth-theoretic models that explore the competition/innovation relationship model "more competition" as a parameter that shifts downward the ex post demand function facing the innovator. They do not model "more competition" as an increase in contestability or appropriability. Instead, "more competition," meaning more imitation, involves *reduced* appropriability and thus lower profit margins for the innovator.

To see how this literature models competition, consider the benchmark model of innovation and productivity growth presented by Aghion and Griffith (2005).[14] In that model, "competition" is measured by the cost advantage of an innovator over a competitive fringe of imitators. I regard this as a measure of the strength of intellectual property protection, or as a measure of the spillovers associated with innovation. It is certainly not a measure of contestability or rivalry to innovate and thus win sales. Clearly, more "competition" in the sense used in this literature equates to less appropriability. It is entirely unsurprising that imitation reduces innovation incentives. Unfortunately, Aghion and Griffith (2005) interpret this finding as follows:

> However, *pro-competition policies* will tend to discourage innovation and growth by reducing χ [the cost advantage of the innovator over the imitators], thereby forcing incumbent innovators to charge a lower limit price. (Aghion and Griffith 2005, 18, emphasis supplied)

So far as I can tell, these so-called "procompetition policies" involve weaker intellectual property rights, or perhaps mandatory licensing or price controls, neither of which can properly be called "competition policies," at

14. See pp. 16–19: "This serves as a basis for the theoretical extension we will present in later chapters of this book."

least in the United States today.[15] But, unfortunately, the idea sticks: competition and procompetition policies discourage innovation and growth.

Aghion and Griffith do not rest at this point and conclude that competition discourages growth. To the contrary, they press forward, seeking to reconcile theory and evidence, emphasizing what they call the "escape competition effect." In my lexicon, this is a form of contestability: a firm that fails to innovate will find its margins squeezed, while innovating preserves these margins. However, their extension models also equate "more competition" with more complete imitation of a process innovation. For that reason, their analysis strikes me as far more relevant for policies that influence the strength of intellectual property rights than for competition policy.[16]

Let me be clear: there is nothing inherently incorrect about modeling "more competition" as "more imitation." Imitation does reduce the demand facing an innovator, and it certainly constitutes "more competition" from that firm's perspective. Furthermore, imitation can be a very important consideration when firms make R&D investment decisions, especially for product or process innovations that are difficult to protect using patents or trade secrets.[17]

Nonetheless, the statement that "more competition discourages innovation" can all too easily be misunderstood or misused in the context of competition policy, not to mention innovation policy more broadly.[18] The statement, "more rapid and complete imitation tends to discourage innovation" seems more reasonable and far more accurate.

Clarity and precision in defining "competition" can reduce perceived complexity regarding the impact of competition on innovation.

7.2.3 Equating "More Competition" with "Lower Market Concentration"

Industrial organization economists have long used product market concentration as a proxy for competition, with higher concentration indicating less competition with respect to price and output. We place less weight on this proxy than we did fifty years ago, but it certainly still has value, at least in

15. The impact of imitation on innovation and economic growth is certainly important for policies governing the design and strength of patent rights, as well as policies affecting the protection and enforcement of trade secrets. That discussion is beyond the scope of this chapter. Shapiro (2007) discusses the relationship between the reward to a patent holder and the patent holder's contribution.

16. Of course, as reflected in the Appropriability principle, imitation and spillovers can be very important in antitrust analysis. In particular, a merger that internalizes significant spillovers may promote innovation, as discussed later.

17. Patents and trade secrets are the most relevant forms of intellectual property for the product and process innovations I have in mind here. However, the same argument can be made for creative works, where copyrights typically are the applicable form of intellectual property.

18. Aghion and Griffith (2005) clearly believe their work is relevant to competition policy. In the conclusion to chapter 3, they state: "These predictions have important policy implications for the design of competition policy" (64).

properly defined relevant markets. The recently revised Horizontal Merger Guidelines continue to use Herfindahl-Hirschman Index (HHI) thresholds, with adverse competitive effects viewed as unlikely in markets with a post-merger HHI less than 1,500 and presumed likely for mergers that raise the HHI more than 200 and lead to a postmerger HHI greater than 2,500.

The link between current or recent product market concentration and R&D rivalry has always been weaker than the link between current or recent product market concentration and rivalry to win current sales. A firm's current sales may not be a good proxy for that firm's R&D incentives and abilities. Plus, R&D expenditures normally have the character of a fixed cost, leading to scale economies. If those fixed costs are large relative to sales, significant market concentration is inevitable in equilibrium, as demonstrated by Sutton (1998). Furthermore, as Schumpeter emphasized, the reward to successful innovation is some degree of market power in the technical sense—price above marginal cost—for a sufficient volume of sales to earn a risk-adjusted return on the fixed and sunk R&D costs. Plus, a highly successful innovator may come to dominate a market, in which case observing a high level ex post concentration would hardly imply a lack of ex ante competition, or a lack of innovation. In an industry where innovation has recently occurred, or is ongoing, any measurement of the current or recent market structure inevitably will be a post-innovation measurement. We should not expect to see atomistic market structures in industries that have experienced significant technological progress, and we may see high levels of concentration in markets that have recently experienced significant innovation.

The empirical literature on product market structure and competition has come to recognize all of these points, and recent work (see the following) attempts heroically to account for them. Cohen (2010) summarizes: "Regarding measures, there can be little disagreement with Gilbert's contention that the commonly employed measure of market structure, market concentration, does not accurately reflect the nature or intensity of competition" (156). Yet there remains some tendency to equate "more competition" with "lower product market concentration." Thus, a finding that unconcentrated markets (or markets where firms earn low operating profits relative to sales) are not the ones where we see the most experienced significant innovation may be interpreted—incorrectly—as "too much competition discourages innovation," or as implying that "an intermediate amount of competition is best for promoting innovation." The real lesson is that static measures of market structure can be poor metrics for assessing innovation competition.

Framing the relationship between competition and innovation as one between product market concentration and competition is not dissimilar to the view in the 1950s and 1960s that atomistic markets were the ideal and best promote (pricing and output) competition. That view gave way

long ago to a more nuanced one, which recognizes that when individual firms differ greatly in their efficiency (as they normally do), and when there are significant economies of scale (as there typically are in markets where antitrust enforcement takes place), robust competition is likely to lead to a market structure in which some firms have substantial market shares. Demsetz (1973) powerfully and influentially articulated this point. This very same principle applies with even greater force to innovation: we know there are significant economies of scale, because R&D is a fixed cost, and it would be surprising indeed if firms did not differ substantially in their ability to innovate. Accounting for the inherent uncertainty associated with innovation only strengthens the point: even if several firms have comparable ex ante incentive and ability to innovate, ex post some will strike gushers and others just dry wells.

Again, there is nothing inherently wrong with observing and reporting that many highly innovative industries do not have atomistic market structures: it is helpful to know not to expect, or strive for, atomistic market structures in those industries.[19] But there is no tension between established competition policy principles and the Schumpeterian observation that successful innovators often are able to price well above marginal cost and often gain substantial market shares. The US antitrust law has understood for a very long time that the market power resulting from successful innovation is an important and inevitable part of the competitive process. As Judge Learned Hand famously observed: "the successful competitor, having been urged to compete, must not be turned upon when he wins."[20] Furthermore, of course, merger enforcement policy does not strive for atomistic markets: under the recently revised Horizontal Merger Guidelines, merger adverse competitive effects are considered unlikely if the post-merger HHI is less than 1,500, and the merger enforcement statistics show that the Department of Justice (DOJ) and the FTC often allow horizontal mergers, leading more concentrated markets to proceed without challenge.

7.3 What Does the Empirical Evidence Really Tell Us?

There is a very substantial body of empirical evidence supporting the general proposition that "more competition," meaning greater contestability of sales, spurs firms to be more efficient and to invest more in R&D. For

19. Even in concentrated industries, start-up firms can play a very positive and powerful role in spurring innovation. If they are rapidly acquired by large incumbents, or if their ideas are copied by large incumbents, their role may never be reflected in a decline in market concentration. Even if antitrust does not stand in the way of mergers that cause moderate increases in concentration, it may need still to intervene to protect customers from unilateral conduct by dominant firms that stifles disruptive innovation by start-up firms.

20. US v. Aluminum Company of America, 148 F.2d 416 (1945).

our purposes, "innovation" encompasses a wide range of improvements in efficiency, not just the development of entirely novel processes or products.

Detailed case studies of businesses operating in diverse settings almost invariably conclude that companies insulated from competition—that is, firms operating in environments in which relatively few sales are contestable—are rarely at the cutting edge in terms of efficiency and can be woefully inefficient. Porter (1990) assembles a raft of evidence showing that companies protected from international competition tend to fall behind and lose their ability to compete in export markets. Porter has long emphasized the importance of competition in spurring innovation, as reflected in this passage from Porter (2001):

> Innovation provides products and services of ever increasing consumer value, as well as ways of producing products more efficiently, both of which contribute directly to productivity. Innovation, in this broad sense, is driven by competition. While technological innovation is the result of a variety of factors, there is no doubt that healthy competition is an essential part. One need only review the dismal innovation record of countries lacking strong competition to be convinced of this fact. Vigorous competition in a supportive business environment is the only path to sustained productivity growth, and therefore to long term economic vitality. (Porter 2001, 923)

In another wide-ranging international study, Lewis (2004) finds that competitive markets are the key to economic growth. His central conclusion is that competition drives innovation:

> Most economic analysis ends up attributing most of the differences in economic performance [across countries] to differences in labor and capital markets. *This conclusion is incorrect. Differences in competition in product markets are much more important.* (Lewis 2004, 13, emphasis in original)

In discussing the relationship between competition and innovation, it is important to bear in mind the enormous differences across firms in their efficiency, even among firms in the same industry. Bartelsman and Doms (2000) survey the literature on firm-level productivity, writing:

> Of the basic findings related to productivity and productivity growth uncovered by recent research using micro-data, perhaps most significant is the degree of heterogeneity across establishments and firms in productivity in nearly all industries examined. (Bartelsman and Doms 2000, 578)

In a more recent survey, Syverson (2011) starts by stating: "Economists have shown that large and persistent differences in productivity levels across businesses are ubiquitous." He reports studies (35–48) showing how competition acts to improve productivity both through a Darwinian selection

effect and by inducing firms to take costly actions to raise their productivity. He also reports studies showing how additional competition arising from trade liberalization enhances productivity. These are first-order effects that serve to remind us that the relevant notion of "innovation" is quite broad, encompassing the adoption and diffusion of best practices. Innovation is not confined to the invention of new products or new methods of production.

Leibenstein (1966) famously asked why so many firms are operated inefficiently and thus appear not to maximize profits. Economic theory has yet to fully explain why firms fail to undertake what appear to be profitable investments to improve their efficiency, but empirical evidence consistently shows that firms are more likely to make such investments when placed under competitive pressure.[21] Holmes, Levine, and Schmitz (2008) argue creatively that competition spurs innovation by reducing margins on existing products and thus reducing the opportunity cost of innovation that involves "switchover disruptions" for suppliers.

Numerous studies show specifically that increased competitive pressure resulting from lower regulatory barriers to entry generally enhances productivity and accelerates innovation. Holmes and Schmitz (2010) provide a recent review of a number of these studies, concluding:

> Nearly all the studies found that increases in competition led to increases in industry productivity. Plants that survived these increases in competition were typically found to have large productivity gains, and these gains often accounted for the majority of overall industry gains. (Holmes and Schmitz 2010, 639)

Syverson (2004) is especially instructive regarding the relationship between competitive pressure and firm-level efficiency. Studying the concrete industry, he shows that average productivity is higher, and productivity differences across firms are smaller, in local markets that are more competitive. Here "more competitive" means that the producers are more densely clustered, increasing spatial substitutability. Syverson finds that relatively inefficient firms in the concrete industry have greater difficulty operating in the more competitive local markets.

In contrast to Syverson's in-depth study of one industry, Bloom and Van Reenen (2007) examine management practices across a wide range of industries by surveying managers from over 700 medium-sized firms. They find very large differences in productivity across firms and conclude that "poor management practices are more prevalent when product market competition is weak." They explain that

> [H]igher levels of competition (measured using a variety of different proxies, such as trade openness) are strongly associated with better management practices. This competition effect could arise through a number of

21. See the survey by Holmes and Schmitz (2010), as well as the other surveys cited earlier.

> channels, including the more rapid exit of badly managed firms and/or the inducement of greater managerial effort. (Bloom and Van Reenen 2007, 1351)

Similarly, Bloom and Van Reenen (2010) observe that "firms with 'better' management practices tend to have better performance on a wide range of dimensions: they are larger, more productive, grow faster, and have higher survival rates" (204–205). They report that strong product market competition appears to boost average management practices through a combination of eliminating the tail of badly managed firms and pushing incumbents to improve their practices.

In addition to these studies, which collectively are quite convincing, there is a very large empirical literature examining the relationship between (a) firm size and innovation, and (b) product market concentration and innovation. Cohen (2010) surveys this literature.[22]

Regarding business unit size and innovation, Cohen writes:

> Thus, the robust empirical patterns relating to R&D and innovation to firm size are that R&D increases monotonically—and typically proportionately—with firm size among R&D performers within industries, the number of innovations tends to increase less than proportionately than firm size, and the share of R&D effort dedicated to more incremental and process innovation tends to increase with firm size. (Cohen 2010, 137)

As Cohen explains (138), these findings are consistent with the view that larger business units expect to be able to apply process innovations over a larger scale of output, because firms chiefly exploit their process innovations internally and often anticipate limited growth due to innovation. In contrast, Cohen writes that "the returns to more revolutionary (i.e., substitute) innovations are less tied to a firm's prior market position" (139).

Regarding the connection between market power and innovation, Cohen observes: "The empirical literature has focused principally on the effects of market concentration on innovative behavior. The literature has thus directly tested Schumpeter's conjectures about the effects of ex ante market structure" (140). Cohen further notes that "the potential for achieving ex post market power through innovation has been characterized under the general heading of appropriability conditions and measured by survey-based indicators of appropriability" (141). Cohen is thus careful to avoid conflating "more competition" with "more imitation."

Lee (2005) offers this view of a key stylized factoid that has long captured the imagination of industrial organization economists:

> The conventional wisdom from the literature postulates an inverted-U relationship between market structure, measured by seller concentration on the horizontal axis, and industry R&D intensity (i.e., R&D-to-sales

22. See Cohen and Levin (1989) and Cohen (1995) for earlier surveys of this literature.

> ratio) on the vertical axis. The inverted-U hypothesis says that moderately concentrated industries engage more intensively in R&D activity than either atomistically competitive or highly concentrated industries. (Lee 2005, 101)

This inverted-U shaped relationship between market concentration and innovation has not held up well under scrutiny, especially after correcting for industry differences in technological opportunity and for the endogeneity of product market structure. I do not intend to wade into that debate, which I do not expect to be resolved definitely one way or the other during my lifetime, either theoretically or empirically, for the reasons given earlier. Meanwhile, the message received by nonspecialists and policymakers is that we know rather little about the relationship between "competition" and innovation, notwithstanding the very powerful evidence about firm-level productivity cited before.

Lee (2005) distinguishes industries based on appropriability and emphasizes that the notions of "more competition" and "more imitation" are very different:

> [T]he concentration-R&D relationship differs depending on the strength of the link or simply the appropriability of R&D in terms of market share: A positive relationship is predicted for low-appropriability industries, where market concentration supplements low R&D appropriability, while a negative or an inverted U-shaped relationship for high-appropriability industries. An empirical analysis of data, disaggregated at the five-digit SIC level, on R&D and market concentration of Korean manufacturing industries provides supportive evidence for the predictions. (Lee 2005, 101)

Attempting to move the debate forward, and recognizing the limitations of market concentration as a proxy for the intensity of competition, the empirical literature has made progress in using measures other than market concentration as a proxy for the intensity of competition. Notably, Nickell (1996) uses a modified Lerner Index as a proxy for competition.[23] Nickell states: "I present evidence that competition, as measured by increased numbers of competitors or by lower levels of rents, is associated with a significantly higher rate of total factor productivity growth (724)."[24] More recently, Aghion et al. (2005), also using a modified Lerner index as their

23. Nickell also uses results from a one-time survey in which management was asked whether it had more than five competitors in the market for its product. He discusses the limitations of his proxies for competition (732). Nickell also uses a measure of market share, with three-digit industry sales in the denominator. Nickell notes that "the three digit industry does not represent anything like a 'market,'" and thus has little value as a cross-section measure of market power, but he argues that it is useful as a time-series measure.

24. Blundell, Griffith, and Van Reenen (1999) state: "We find a robust and positive effect of market share on observable headcounts of innovations and patents although increased product market concentration in the industry tends to stimulate innovative activity" (529). They measure innovation by counting the number of technologically significant and commercially

measure of competition, have challenged Nickell's conclusions. They find instead an inverted U-shaped relationship between product market competition and innovation.

> This paper investigates the relationship between product market competition and innovation. We find strong evidence of an inverted-U relationship using panel data. We develop a model where competition discourages laggard firms from innovating but encourages neck-and-neck firms to innovate. (Aghion et al. 2005, 701)

Aghion et al. (2005) look at two-digit Standard Industrial Classification (SIC) industries. They measure innovation using the number of citation-weighted patents. Their measure of the Lerner Index averages 4 percent, and generally falls between zero and 10 percent, with the peak of the inverted-U occurring at a Lerner index of around 5 percent. Whatever one makes of these findings, they do not challenge the extensive empirical evidence cited earlier about innovation and firm-level efficiency. Nor do they call into question the Contestability, Appropriability, and Synergies principles. In any event, they are not directly relevant to analyzing the effects of proposed mergers on innovation.

Cohen (2010) reports a number of other studies that support the general proposition that greater competitive pressure spurs firms to innovate to get ahead of their rivals. For example, he notes that "Lee (2009), using World Bank survey data for nine industries across seven countries, finds that intensity of competition may stimulate more capable firms to invest more heavily in R&D, while less capable firms may invest less" (16). Of special relevance for competition policy, Cohen reports work suggesting that entry causes innovation (144). However, this is a tricky area empirically, since high technological opportunity in an industry tends to cause both more entry and faster innovation in that industry. In summarizing the literature on market structure and innovation, Cohen (2010) states: "Moving on to our consideration of the relationship between market structure and R&D, the empirical patterns are mixed, and not terribly informative" (154). Again, this is unsurprising, given what we can measure, given the endogeneity of market structure, and given that increased market concentration may or may not go along with greater contestability.

Of particular interest here, Gilbert (2006) provides an extensive discussion of what this empirical literature implies for competition policy. As he points out, product market concentration is "a commonly used, but highly imperfect, surrogate for competition" (187). I note in particular that relevant antitrust markets do not match up well with the publicly available sales data,

important innovations. They define an industry at the three-digit level. Their metrics for competition are the proportion of industry sales made by the five largest domestic firms and the value of imports in proportion to home demand.

making the measurement of meaningful market shares difficult or impossible for academic researchers. Two-digit SIC industries are very far indeed from relevant antitrust markets. Likewise, academic researchers often have difficulty measuring true economic operating profits or price/cost margins using publicly available accounting data.

Gilbert concludes that these studies have failed to establish a general and robust relationship between product market concentration and innovation, once one controls for the underlying technological environment.

> Empirical studies that use market concentration as a proxy for competition fail to reach a robust conclusion about the relationship between market concentration and R&D when differences in industry characteristics, technological opportunities, and appropriability are taken into account. (Gilbert 2006, 206)

Gilbert notes several reasons for these negative results: limited data on innovative activity and market concentration, including the high level of aggregation at which market concentration is usually measured; failure to distinguish exclusive from nonexclusive property rights and between product and process innovations; differences in technological opportunities across industries and over time; and failure to control for other confounding factors.

The lack of robust results in this particular line of empirical work is understandable, given the measurement difficulties and conceptual complexities already discussed. However, given the very extensive empirical evidence showing that competitive pressure forces firms to be more efficient, and given the robust theoretical points relating innovation incentives to the contestability of future sales, the negative results in this particular area should not be interpreted as implying that "we just don't know anything about the relationship between competition and innovation." To the contrary, the empirical evidence overall gives powerful support for the proposition that heightened competitive pressure causes firms to invest more to improve their efficiency. Another advantage of having multiple firms seeking to innovate in a given area is that such decentralization supports greater innovation diversity.[25]

7.4 Competition and Innovation: Toward Robust Principles

When considering the impact of competition on innovation, rather than equating "more competition" with "less product differentiation," "more imi-

25. Even if overall profit maximization at the dominant firm entails pursuing multiple distinct approaches to developing next-generation products, organizational obstacles to doing so can be significant, especially when opinions differ greatly about which approach is most promising. Grove (1996) explains how Intel found it very difficult to pursue two distinct microprocessor architectures, CISC and RISC, at the same time. Christensen (1997) discusses the limitations of "skunk works."

tation," or "lower product market concentration," I suggest that the term "more competition" be reserved for market characteristics that correspond greater *rivalry* to serve the needs of customers. This is how the concept of "more competition" is generally applied in the area of competition policy: the competitive process is working well if there is healthy rivalry, on the merits, to win the patronage of customers by offering them superior value. Effective competition is about the *competitive process,* not the outcome. More important than terminology, assessing competition based on rivalry allows us to articulate and employ practical principles regarding innovation that are theoretically and empirically robust.

Rivalry in the current context is driven by the incentive and ability of firms to engage in innovation, broadly defined to include increased efficiency as well as the development of entirely new products and processes. The Contestability and Appropriability principles relate to innovation incentives, and the Synergies principle relates to innovation ability.

What basic factors govern an individual firm's incentive to innovate? Consider the following highly simplified model of the impact on a given firm's operating profits if that firm achieves a given product or process innovation. For simplicity, suppose the firm produces a single product, whether or not it innovates, although the firm will offer an improved product if it succeeds in achieving a product innovation. Denote the product's price by P, its output by X, and its (constant) marginal cost by C, so the profit margin on incremental units is given by $M \equiv P - C$. The firm's operating profits are $\pi = (P - C)X = MX$.[26] Whether or not the firm in question successfully innovates, it sets its price to maximize its operating profits.

Let the subscript "0" denote the situation in which the firm does not successfully innovate, and the subscript "1" denote the situation in which the firm does successfully innovate. The "no innovation" state will typically not be the pre-innovation status quo, since *other* firms may well successfully innovate even if the firm in question does not. This allows us to account for the added competitive pressure faced by the firm in question if it fails to innovate and its rivals succeed: X_0 and/or M_0 are reduced when rivals innovate. The "innovation" state incorporates rivals' reactions to the firm's innovation, including price adjustments and imitation. This setup allows us to examine the innovation incentives facing one firm, given the actions and reactions of other firms in terms of their own pricing, product offerings, efficiency, and R&D investments.

Successful innovation increases the firm's profits by $\Delta\pi = \pi_1 - \pi_0 = M_1X_1 - M_0X_0$, which can be written as

$$\Delta\pi = X_0\Delta M + M_0\Delta X + \Delta M\,\Delta X.$$

26. By operating profits I mean profits gross of R&D expenses and other costs that are fixed with respect to the firm's output level over the relevant time frame.

This expression for $\Delta\pi$ is simple, and not deep, but it does serve to remind us of the basic factors at play that govern the firm's innovation incentives.

The first term reflects the extra margin the firm earns as a result of innovating. These margins are applied to the firm's without-innovation output level, X_0. This extra margin can come from lower costs (for a process innovation) and/or from a higher price (for a product innovation). This term encompasses the "escape competition" effect in the literature.

The second term reflects the extra unit sales the firm makes by successfully innovating. These sales are valued at the firm's without-innovation margin, M_0. Other things equal, a firm that would make substantial sales *without* innovating will have a smaller sales boost from innovating, ΔX, and thus a smaller incentive to innovate. This is the Arrow replacement effect at work.

The third term is a positive interaction term between higher incremental margins and higher incremental sales. Since the firm picks its own price, the firm can choose how best to capture the rewards from innovation, as between higher margins and greater unit sales.

If successful innovation will do little to increase the firm's unit sales, ΔX is small and we have $\Delta\pi \approx X_0\Delta M$. Under these conditions, initially larger firms have greater incentives to innovate. This is a standard observation in the literature: the benefit of lowering marginal cost is proportional to output.[27] These conditions tend to apply when demand is sticky, so one firm cannot gain many sales even as a result of successful innovation, or when the innovating firm faces lasting capacity constraints. A similar situation arises for process innovations if the firm's rivals would react strongly (were the firm to lower its prices) by lowering their own prices. In that situation, the firm in question will gain few sales by lowering its own price, so the firm will tend to take the rewards from innovation in the form of higher margins on existing sales, rather than by lowering its price to expand its sales; this too implies that $\Delta\pi \approx X_0\Delta M$.

Additional insights can be obtained by examining how the firm's operating profits are boosted by incremental innovation. Denote by θ the level of innovation achieved by the firm. The innovation can involve an improvement in efficiency, or a process innovation, either of which lowers the firm's cost. To capture this, we write the firm's marginal cost as $C(\theta)$, with $C'(\theta) < 0$. The innovation also can involve an increase in the quality of the firm's product. To capture this, we write the firm's demand as $D(P,\theta,z)$, where $D_\theta(P,\theta,z) > 0$.

The variable z in $D(P,\theta,z)$ captures the attractiveness of the products

27. The simple formulation used here does not include licensing revenues. Licensing breaks the connection between the firm's own sales and the base on which higher margins can be earned.

offered by the firm's rivals, so $D_z(P,\theta,z) < 0$. Rivals can react to the firm's price and level of innovation, so $z = z(P,\theta)$. Successful innovation can weaken the firm's rivals, or even drive them from the market, so z_θ can be negative. However, we are more interested here in situations in which the rivals respond to the firm's innovation by improving their own offerings, either by lowering their prices or improving their own products (perhaps through imitation), in which case we have $z_\theta > 0$.

The firm's profits are given by $\pi(P,\theta) = D(P,\theta,z(P,\theta))(P - C(\theta))$. Applying the envelope theorem to the firm's price, achieving marginally more innovation raises operating profits by

$$\pi_\theta(P,\theta) = D(P,\theta,z)|C'(\theta)| + (P - C(\theta))[D_\theta + D_z z_\theta].$$

The first term in this expression captures the margin boost resulting from lower costs. The benefit of lower costs is proportional to the firm's output. The second term captures the sales boost resulting from product improvement. The impact of these incremental sales on profits is proportional to the gap between the firm's price and marginal cost. The sales boost consists of two terms: (1) the D_θ term reflects the increased demand given the prices and product offerings of rivals, and (2) the $D_z z_\theta$ term reflects rivals' responses to the firm's innovation.

We next show how the Contestability and Appropriability principles relate to this expression.

7.4.1 Contestability

The Contestability principle focuses on the ability of an innovating firm to gain or protect profitable sales by providing greater value to customers. This principle directs our attention to the incremental profits associated with innovation, taking as given the price and product offerings of other firms. Holding z fixed at $\bar{z}$, the incremental profits resulting from innovation are

$$\pi_\theta(P,\theta)\big|_{z=\bar{z}} = D(P,\theta,\bar{z})|C'(\theta)| + (P - C(\theta))D_\theta(P,\theta,\bar{z}).$$

The first term is the standard benefit to the firm from lowering its costs, which is proportional to the firm's output level. The second term is the boost in the firm's unit sales as a result of offering a better product, multiplied by the firm's price/cost margin. This second term captures the fundamental idea that a firm has greater innovation incentives if successful innovation allows the firm to gain, or protect, profitable sales. Sales are highly contestable—in the sense relevant for innovation—if a firm that provides greater value to customers gains substantial unit sales from its rivals; that is, if $D_\theta(P,\theta,\bar{z})$ is large.

An unconcentrated market is highly contestable if an innovator can gain substantial market share at a healthy margin by providing a better product or setting a lower price. In contrast, for product innovations, an unconcentrated market is not highly contestable if customers exhibit strong brand

preferences, or have high switching costs, so any one firm that develops an improved product will gain few sales from its rivals.

The Arrow "replacement effect" is driven by contestability. In Arrow (1962), innovation allows a firm initially operating in a highly competitive market to take over the entire market at a margin reflecting its cost advantage. In contrast, the incumbent monopolist has far fewer sales to gain from innovation, and its without-innovation sales are not at risk since (by assumption) only the monopolist can innovate, so contestability is far lower.[28]

The robustness of the Contestability principle is nicely illustrated by seeing how it fares in the model of continual process innovation used by Aghion et al. (2005). They use this model to argue for an inverted U-shaped relationship between competition and innovation. Such a nonmonotonic relationship might appear to defy the Contestability principle. It does not.

In the Aghion et al. model, each industry is a duopoly, with no possibility of entry. The two firms sell a homogeneous product, so the only possible source of competitive advantage is a cost advantage. The duopolists can invest in R&D to lower their costs; such process innovations come in discrete steps. At any point in time, if the two firms have equal costs, the industry is said to be "neck-and-neck." Aghion et al. assume that spillovers allow a firm falling two steps behind immediately and costlessly to narrow the gap to one step, so the only other possible state of the market is for one firm to be the leader and the other the laggard, one step behind. This assumption also implies that the leader never invests in R&D, since it cannot extend its lead and since the leader's profits only depend upon the cost gap between the two firms, not on their absolute cost levels.

Aghion et al. state: "We define the degree of product market competition inversely by the degree to which the two firms in a neck-and-neck industry are able to collude" (713). A neck-and-neck firm has a stronger incentive to innovate, the greater the degree of product market competition. They call this the "escape the competition" effect, which I think of as the flip side of the Arrow replacement effect. In contrast, a laggard firm has a *weaker* incentive to innovate, the greater the degree of product market competition, since successful innovation leads to the less profitable neck-and-neck state. They call this a Schumpeterian effect. Aghion et al. cleverly exploit these mixed effects to obtain an inverted U-shaped relationship between equilibrium steady-state innovation rates (aggregated across many sectors) and the degree of product market competition (i.e., inability to collude). The model is elegant and instructive—major virtues in my view—but it is worth noting some of the strong assumptions underlying its prediction of

28. In Arrow's model, only a single firm can innovate, so the incumbent monopolist faces no danger of losing its monopoly if it is the designated innovator. If the monopolist can be dethroned, it has highly profitable sales to protect by innovating first; this is the central point in Gilbert and Newbery (1982), who allow for innovation rivalry.

the inverted U-shaped relationship between competition and innovation: there are only two firms in each industry, with no possibility of entry; the two firms sell a homogeneous product; the laggard firm cannot innovate in a different direction, for example, to differentiate its product, or take a riskier approach that might leapfrog the leader; and (due to imitation) the leader does not benefit at all from further lowering its costs.

Whether or not these conditions are realistic, the basic forces modeled by Aghion et al. (2005) fit comfortably with the Contestability principle. In particular, the inverted U-shaped relationship they uncover between "competition" and innovation does *not* correspond to a nonmonotonic relationship between contestability and innovation. In their model, "more competition" means less effective collusion when the duopolists are neck-and-neck. Their notion of "more competition" translates to more contestability when the firms are neck-and-neck: each firm has *more* to gain from pulling ahead, the more vigorously the two firms are competing. However, critically, their notion of "more competition" translates to *less* contestability when the firms are in the leader/laggard state: the laggard (the only innovator in this state) earns zero profits regardless of the degree of competition and *smaller* profits by catching up, the more vigorously the firms compete when neck-in-neck. Both states in their model are perfectly consistent with the Contestability principle.

This is a good point to elaborate on the connection between the notion of "more competition" and the operation of competition policy. Taking the Aghion et al. (2005) model at face value, it suggests that allowing some degree of collusion is desirable to spur innovation because it provides greater incentives to laggard firms to catch up so they can collude with their rival. However, in their model allowing a great deal of collusion is undesirable for innovation because the duopolists would then be more content to rest comfortably once they are neck-and-neck and effectively colluding. I am not aware of anyone actually proposing such a policy toward collusion, and for good reason. Among other problems, if the firms were given latitude to communicate and collude, they might also find a way to maximize joint profits by agreeing to stop spending money on R&D. In any event, a more relevant question for competition policy is whether reducing competition by allowing the two firms to *merge* would accelerate or retard innovation. In the Aghion et al. model, a merger between the two firms would be disastrous for innovation. Assuming that knowledge spillovers continue to limit the merged firm's competitive advantage to one step, the merged firm would immediately cease all innovation and coast along indefinitely with a one-step advantage over the imitating fringe.

7.4.2 Appropriability

Any analysis of competition and innovation needs to pay close attention to the conditions of appropriability; that is, the extent to which inno-

vators can appropriate the social benefits their innovations have caused. The conditions of appropriability can greatly affect innovation incentives.[29] Appropriability is heavily influenced by the strength of intellectual property rights. Appropriability for a given firm is reduced by spillovers to noninnovating firms (e.g., through imitation). Some causal factors, such as low entry barriers combined with weak intellectual property rights, can lead to both more competition and more imitation. But any analysis of competition and innovation should avoid conflating "low appropriability" with "more competition."

Define the appropriability ratio as $\alpha \equiv (d\pi/d\theta)/(dW/d\theta)$, where W measures total welfare. In a model of a single firm, the appropriability ratio is less than unity under the mild condition that customers benefit from the innovation. In the special case of a single firm offering one product at a uniform price, textbook monopoly pricing theory tells us that the appropriability ratio for innovations that lower the firm's marginal cost is given by $\alpha = 1/(1 + \rho)$, where $\rho = dp/dc$ is the rate at which the firm passes cost changes through to price changes. However, analysis of the appropriation ratio becomes much more complex when the firm offers multiple products, engages in price discrimination, or faces rivals that are not price-taking firms. Fortunately, for the purposes of antitrust analysis, we typically do not need to measure the appropriability ratio; we are more interested in whether a particular merger or business practice internalizes important spillovers and thus increases appropriability.

The Appropriability principle builds on the Contestability principle by taking into account how rivals will respond to a given firm's successful innovation. In some cases, rivals respond passively by reducing their own R&D efforts or even exiting the market, adding to the rewards to the successful innovator.[30] In other cases, rivals respond aggressively by lowering their price or redoubling their own innovative efforts, either (1) improving their own efficiency, thus lowering their costs and their price, or (2) making their own product improvements, perhaps by imitating the first firm. In such cases, appropriability is reduced because the total benefits caused by the firm's innovation are larger and because that firm's rewards are reduced according to the $(P - C(\theta))D_z z_\theta$ term, which is negative if rivals improve their product offerings in response to the firm's innovation; that is, if $z_\theta > 0$.

Aggressive rivals' responses reduce appropriability by shifting the benefits of innovation to rivals and/or to customers. For example, if rivals will quickly imitate the product improvements introduced by a pioneering firm, that firm may gain little from leading the way. If that product improvement

29. Increasing appropriability for one firm can reduce it for others, especially when multiple innovating firms supply complements.

30. Responses of this type can create or bolster business-stealing effects leading to an appropriability ratio in excess of unity.

would not have been introduced without the pioneering firm taking the lead—a critical qualification—then appropriability is low in this situation. In this example, many of the benefits of innovation will flow to customers, or to suppliers of complements, not to the pioneering firm. However, one must be careful not to take this argument too far: if several firms are introducing a certain type of product improvement, little or none of the social benefits associated with that improvement are properly attributable to any one of those firms and ongoing competition to offer that improvement does not indicate any lack of appropriability.

Appropriability can be enhanced by mergers or business practices that internalize positive externalities, aka spillovers. Spillovers can occur between direct rivals through imitation, so these considerations can come into play in the analysis of horizontal mergers. Spillovers also arise between suppliers of complements, in which case the Appropriability principle reinforces the Synergies principle: combining complements can increase both innovation incentives and innovation capabilities.

7.4.3 Synergies

The Synergies principle recognizes that combining complementary assets can enhance innovation capabilities. As a classic example, in the pharmaceutical industry the process of bringing new drugs successfully to market requires an effective R&D program to identify and develop promising new compounds, the skills necessary to navigate the long and complex FDA testing and approval process, possibly demanding manufacturing capabilities, and effective marketing and distribution. Assembling these various skills, whether through contract, joint venture, strategic alliance, or integration, can lead to enhanced innovation capabilities.

7.5 Merger Enforcement

We are now ready to see what all of this implies for merger enforcement in cases where innovation effects are involved. This is no small matter, since merger enforcement is central to the work of the antitrust agencies and since many DOJ and FTC merger investigations and enforcement actions over the past fifteen years have involved innovation.[31] Here I follow in the footsteps of Katz and Shelanski (2005) and (2007), who offer an extensive

31. Katz and Shelanski (2005) and Gilbert (2006) note the growing importance of innovation in merger analysis. Katz and Shelanski (2005) also discuss a number of specific merger cases in which innovation has been an important factor. Gilbert and Tom (2001) discuss the rising importance of innovation in DOJ and FTC antitrust enforcement more generally during the 1995–2000 time period. Porter (2001) argues that antitrust treatment of mergers should focus on productivity growth. The 2010 Horizontal Merger Guidelines include, for the first time, a section devoted to innovation.

and thoughtful discussion of how merger enforcement does, and should, take account of innovation.[32]

Analysis of horizontal mergers involves predicting the effects of a specific, discrete change in industry structure, namely the joining of two former rivals under common ownership. As a practical matter, most mergers that receive serious antitrust scrutiny based on a theory of innovation effects involve two of a small number of companies with products, R&D programs, or capabilities in a given area. Usually, but not always, the two merging firms are important premerger rivals in the product market. The merger cases of greatest interest in which innovation effects are important typically fit into one of the following fact patterns:

- *Two product market rivals:* The merging firms are rivals in the relevant product market. One or both of them is investing in R&D to strengthen its position in the market.
- *Incumbent and potential entrant:* One merging firm has a strong position in the product market. The other merging firm has no current offering in the product market but is investing in R&D and will enter the product market if that R&D is successful.
- *Pure innovation rivals:* Neither merging firm has a current offering in the product market, but both are developing products to serve the market.

When examining a horizontal merger with possible innovation effects, we generally are interested in some version of this question:

> Will a merger between two rivals significantly reduce their *incentive* to innovate? If so, will the merger enhance their *ability* to innovate sufficiently to offset the reduced incentive?

The Contestability and Appropriability principles are directed at the first of these questions. The Synergy Principle applies to the second.

The overall relationship between market concentration and innovation is not especially relevant to this inquiry, especially since merger enforcement only takes place in moderately or highly concentrated markets. In particular, since merger analysis is not about a generalized increase in "competition," such as a reduction in the extent of product differentiation or an increase in imitation, the literature relating the (exogenous) degree of product differentiation to innovation is of little or no relevance to merger analysis. The Schumpeterian proposition that an ex post atomistic market structure is not conducive to innovation also is not directly relevant to merger enforcement, which involves a discrete change, usually a substantial increase in concentra-

32. Katz and Shelanski (2007) make the useful distinction between "innovation impact" and "innovation effects." My focus here is on "innovation effects."

tion, in ex ante market structure. The empirical literature on firm size and R&D is potentially more relevant, to the extent that it can inform us about the merger-specific efficiencies relating to innovation that are likely to arise when two competing business units are combined to form a larger business unit. However, the analysis of merger synergies is highly fact-specific. So far at least, general findings about firm size and innovation have not proven helpful for assessing merger-specific R&D efficiencies.

In subsection 7.5.1, I briefly explain what the recently revised Horizontal Merger Guidelines say about innovation effects. The guidelines utilize the Contestability, Appropriability, and Synergies principles. Subsections 7.5.2, 7.5.3, and 7.5.4 apply these principles to three merger cases in which innovation effects were central.

7.5.1 Innovation Effects Under the Merger Guidelines

The recently revised Horizontal Merger Guidelines contain Section 6.4, "Innovation and Product Variety." Innovation effects had not been explicitly addressed in the predecessor, 1992 Horizontal Merger Guidelines (see http://ftc.gov/os/2010/08/100819hmg.pdf). Section 6.4 begins this way:

> Competition often spurs firms to innovate. The Agencies may consider whether a merger is likely to diminish innovation competition by encouraging the merged firm to curtail its innovative efforts below the level that would prevail in the absence of the merger. That curtailment of innovation could take the form of reduced incentive to continue with an existing product-development effort or reduced incentive to initiate development of new products. The first of these effects is most likely to occur if at least one of the merging firms is engaging in efforts to introduce new products that would capture substantial revenues from the other merging firm.

This question is a direct application of the Contestability principle. Consider how the two firms are affected if Firm A introduces a new and improved product. The new product will increase Firm A's operating profits (measured gross of its R&D expenditures). If Firm B offers products that compete against Firm A's new product, the introduction of Firm A's new product will lower Firm B's operating profits. We can ask what fraction of Firm A's extra profits come at the expense of Firm B's profits. Farrell and Shapiro (2010) call this the "innovation diversion ratio."

How will this change if Firm A acquires Firm B? Applying the Contestability principle, the merger reduces the incentive to introduce this new product by more, the more profitable sales Firm A would capture from Firm B. Postmerger, sales gained at the expense of Firm B's products are no longer incremental to the merged firm: they cannibalize Firm B's profits. Put differently, the merger internalizes what had been a pecuniary negative externality. The merger turns the lost profits on Firm B's products into an opportunity cost borne by the merged firm when introducing Firm A's new

product. The magnitude of the resulting "tax" on the profits from Firm A's new product is, by definition, the innovation diversion ratio.

While the innovation diversion ratio is not typically amenable to precise measurement, because it involves products not yet introduced, the marketing and financial documents of merging firms, along with other evidence, can indicate the products from which a new product is expected to gain sales. Even when the innovation diversion ratio is not amenable to measurement, it is still conceptually central to evaluating the impact of the merger on Firm A's incentive to introduce its new product. When the innovation diversion ratio is high, the merger significantly reduces the contestability associated with the new product in question.

The guidelines reflect these ideas, along with the possibility of offsetting innovation synergies:

> The Agencies evaluate the extent to which successful innovation by one merging firm is likely to take sales from the other, and the extent to which post-merger incentives for future innovation will be lower than those that would prevail in the absence of the merger. The Agencies also consider whether the merger is likely to enable innovation that would not otherwise take place, by bringing together complementary capabilities that cannot be otherwise combined or for some other merger-specific reason. (Section 6.4)

As an example of merger-specific efficiencies relating to innovation, suppose that Firm A is considering investing in R&D to develop an improved process that will lower its unit costs. Suppose also that Firm A does not expect to expand its unit sales much as a result of these lower costs.[33] If the merger will enable the process innovation to be applied to Firm B's output, and if Firm A would not license its process innovation to Firm B in the absence of the merger, the merger can enhance Firm A's incentives to develop this process innovation. Of course, any such merger synergy must be weighed against the innovation diversion effects discussed earlier. In terms of the Contestability principle, the merger can increase innovation incentives by expanding the base of sales on which lower costs can be achieved. This effect is captured by a larger value of $D(P, \theta, z)$ in the $D(P, \theta, z)|C'(\theta)|$ term that is part of the innovation reward expression. This reflects the robust idea in the literature that smaller firms have lower incentives to engage in process innovations. However, offsetting this effect is the internalization of sales captured at the expense of Firm B's product, which reduces the $D_\theta(P, \theta, z)$ term in this same expression when viewed from the perspective of the merged firm.[34]

33. As discussed before, this can occur because the firm faces binding capacity constraints or because consumers have strong brand preferences and the firm will gain relatively few sales even if it lowers its price to fully pass through its lower costs.

34. For the merged firm, this term is given by the net gain in the combined sales of the two products, weighted by their margins.

Similar ideas can be used to evaluate the longer-term impact of a merger on innovation. The guidelines state:

> The second, longer-run effect is most likely to occur if at least one of the merging firms has capabilities that are likely to lead it to develop new products in the future that would capture substantial revenues from the other merging firm. The Agencies therefore also consider whether a merger will diminish innovation competition by combining two of a very small number of firms with the strongest capabilities to successfully innovate in a specific direction.

This line of inquiry also is directly related to the Contestability principle, but applies over a longer time frame, over which the firms' durable capabilities can be more informative than are their current offerings. These effects can arise even if the merging firms are not premerger product market rivals, as in the Genzyme/Novazyme and Thoratec/HeartWare cases discussed later.

Evaluating a firm's innovation capabilities is inherently difficult, and the importance of the R&D rivalry between the merging firms can be very difficult to assess if the attributes of the products likely to result from their R&D projects are unknown. Katz and Shelanski (2005) note that many of the merger cases in which R&D rivalry was central have involved pharmaceutical mergers. The FDA approval process often makes it possible to know well in advance which firms are in the best position to introduce drugs or medical devices soon in a specific therapeutic area.

Often, the firms with the greatest ability to innovate in a given area are those that have successfully innovated in similar areas in the past, or who own the complementary assets necessary to commercialize innovations. Such firms often have a strong ex ante market position. Historical R&D successes and current market position are thus two common indicators of a firm's innovation capabilities.

The guidelines incorporate the Appropriability and Synergies principles more explicitly in Section 10, "Efficiencies."

> When evaluating the effects of a merger on innovation, the Agencies consider the ability of the merged firm to conduct research or development more effectively. Such efficiencies may spur innovation but not affect short-term pricing. The Agencies also consider the ability of the merged firm to appropriate a greater fraction of the benefits resulting from its innovations.

The guidelines specifically ask whether the merger is likely to enable merger-specific efficiencies by combining complementary capabilities within a single firm. For example, a merger can enable cross-fertilization between the research teams of the two merging firms. Likewise, a merger can enable valuable information sharing between the regular operations of one merging

firm and the researchers at the other firm. Similarly, a merger can combine complementary assets such as a new product by a small start-up firm and the existing manufacturing or distribution assets of a larger, more established firm. However, merger synergies are far easier to claim than to achieve. The guidelines require that efficiencies be merger-specific and verified to be credited.

7.5.2 Genzyme/Novazyme

Genzyme Corporation acquired Novazyme Pharmaceuticals Inc. in September 2001.[35] Genzyme and Novazyme were the only companies pursuing enzyme replacement therapies for the treatment of Pompe disease, a rare and often fatal genetic disorder afflicting several thousand individuals in the United States, mostly infants and children. The FTC reviewed this merger after it was consummated but closed its investigation in January 2004, taking no action. The closing statement issued by FTC Chairman Timothy Muris stated: "The facts of this matter do not support a finding of any possible anticompetitive harm" (1). This was a striking assertion, since the merger created a monopoly in the market for Pompe enzyme replacement therapies.

The essential facts are as follows.[36] At the time of the merger, no treatments for Pompe disease had been approved by the Food and Drug Administration (FDA). Since Pompe disease is rare, under the Orphan Drug Act the first innovator to obtain FDA approval for a therapy is awarded seven years of exclusivity. By design, this regulatory structure rewards the first company to obtain FDA approval, even if patent protection is not available. However, this exclusivity may be lost if another innovator develops a superior treatment. This latter provision provides an incentive for other companies to continue their efforts to develop a superior treatment.

In the years leading up to the merger, Genzyme had invested heavily in developing a treatment for Pompe disease. At the time of the merger, Genzyme was pursing three treatments: one arising from a 1998 joint venture with Pharming, one arising from a 2000 joint venture with Synpac, and one that Genzyme had developed internally starting in 1999. The Synpac enzyme was in clinical trials and Genzyme was ramping up its own internal research program.

Novazyme had been developing its own Pompe treatment. At the time of the merger, the Novazyme treatment was not yet in clinical trials, but it had shown some promising results in mice. Novazyme was an especially aggressive innovation rival. The CEO of Novazyme, John Crowley, was

35. I rely primarily on the Federal Trade Commission (2004) for the facts of this case, but also on Anand (2006). The afterword in Anand (2006) provides an update as of October 2009.

36. I do not have access to the extensive confidential record that was available to the Federal Trade Commission. My comments here focus on the pertinent economic principles, not the FTC's enforcement decision in this case.

the father of two children with Pompe disease. His efforts to develop a treatment to save his children are documented in Anand (2006) and in the 2010 movie *Extraordinary Measures,* starring Harrison Ford. Prior to the merger, Novazyme projected that its treatment would reach clinical trials by the end of 2001. At the time of the acquisition, Genzyme announced that the Novazyme treatment would reach clinical trials in the first half of 2002.[37]

Soon after the merger, Genzyme reviewed all four treatments and decided to move forward to clinical trials with only the most promising one, which Genzyme determined to be its own internal program.[38] Anand notes, "Instead of being moved to human clinical trials, Novazyme's technology and experimental enzyme treatments were being sent back to the research labs."[39] Under Genzyme ownership, the Novazyme approach was slowed down, becoming a candidate for a superior, second-generation treatment.[40] Clinical trials for the Novazyme enzyme were substantially delayed. By the time of the FTC review in 2003, this date had been pushed back to between 2009 and 2011.[41] John Crowley left Genzyme in fall 2002. The internal Genzyme program commenced clinical trials in 2003, roughly one year after the merger with Novazyme.[42]

On its face, Genzyme's acquisition of Novazyme appears to have short-circuited a race between the two companies to be the first to obtain FDA approval of an enzyme treatment for Pompe disease. Applying the Contestability principle, all of the sales and profits accruing to the winner of this race were contestable prior to the merger. After the merger, however, far fewer sales and profits were contestable: Genzyme still had some incentive to gain FDA approval so it could begin earning profits from its treatment, but it no longer had to fear losing the race to Novazyme.

Furthermore, even if one assumes that there was no real race between the two companies, because Novazyme had no chance of gaining FDA approval before Genzyme, the merger still eliminated Novazyme as a competitor with a superior, second-generation treatment. Genzyme's incentive to develop a superior second-generation treatment would be far smaller than Novazyme's would have been, since sales of the second-generation treatment would come largely at the expense of the first-generation treatment. This is just the type of "replacement effect" identified by Arrow (1962) fifty years ago.

Application of the Contestability principle—following the approach de-

37. Federal Trade Commission (2004, 5), Dissenting Statement of Commissioner Thompson.

38. Anand (2006), chapter 23, "The Mother of All Experiments," describes Genzyme's evaluation of the four treatments.

39. Ibid., 261.

40. Ibid., 263.

41. Federal Trade Commission (2004, 5), Dissenting Statement of Commissioner Thompson.

42. The Genzyme treatment eventually gained FDA approval in spring 2006 under the brand name Myozyme. The treatment costs on average about $200,000 annually per patient. See Anand (2006, 316–17).

scribed in the 2010 Horizontal Merger Guidelines—strongly suggests that the merger had a significant adverse effect on innovation incentives. That conclusion appears to be further supported by the postmerger evidence available to the FTC at the time of its review. By 2003, it was clear that Genzyme's progress toward commercializing an enzyme treatment for Pompe disease had slowed down after the merger. As predicted by economic theory, Genzyme had delayed the development of the Novazyme treatment, pursing alternative treatments in *series* rather than in parallel.

How, then, did Chairman Muris conclude that the merger would not cause any anticompetitive harm? Muris begins by relying on "the lack of any clear theoretical or empirical link between increased concentration and reduced innovation" (2) to argue that there should be no presumption, even in a merger to monopoly such as this one, that innovation will be harmed. As discussed earlier, the overall cross-sectional relationship between market concentration and innovation is very difficult to discern for a number of reasons, including the lack of good data on concentration in relevant antitrust markets. Plus, even if one could measure this relationship, it is not directly relevant for analyzing mergers in which innovation effects are paramount, especially mergers to monopoly.

Chairman Muris does go on to examine the impact of the merger on Genzyme's incentive and ability to develop Pompe treatments. He denies that the two companies were racing for FDA approval, explaining:

> Shortly after the merger, Genzyme stated that comparative testing showed that its internal Pompe enzyme could be developed and commercialized most quickly. Genzyme also stated that the promise of the Novazyme technology was to provide a basis for an improved second-generation therapy. (Muris Statement, 12)

However, these statements, made by the merged firm itself in the face of antitrust review by the FTC, are perfectly consistent with the premerger Genzyme being spurred by the Novazyme threat to develop its treatment more quickly. In his dissenting statement, Commissioner Thompson, referring to competition between Genzyme and Novazyme states: "This competition was important because it created a race between Genzyme and Novazyme to develop Pompe ERTs, thus increasing the pace of innovation" (4). Given the inherent uncertainties associated with the new drug development process (and noting that Genzyme's treatment did not in fact gain FDA approval until 2006), it would seem hard to dismiss the possibility that, but for the merger, Genzyme would have been driven to move forward more quickly to gain FDA approval out of fear that Novazyme's treatment would gain FDA approval first.[43]

Even if one concludes that the merger did not reduce Genzyme's incen-

43. Anand (2006), chapter 20, "The Deal," offers evidence that Genzyme feared competition from Novazyme, and that these fears were a critical factor in Genzyme's decision to pay

tive to gain FDA approval for its first Pompe therapy, the merger reduced Genzyme's incentive to gain FDA approval for a *second* Pompe therapy. Chairman Muris explicitly notes this danger:

> If Genzyme has one Pompe therapy on the market, it might then have less incentive to market a second therapy than would an independent company that does not already have a product on the market. Because the second therapy would cannibalize sales of Genzyme's internal product, a merger with Novazyme could have caused Genzyme to reduce its investment in the second therapy. Moreover, Genzyme might have an incentive to delay introduction of the second therapy until the end of its initial seven years of market exclusivity in order to obtain a total of 14 years of exclusivity under the ODA. (Muris Statement, 13)

Muris dismisses this theory as well, noting that Genzyme would still have *some* incentive to develop and introduce a superior second treatment (14). However, by this argument one would never worry about the effect of a merger to monopoly on innovation because even a monopolist has *some* incentive to improve its product. Based on this dubious reasoning, Muris then states:

> In short, an analysis of Genzyme's incentives in this case does not clearly indicate whether Genzyme would have an incentive to delay the second Pompe product in the event that the first proved successful. (Muris Statement, 15)

Muris assigns a 75 percent probability to Genzyme's internal treatment gaining FDA approval, but concludes that the merger will not harm innovation to develop a superior treatment:

> There is no basis in the record for concluding that the circumstances that would give Genzyme an incentive to delay—concerns about cannibalization of sales of its internal product without sufficient offsetting expansion in demand, reduction in costs, or extension in product line—amount to anything more than a bare theoretical possibility. (Muris Statement, 19–20)

This statement appears to place no weight on Genzyme's reduced incentive to develop a superior treatment, and is peculiar given that Genzyme substantially delayed the Novazyme program during the time when the FTC was conducting its investigation.

Moving on to the Appropriability and Synergies principles, the merger does not appear to have solved an appropriability problem, or created merger-specific synergies, sufficient to offset the basic anticompetitive effects identified using the Contestability principle. Chairman Muris asserts: "By accelerating the Novazyme program, the merger may have increased its odds

---

$137.5 million (plus an additional $87.5 million on a contingent basis) for Novazyme, a company with no products or revenues.

of success" (17). However, as just described, the Novazyme program was greatly delayed after the merger.[44]

Muris also points to the comparative postmerger experiments conducted by Genzyme as a merger synergy. However, according to Anand (2006), Genzyme used these experiments to pick *one* Pompe treatment to pursue (its own internal program) and drop or delay the others. Without the merger, Genzyme could have performed comparative experiments among the three programs it controlled, and Novazyme could have continued with its own program, either alone or with another partner. That would have been a more innovative outcome.[45] Lastly, Anand (2006), reports that when Genzyme was bidding to acquire Novazyme in 2001, Genentech was offering to invest $22.5 million to acquire 10 percent of Novazyme and to fund the majority of the future development costs for the Novazyme Pompe treatment (224). Therefore, any benefits to Novazyme of gaining additional financing and moving forward with a major sponsor and partner were not specific to the Genzyme acquisition.

7.5.3 Thoratec/HeartWare

In February 2009, Thoratec and HeartWare signed an agreement under which Thoratec would acquire HeartWare for $282 million.[46] Thoratec was the only company offering a left ventricular assist device (LVAD) approved by the FDA for sale in the United States. According to the FTC Complaint, "LVADs are a life-sustaining technology for treating end-stage heart failure patients who have failed other courses of treatment and are likely to die while waiting for a donor heart or are ineligible for a heart transplant" (1). At the time of the proposed merger, HeartWare was developing its own LVAD, which was in the latter stages of clinical development.

The FTC challenged Thoratec's proposed acquisition of HeartWare in July 2009. According to the FTC Complaint, HeartWare was "the one company poised to seriously challenge Thoratec's monopoly of the US left ventricular assist device ('LVAD') market" (1). The FTC alleged that competition from HeartWare had already forced Thoratec to innovate and that the merger would eliminate innovation competition.

As with the Genzyme/Novazyme merger, we do not need to know about the overall cross-sectional relationship between market concentration and innovation to evaluate this merger. The Contestability principle tells us that

44. Muris later states: "it appears that the merger has accelerated the Novazyme program" (19). This assertion is difficult to reconcile with the description given in Anand (2006) and with Genzyme's public statements to investors about delays in the Novazyme program, as cited by Commissioner Thompson in his dissent (5).

45. According to Anand (2006), the Genzyme scientists had been skeptical of the Novazyme approach from the outset. This case thus also illustrates the advantages of independent ownership and competition for preserving innovation diversity when there are differences of opinion about which research tracks are the most promising.

46. I rely on the Federal Trade Commission (2009) for the facts reported here.

the merger would have substantially reduced rivalry, since many of the sales HeartWare stands to gain by obtaining FDA approval would come at the expense of Thoratec. There is no indication in the FTC complaint that the merger would have solved a substantial appropriability problem, or that the merger would have generated extraordinary merger-specific synergies.[47]

7.5.4 Ticketmaster/Live Nation

In February 2009, Ticketmaster and Live Nation announced their plans to merge.[48] For over two decades, Ticketmaster had been the dominant primary ticketing provider in the United States to major concert venues. The DOJ estimated Ticketmaster's share of primary ticketing to major concert venues at more than 80 percent, with the next closest competitor less than 4 percent.[49] Ticketmaster had also been slow to innovate and pass along lower costs to consumers:

> Ticketmaster's costs for distributing a ticket have been decreasing as consumers increasingly purchase tickets through the Internet. The cost-per-ticket to Ticketmaster for tickets sold through its website is significantly lower than the cost-per-ticket to Ticketmaster for tickets sold over the telephone or at a retail outlet. However, ticketing fees retained by Ticketmaster have not fallen as its distribution costs have declined. (DOJ Complaint, 11)

Live Nation was the largest concert promoter in the United States, also controlling over seventy-five concert venues in the United States. Live Nation had been Ticketmaster's largest primary ticketing client for a number of years. However, in 2007, Live Nation announced that it would not renew its contract with Ticketmaster and would instead become a direct competitor to Ticketmaster in primary ticketing once its Ticketmaster contract expired at the end of 2008.

In late December 2008, after nearly two years of preparation, Live Nation launched its ticketing service for its own venues and for potential third-party major concert venue clients. Live Nation represented an innovative threat to Ticketmaster's dominance in primary ticketing for major concert venues. By merging with Live Nation, Ticketmaster would have nipped that emerging threat in the bud. From the perspective of Live Nation, a large quantity of ticketing revenues were contestable, because Live Nation could capture those revenues from Ticketmaster. As initially proposed, the merger would have substantially reduced the contestability of ticketing revenues at major concert venues. However, Ticketmaster and Live Nation argued that the merger would also generate significant synergies through the vertical

47. "Any merger-specific and cognizable efficiencies resulting from the transaction will not offset the transaction's profound anticompetitive effects" (FTC Complaint, 2).

48. I rely on the Department of Justice (2010) for the facts reported here.

49. Department of Justice (2010, 10).

integration of promotion, venues, and ticketing. The Department of Justice eventually approved the merger subject to some substantial divestitures and other remedies.[50]

7.6 Exclusionary Conduct by Dominant Firms

Antitrust law in the United States has grappled for more than a century with where to draw the boundary between legitimate competition and exclusionary conduct by a dominant firm.[51] Considerable progress has been made on topics such as predatory pricing, but substantial controversy remains. Notably, the report on this topic issued by the Department of Justice in September 2008, "Competition and Monopoly: Single-Firm Conduct Under Section 2 of the Sherman Act," immediately drew sharp criticism from the Federal Trade Commission and was officially withdrawn in May 2009. My discussion here merely touches very lightly on the treatment of exclusionary conduct, focusing on innovation.

The highest-profile monopolization case in recent years, the case brought by the Department of Justice against Microsoft, centered on innovation effects. That case fit into the following general pattern: Firm M (the monopolist) is currently dominant in the market but faces the threat that Firm E (the entrant) will develop a new and improved product and overthrow Firm M as the market leader. Firm M engages in some type of conduct that impedes Firm E from developing new products, entering the market, or gaining scale. How does one determine whether Firm M's conduct is legitimate or exclusionary under the antitrust laws?

The empirical literature discussed earlier makes it clear that ongoing innovation by an incumbent is promoted if the incumbent fears that failure to improve its own product will place it at risk of being displaced as the market leader.[52] Likewise, innovation by entrants is promoted if an entrant that introduces a superior product will indeed gain substantial profitable sales, and perhaps even a dominant market position, at least for some period of time. Arrow was right that disruptive entrants with little or no financial interest in the status quo are critical to the innovative process. Schumpeter was also right that the prospect of gaining a temporary monopoly is a powerful inducement to innovate, for established firms and entrants alike.

The Contestability and Appropriability principles can go a long way—albeit at a high level—to inform the antitrust treatment of conduct by a

50. See Department of Justice (2010), Final Judgment, July 30, 2010. The author participated in this case at the DOJ.

51. For an entré to this literature that focuses on economic principles, see Kaplow and Shapiro (2007) and the references cited therein.

52. The cross-sectional relationship between market concentration and innovation is not directly relevant, especially inasmuch as the observations used to estimate that relationship involve concentration levels far lower than those associated with dominant firms facing a fringe of smaller rivals or entrants.

dominant incumbent firm in a market subject to technological change. Innovation by both incumbents and entrants is spurred if tomorrow's sales are contestable, in the sense that multiple firms are vying to win those sales and the lion's share of tomorrow's sales goes to the firm that succeeds in developing the best product. In the extreme case, one firm dominates the market at any point in time, but there is ongoing intense competition "for the market" that leads to rapid innovation. Innovation by both incumbents and entrants also is spurred if the successful innovator can appropriate a significant portion of the social benefits actually caused by its innovation.

Some have argued for a laissez faire antitrust policy in industries subject to technological change on the grounds that monopoly power in these industries is fleeting. However, this argument is seriously incomplete, since exclusionary practices (such as tying or exclusive dealing), if not checked by antitrust law, can make current monopoly power more durable by deterring innovative entrants. Others have argued for a laissez faire antitrust policy in industries subject to technological change on the grounds that such a policy would spur innovation by increasing the size of the prize won by the firm that obtains a dominant position. In a very important recent work, Segal and Whinston (2007) show that this argument also is seriously incomplete. In a model where two firms compete over time for market leadership by innovating, they provide surprisingly general conditions under which antitrust policies that protect entrants raise the rate of innovation.[53] Their analysis applies to a range of business practices by dominant firms, including long-term exclusive contracts with customers, compatibility decisions in a network industry, conduct that deters the R&D activities of entrants, and predatory activities.

7.7 Conclusions

Yes, Arrow *did* hit the bull's eye: a firm with a vested interest in the status quo has a smaller incentive than a new entrant to develop or introduce new technology that disrupts the status quo. Schumpeter was also quite correct: the prospect of obtaining market power is a necessary reward to innovation. There is no conflict whatsoever between these two fundamental insights.

The unifying principle, richly supported by the empirical literature, is that innovation, broadly defined, is spurred if the market is contestable; that is, if multiple firms are vying to win profitable future sales. This basic principle can take us a long way in evaluating the impact on innovation of horizontal mergers and of unilateral conduct by dominant firms.

53. Gans (2011) draws out some of the implications of the Segal and Whinston model for antitrust and innovation. In a related model of cumulative innovation, Raskovich and Miller (2010) provide conditions under which monopoly "extension" activities, which delay entry by the next incumbent, reduce the rate of innovation.

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics* 120(2): 701–28.

Aghion, Philippe, and Rachel Griffith. 2005. *Competition and Growth: Reconciling Theory and Evidence.* Cambridge, MA: MIT Press.

Aghion, Philippe, Christopher Harris, Peter Howitt, and John Vickers. 2001. "Competition, Imitation and Growth with Step-by-Step Innovation." *Review of Economic Studies* 68:467–92.

Aghion, Philippe, and Peter Howitt. 2009. *The Economics of Growth.* Cambridge, MA: MIT Press.

Anand, Geeta. 2006. *The Cure: How a Father Raised $100 Million and Bucked the Medical Establishment in a Quest to Save His Children.* New York: HarperCollins.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources to Invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors,* edited by the Universities-National Bureau Committee for Economic Research and the Committee on Economic Growth of the Social Science Research Councils, 609–26. Princeton, NJ: Princeton University Press.

Baker, Jonathan. 2007. "Beyond Schumpeter vs. Arrow: How Antitrust Fosters Innovation." *Antitrust Law Journal* 74:575–602.

Baker, Jonathan, and Carl Shapiro. 2008. "Reinvigorating Horizontal Merger Enforcement." In *How the Chicago School Overshot the Mark: The Effect of Conservative Economic Analysis on U.S. Antitrust,* edited by Robert Pitofsky. New York: Oxford University Press.

Bartelsman, Eric, and Mark Doms. 2000. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature* 38:569–94.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *Quarterly Journal of Economics* 122:1351–408.

———. 2010. "Why Do Management Practices Differ Across Firms and Countries?" *Journal of Economic Perspectives* 24:203–24.

Blundell, Richard, Rachel Griffith, and John Van Reenen. 1999. "Market Share, Market Value and Innovation in Panel of British Manufacturing Firms." *Review of Economic Studies* 66:529–54.

Boone, Jan. 2000. "Competitive Pressure: The Effects on Investments in Product and Process Innovation." *Rand Journal of Economics* 31:549–69.

———. "Intensity of Competition and the Incentive to Innovate." *International Journal of Industrial Organization* 19:705–26.

Christensen, Clayton. 1997. *The Innovator's Dilemma.* Boston: Harvard Business School Press.

Cohen, Wesley. 1995. "Empirical Studies of Innovative Activity." In *Handbook of the Economics of Innovation and Technological Change,* edited by Paul Stoneman, 1059–1107. Basil Blackwell.

———. 2010. "Fifty Years of Empirical Studies of Innovative Activity and Performance." In *Handbook of Economics of Innovation,* vol. 1, edited by Bronwyn Hall and Nathan Rosenberg, 129–213. Amsterdam: North Holland.

Cohen, Wesley, and Richard Levin. 1989. "Empirical Studies of Innovation and Market Structure." In *Handbook of Industrial Organization,* edited by Richard Schmalensee and Robert Willig. New Holland.

Davis, Ronald. 2003. "Innovation Markets and Merger Enforcement: Current Practice in Perspective." *Antitrust Law Journal* 71:695–703.

Demsetz, Harold. 1973. "Industry Structure, Market Rivalry, and Public Policy." *Journal of Law and Economics* 16:1–9.
Department of Justice, Antitrust Division. 2008. "Competition and Monopoly: Single-Firm Conduct Under Section 2 of the Sherman Act." September. http://www.justice.gov/atr/public/reports/236681.pdf.
———. 2010. "Ticketmaster Matter Materials." http://www.justice.gov/atr/cases/ticket.htm.
Farrell, Joseph, and Carl Shapiro. 2010. "Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition." *B. E. Journal of Theoretical Economics* 10, Article 9. http://www.bepress.com/bejte/vol10/iss1/art9.
Federal Trade Commission Staff Report. 1996. *Anticipating the 21st Century: Competition Policy in the High-Tech Global Marketplace.* http://www.ftc.gov/opp/global/report/gc_v1.pdf.
Federal Trade Commission. 2004. "Genzyme/Novazyme Matter Materials." http://www.ftc.gov/opa/2004/01/genzyme.shtm.
———. 2009. "Thoratec/Heartware Matter Materials." http://www.ftc.gov/opa/2009/07/thoratec.shtm.
Gans, Joshua. 2011. "When Is Static Analysis a Sufficient Proxy for Dynamic Considerations? Reconsidering Antitrust and Innovation." In *Innovation Policy and the Economy,* vol. 11, edited by Josh Lerner and Scott Stern, 55–78. Chicago: University of Chicago Press.
Gilbert, Richard. 2006. "Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate." In *Innovation Policy and the Economy,* vol. 6, edited by Adam Jaffe, Josh Lerner, and Scott Stern, 159–215. Chicago: University of Chicago Press.
Gilbert, Richard, and David Newbery. 1982. "Preemptive Patenting and the Persistence of Monopoly." *American Economic Review* 72:514–26.
Gilbert, Richard, and Willard Tom. 2001. "Is Innovation King at the Antitrust Agencies? The Intellectual Property Guidelines Five Years Later." *Antitrust Law Journal* 69:43–86.
Grove, Andrew. 1996. *Only the Paranoid Survive.* New York: Random House.
Holmes, Thomas, David Levine, and James Schmitz. 2008. "Monopoly and the Incentive to Innovate When Adoption Involves Switchover Disruptions." NBER Working Paper no. 13864. Cambridge, MA: National Bureau of Economic Research, March.
Holmes, Thomas, and James Schmitz. 2010. "Competition and Productivity: A Review of Evidence." *Annual Review of Economics* 2:619–42.
Kaplow, Louis, and Carl Shapiro. 2007. "Antitrust." In *Handbook of Law and Economics,* vol. 2, edited by A. Mitchell Polinsky and Steven Shavell, 1073–225. Elsevier.
Katz, Michael, and Howard Shelanski. 2005. "Mergers Policy and Innovation: Must Enforcement Change to Account for Technological Change?" In *Innovation Policy and the Economy,* vol. 5, edited by Adam Jaffe, Josh Lerner, and Scott Stern, 109–65. Chicago: University of Chicago Press.
———. 2007. "Mergers and Innovation." *Antitrust Law Journal* 74:1–85.
Lee, Chang-Yang. 2005. "A New Perspective on Industry R&D and Market Structure." *Journal of Industrial Economics* 53:101–22.
———. 2009. "Competition Favors the Prepared Firm: Firm's R&D Response to Competitive Market Pressure." *Research Policy* 38:861–70.
Leibenstein, Harvey. 1966. "Allocative Efficiency vs. 'X-Efficiency.'" *American Economic Review* 56:392–415.
Lewis, William. 2004. *The Power of Productivity: Wealth, Poverty, and the Threat to Global Stability.* Chicago: University of Chicago Press.

Motta, Massimo. 2004. *Competition Policy: Theory and Practice.* Cambridge: Cambridge University Press.
Nickell, Stephen. 1996. "Competition and Corporate Performance." *Journal of Political Economy* 104 (4): 724–46.
Porter, Michael. 1990. *The Competitive Advantage of Nations.* New York: McMillan Press.
———. 2001. "Competition and Antitrust: Towards a Productivity-Based Approach to Evaluating Mergers and Joint Ventures." *Antitrust Bulletin* 46:919–58.
Raskovich, Alexander, and Nathan Miller. 2010. "Cumulative Innovation and Competition Policy." Antitrust Division, US Department of Justice, EAG Discussion Paper 10-5, September. http://www.justice.gov/atr/public/eag/262643.pdf.
Sacco, Dario, and Armin Schmutzler. 2011. "Is There a U-Shaped Relation Between Competition and Investment?" *International Journal of Industrial Organization* 29 (1): 65–73.
Schmutzler, Armin. 2010. "The Relation Between Competition and Innovation: Why Is It Such a Mess?" Centre for Economic Policy Research, Discussion Paper no. DP-7640, January. University of Zurich.
Schumpeter, Joseph. 1942. *Capitalism, Socialism and Democracy.* New York: Harper & Brothers. (Citations to Harper Perennial Modern Thought Edition, published 2008.)
Scopelliti, Alessandro. 2010. "Competition and Economic Growth: A Critical Survey of the Theoretical Literature." *Journal of Applied Economic Sciences* 11:70–93.
Segal, Ilya, and Michael Whinston. 2007. "Antitrust in Innovative Industries." *American Economic Review* 97:1703–30.
Shapiro, Carl. 2007. "Patent Reform: Aligning Reward and Contribution." In *Innovation Policy and the Economy,* vol. 8, edited by Adam Jaffe, Josh Lerner, and Scott Stern, 111–56. Chicago: University of Chicago Press.
Sutton, John. 1998. *Technology and Market Structure.* Cambridge, MA: MIT Press.
———. 2007, "Market Structure: Theory and Evidence." In *Handbook of Industrial Organization,* vol. 3, edited by Mark Armstrong and Robert Porter, 2301–368. Elsevier.
Syverson, Chad. 2004. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy* 112:1181–222.
———. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49 (2): 326–65.
Tirole, Jean. 1997. *The Theory of Industrial Organization.* Cambridge, MA: MIT Press.
Vives, Xavier. 2008. "Innovation and Competitive Pressure." *Journal of Industrial Economics* 61:419–69.

Comment Michael D. Whinston

It is a pleasure to discuss a chapter of Carl's. The chapter focuses on an important but quite specific issue concerning innovation, namely the anti-

Michael D. Whinston is the Robert E. and Emily H. King Professor of Business Institutions at Northwestern University and a research associate of the National Bureau of Economic Research.

trust review of mergers in innovative industries. In the chapter, Carl makes two basic points: first, he argues that a merger's likely effects on innovation can often be discerned despite the seemingly negative lesson from the recent R&D and growth literatures, in which the level of innovation has no clear relation to the level of competition. Second, Carl suggests some principles that he feels can usefully guide such merger reviews. Here I will discuss these points in turn.

Let's start with the "complex relationship" between the level of competition and the rate of innovation, upon which the R&D and growth literatures have recently focused. What drives this complexity? In fact, you can see an important source of it by thinking about Arrow (1962) and Schumpeter (1942). Roughly speaking, there are two different times at which we might be concerned with market structure: ex ante (before the innovation) and ex post (after the innovation). Arrow showed that ex ante market structure is important, and that greater ex ante competition encourages innovation. The reason is simple: more ex ante competition destroys profits in the ex ante state, which gives firms a greater incentive to innovate to escape from that state. Schumpeter instead argued that competition is bad for innovation, but did so focusing on ex post market structure: destroying profits ex post reduces firms' incentives to innovate to get into that state. In essence, in the more recent models in this literature, competition is changed in *both* ex ante *and* ex post states. Because of this, things get complicated, and this tension between ex ante and ex post effects shows up in the varied effects observed in a lot of the literature.

Carl nicely illustrates this point in his discussion of the Aghion et al. (2005) paper. In that paper, the meaning of "less competition" is that there is less intense pricing rivalry when firms are in the neck-and-neck state in which they have the same technological capabilities. The neck-and-neck state is the ex post state when we look at R&D by the trailing firm when one firm is ahead and the other is behind,[1] but it is the ex ante state when we think about the R&D that occurs when the two firms are neck and neck. As a result, there are two opposing effects of more intense competition on innovation: an increase in innovation in the neck-and-neck state but a reduction in the state in which one firm is ahead. This fact then leads to an inverted U-shaped relationship between competition and innovation, where innovation is greatest at intermediate levels of competition. The reason for the inverted U is that the industry tends to spend more of its time in the state in which innovation is lowest, because that is the state firms tend not to move out of. Specifically, when there is little competition, there is little innovation in the neck-and-neck state, and a lot in the state where one firm is behind. As a result, firms are much more likely to be in the neck-and-neck state, which

1. Aghion at al. assume that a leader cannot be more than one step ahead; as a result, only the follower will do R&D in this state.

means that if we increase competition the (average) response of innovation is dominated by the response in the neck-and-neck state, which is positive. Similar reasoning implies that when competition is high in this sense, the industry is much more likely to be in the state where one firm is behind, so an increase in competition will reduce R&D on average.

While this inverted U-shaped relationship is certainly interesting and useful for understanding what we see in industry data, does it mean that we cannot predict the likely effects of a merger in an innovative industry? Carl argues no, and I agree. A key reason is that if you are thinking about mergers, the comparative statics exercise that is of interest to you—how this merger will affect the rate of innovation and welfare—differs from the comparative statics exercise that is conducted in this literature. To shamelessly plug some of my own work, a few years ago Ilya Segal and I wrote a paper (Segal and Whinston 2007) on antitrust in innovative industries. There we focused primarily on exclusionary behavior rather than on mergers, but a similar issue came up. We put the point as follows:

> The growth literature often considers how changes in various parameters will affect the rate of innovation, sometimes even calling such parameters measures of the degree of "antitrust policy". . . . Here we are much more explicit than is the growth literature about what antitrust policies toward specific practices do. This is not a minor difference, as our results differ substantially from those that might be inferred from the parameter changes considered in the growth literature. As one example, one would get exactly the wrong conclusion if one extrapolated results showing that more inelastic demand functions lead to more innovation (e.g., Aghion and Howitt 1992) to mean that allowing an incumbent to enhance its market power through long-term contracts leads to more innovation. (Segal and Whinston 2007, 1704)

Let's consider two examples to illustrate how the presence of a seemingly "complex relationship" between competition and R&D need not prevent definitive answers to specific competition policy questions. Consider first the model with Ilya. It was a quality ladder model of innovation similar to those in the growth literature. There was an entrant—if successful in its R&D, the entrant came in and competed for one period before displacing the incumbent monopolist. The entrant would then be an uncontested monopolist until he himself ultimately faced a successful new entrant and was displaced.

In this setting we asked whether allowing incumbents to deter entry through exclusive contracts with buyers would encourage or discourage innovation. (The question was motivated in part by the *Microsoft* case, where Microsoft wrote partially exclusive contracts with buyers and providers of complementary goods.) Exclusive contracts reduce the number of buyers who are free to purchase from an entrant, which tends to reduce innovative effort by prospective entrants. However, once an entrant displaces

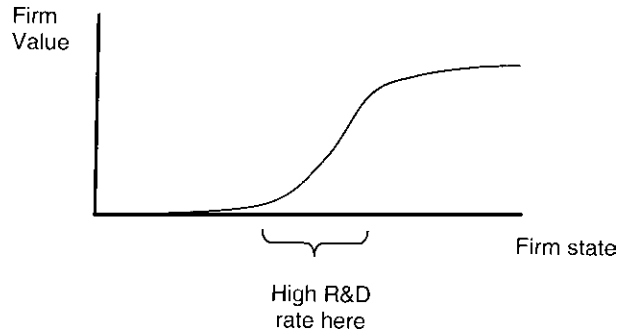

Figure 7C.1 A firm's value function in the Pakes-McGuire model

the incumbent and becomes the new monopolist, it is more profitable if it can deter entry, so allowing such deterrence could also raise the incentive to innovate. As a result, it might seem like one cannot say anything about which way the overall effect comes out. Nonetheless, we showed that fairly generally the use of exclusives lowers the rate of innovation (and both consumer and aggregate surplus).

Now consider a different dynamic model of innovation due to Pakes and McGuire (1994) (see also Ericson and Pakes 1995). In this model, there is a differentiated product oligopolistic industry in which, in each period, firms engage in price competition and can also invest in stochastic product improvement. Both entry and exit are also possible. A firm's value function in this model typically looks as in figure 7C.1, where the horizontal axis measures the firm's state (innovation can increase a firm's state, which raises its product's value to consumers) and the vertical axis measures the firm's value. The graph of the value function in the figure holds the states of the firm's rivals fixed.

As can be seen in the figure, the value function is S-shaped: relatively flat at low and high states, with a steep section in the middle. Innovation will be high when the firm is in a state at which this curve is steep (the returns to product improvement are then large). The steep section is like the neck-and-neck state in Aghion et al. (2005). Although Pakes and McGuire do not do this, I think if you actually looked at this model and had a bunch of these industries in different states, you likely would get an inverted U-shaped relationship between the rate of innovation and the intensity of competition. At the very least, the relationship would be "complex."

Nonetheless, when Pakes and McGuire simulate the effect of a merger in the Markov perfect equilibrium of their model, its impact on consumers is very clear. Table 7C.1 shows the levels of industry profit, consumer surplus, and aggregate surplus in three cases: the first best, the oligopolistic Markov perfect equilibrium, and a fully collusive outcome. The fully col-

Table 7C.1 **Profit, consumer surplus, and aggregate surplus in the Pakes-McGuire model**

	Industry profit	Consumer surplus	Aggregate surplus
First best			377
Markov perfect equilibrium	70	301	369
Collusion (industry-wide merger)	218	115	332

lusive outcome can be thought of as the result of an industry-wide merger (including all potential entrants). The first-best aggregate surplus is 377. There is a small loss in aggregate surplus in the Markov perfect equilibrium: consumer surplus is 300 and industry profit is 70. (This is an industry where, on average, three or four firms are active.) With the industry-wide merger, aggregate surplus falls 10 percent compared to the Markov perfect equilibrium and consumers do really badly: their surplus falls by almost two-thirds. (The rate of innovation also falls dramatically.) Thus, despite any general complexity of the relation between the level of competition and the rate of innovation, this merger is evidently very bad for consumers. Gowrisankaran (1995) also finds negative effects on consumers (and a reduction in R&D) in a closely-related model when he allows for (endogenous) nonindustry-wide mergers.[2]

In summary, I think Carl is completely correct in his first point: while the R&D and growth literatures that exhibit "complex" (inverted U-shaped) effects are certainly interesting and valuable contributions, they are often not on point, or only partially so, for the questions we want to ask when evaluating mergers in innovative industries.

Now to Carl's second point. Suppose a merger in an innovative industry faces antitrust review. What can we say about the merger's likely effects on innovation? Carl proposes some principles to aide such analysis. Perhaps it would be most useful if I discuss how I would think about the likely effects on innovation if I were looking at such a merger.[3] (One would also need to think about its overall effect on consumers.)

My starting point would be to assess how the merger changes the R&D incentives for the merging firms, holding fixed the R&D activities of the merging firms' rivals. Here one is assessing how the merger changes the degree to which the firms' profits respond positively to their level of inno-

2. It is worth noting that other interventions to increase "competition" need not be welfare-improving. For instance, Pakes and McGuire also simulate the effect of a rule limiting firms' market shares to be no greater than 65 percent. This rule reduces both consumer and aggregate surplus relative to the Markov perfect equilibrium.

3. Because Carl changed his statement of these principles in the revised draft of his paper, I have modified what follows somewhat from my discussion at the conference. The discussion that follows is, I think, broadly consistent with the approach Carl proposes in the final version of his chapter.

vation. Several factors go into this. The most important seems to me to be the degree to which the merger internalizes externalities arising from the merging firms' R&D. This R&D externality internalization effect of the merger could in principle be positive or negative. For example, in a quality ladder model there is an important positive externality across generations (each innovation enables later ones), so a merger could increase innovation incentives by internalizing this positive externality. On the other hand, in the Pakes and McGuire model, innovation creates only negative externalities across firms, so a merger will most likely reduce innovation incentives. But what is important to note, I think, is that this first critical factor is likely to be reasonably assessed by those reviewing the merger, and is unrelated to the factors contributing to the "complex" relationship just discussed. This is where the fact that we are focusing on the effect of a *merger,* not some other change in "competition," really matters.

Mergers also cause externalities on another set of market participants: consumers. Because the merger internalizes pricing externalities, it can alter the degree to which firms rather than consumers benefit from an innovation, and hence can alter firms' incentives to do R&D. This effect is related to the complex relationship discussed earlier, and is probably harder to assess. My own gut feeling is that in most (though not all) cases, this effect is likely to be less important than the R&D externality internalization effect.

Finally, this first step also needs to incorporate any efficiency effects in R&D production created by the merger.

A second concern is how the merging firms' rivals will react to this change. In particular, are R&D efforts strategic substitutes or strategic complements in the sense of Bulow, Geanakoplos, and Klemperer (1985)? If they are strategic complements and you dull innovation incentives for the merging firms, everyone's R&D goes down. If they are strategic substitutes, then the rivals will increase their R&D in response to the merging firms reducing theirs. In that case, it may seem that the overall effect is unclear. Typically, however, we expect that this countervailing effect does not overwhelm the direct effect—that the other firms do not expand their R&D enough to counterbalance the R&D contraction of the merging firms. Indeed, in most theoretical papers, this is just invoked as a standard assumption. Its import is that, if true, one only needs to look at the direct effect on the merging firms' R&D holding rivals' R&D efforts fixed to discern the overall effect on R&D.

Matters would be more complicated when innovative efforts are not one-dimensional. For example, a merger might enhance incentives for some types of R&D and reduce it for others. Or the R&D of the rivals may differ from that of the merging firms. Nonetheless, in many cases this way of thinking seems likely to get us fairly far in thinking about these issues.

To sum up, this is a worthwhile chapter that should help restore faith among those who need to evaluate mergers in innovative industries, and that also provides some guidance on how to do it.

References

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt. 2005. "Competition and Innovation: An Inverted U Relationship." *Quarterly Journal of Economics* 120:701–28.

Arrow, K. 1962. "Economic Welfare and the Allocation of Resources to Invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors,* edited by Universities-National Bureau Committee for Economic Research and the Committee on Economic Growth of the Social Science Research Councils, 467–92. Princeton, NJ: Princeton University Press.

Bulow, J., J. Geanakoplos, and P. Klemperer. 1985. "Multimarket Oligopoly: Strategic Substitutes and Complements." *Journal of Political Economy* 93:488–511.

Ericson, R., and A. Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Analysis." *Review of Economic Studies* 62:52–82.

Gowrisankaran, G. 1995. "A Dynamic Model of Endogenous Horizontal Mergers." *RAND Journal of Economics* 30:56–83.

Pakes, A., and P. McGuire. 1994. "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model." *RAND Journal of Economics* 25:555–89.

Schumpeter, Joseph. 1942. *Capitalism, Socialism and Democracy.* New York: Harper & Brothers.

Segal, I., and M. D. Whinston. 2007. "Antitrust in Innovative Industries." *American Economic Review* 97:1703–30.