This is a section of doi:10.7551/mitpress/12686.001.0001

Innovation Matters

Competition Policy for the High-Technology Economy

By: Richard J. Gilbert

Citation:

Innovation Matters: Competition Policy for the High-Technology Economy
By: Richard J. Gilbert
DOI: 10.7551/mitpress/12686.001.0001
ISBN (electronic): 9780262358637
Publisher: The MIT Press
Published: 2022

OA Funding Provided By:

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

The MIT Press

4 Dynamics, Cumulative Innovation, and Organizational Theories

> What's the point of focusing on making the product even better when the only company you can take business from is yourself?
> —Steve Jobs, "Voices of Innovation" (2004)

1 Introduction

Kenneth Arrow and Joseph Schumpeter are towering intellectuals associated with the theory of competition and innovation. Their contributions are profoundly important, but neither Arrow nor Schumpeter developed theories that describe innovation incentives that apply more generally to dynamic markets. Innovation is quintessentially a dynamic process. Furthermore, it is typically cumulative, with discoveries providing a knowledge base that enables future discoveries. Yet most theories, including the Arrow replacement effect, assume that innovation is a single discrete event with only limited consequences for the future evolution of industry technology and market structure.

Section 2 of this chapter reviews models of innovation incentives that embed both the Arrow replacement effect and Schumpeterian effects in dynamic settings. These models get complicated quickly and require simplifying assumptions to be analytically tractable. The section begins with a discussion of predictions from a model that has only two periods. Firms invest in research and development (R&D) in the first period to discover new products or new production technologies, and they choose prices in the second period for the products that emerge from the first-period R&D or compete using new production technologies.

A more dynamic formulation discussed in this section is a patent race in which investment in R&D yields a probability of discovery, and the first firm to make that discovery wins a patent. Both two-period models and patent races show that an increase in rivalry often, but not always, increases the probability of discovery. The two-period model shows that rivalry has an additional benefit: It increases expected future price competition when two or more firms develop innovative products that address similar market demand. The benefit of price competition is absent in models of patent races that assume that the patent excludes competition.

Industrial structure shapes the incentives for firms to invest in R&D and is shaped by the innovations that result from those investments. That interaction is absent in two-period models of R&D investment and in "winner-take-all" patent races. Section 2 also surveys dynamic models that allow this interdependence and explores their implications for the relationship between market structure, competition, and innovation. Models in which firms compete to improve technologies in discrete steps show that modest increases in competition can increase the probability of innovation, but intense competition can make innovation less likely. Competition increases the incentive to innovate to escape price constraints imposed by rivals, while also lowering the return to a firm that becomes a successful innovator by reducing the ability of the firm to profit from the innovation. Furthermore, intense competition can lower the probability that an industry is in a structural state that promotes R&D rivalry.

Many firms invest in R&D with the expectation that if their efforts show promise, they will be acquired by an incumbent or else license relevant technology to an established firm. When there is R&D for buyout, innovation incentives depend both on rivalry in R&D and on competition among incumbents that are likely bidders for the innovator or for licenses to its technology. Section 2 describes the implications of this bidding competition for R&D incentives.

Section 3 addresses economic incentives for firms to invest in innovations that build on prior innovations, also called "cumulative innovation." Cumulative innovation is a particularly important issue for the scope of intellectual property rights because strong patent rights provide incentives for early innovators, but potentially impose costs on firms that make infringing follow-up innovations. Xerography was a fundamental discovery that warranted patent protection. But plain-paper copiers were not widely accepted in the workplace until innovations in paper handling made these copiers an efficient office productivity tool. This section does not address the design of intellectual property rights. The focus instead is on antitrust policies that affect the exercise of these rights or other mechanisms that protect early innovators but have consequences for follow-on innovators.

Innovation is a human activity that often occurs in a complicated organization, which has limited information flows among the company's owners, managers, researchers, and customers, and may require structural adaptations to the existing business bureaucracy. Several theories of organizational behavior and bureaucracy address these limitations and their implications for innovation incentives. Section 4 reviews some of the more prominent theories and relates them to economic models.

2 Innovation Dynamics

We begin here with an extension of the discussion of incentives for innovation by exploring the implications from a model in which there are two separate periods. Firms invest in R&D during the first period, and successful innovators compete in price during the second period.

Two-period models of R&D and price competition

This subsection summarizes key predictions from a two-period model presented in Gilbert (2019), which is closely related to models developed by Federico, Langus, and Valletti (2017, 2018) and Letina (2016). In the first period, firms can invest in any number of discrete R&D projects, each of which incurs a fixed cost and has an independent probability of success. For example, if each R&D project has a 50 percent probability of success, then if a firm invests in two projects, the probability that at least one succeeds is 75 percent. Innovation is winner-take-all if only one firm can profit from a discovery. Otherwise, firms can share profits from their discoveries, with competition reducing both the total available profit and each innovator's share of the total profit.

This relatively simple model has some useful implications. Suppose that firms have no profits from discovery that are put at risk and they cannot imitate discoveries made by others. If innovation is winner-take-all, then R&D incentives are independent of the structure of the market.[1] The reason for this result is that an additional R&D project has value only if all other R&D projects in the industry fail, without regard to whether these other projects are held by a single firm or distributed among many firms.[2]

A second implication from the model is that an increase in rivalry typically increases innovation incentives if innovation is drastic (as defined in chapter 3) but not "winner-take-all." If discovery is not winner-take-all, firms have incentives to innovate to take business from rival firms that have also innovated successfully. The opportunity to profit from simultaneous innovation provides a stimulus for R&D investment that increases with the number of potential innovators. Consequently, an increase in rivalry typically increases the total industry profit-maximizing level of R&D effort.[3] The opposite can occur if rivalry is intense. An additional rival lowers each firm's expected profit from R&D, which can more than offset its contribution to industry R&D efforts in some circumstances.

A third implication is that rivalry can reduce innovation incentives if innovations have only incremental benefits (i.e., if innovation is not drastic). Competition from existing products reduces the payoff from new products when innovations are not drastic. Consequently, industry investment in R&D for incremental innovations can peak at an intermediate level of market concentration.

A fourth implication is the importance of rivalry for future price competition. Rivalry can benefit consumers by increasing future price competition even if it lowers or has no effect on innovation incentives.

Fifth, the simple model also demonstrates that profits at risk from innovation reduce a company's incentives to invest in R&D (the Arrow replacement effect). An increase in the number of firms that invest in R&D can mitigate the replacement effect if competition among these firms lowers the profits that are at risk from innovation. A merger can increase the replacement effect and reduce R&D incentives if both firms have profits that are at risk from innovation, even if the firms' existing products do not compete with each other (see chapter 5).

Finally, information spillovers that allow rivals to imitate a discovery more easily can reduce incentives for firms to invest in R&D. It does not follow that such spillovers harm industry innovation. In the presence of information spillovers, an increase in the number of firms that can benefit from a discovery can lower each firm's expected profit from discovery and reduce incentives to invest in R&D; however, there is a compensating benefit because firms that benefit from information spillovers are more likely to innovate.

At a more general level, R&D competition promotes innovation by increasing the number of firms engaged in R&D, but R&D competition can harm innovation by reducing the ability of firms to appropriate value from their discoveries. Appropriation has several dimensions. An increase in competition lowers each firm's ability to profit from a discovery in the presence of other innovators. Furthermore, for innovations that are not drastic, competition from existing products can lower the profits from a discovery. The effect of competition on the value of an innovation is often called a Schumpeterian appropriation effect. An increase in competition also can reduce the ability of innovators to appropriate value from a discovery by increasing the number of potential rivals that can compete with the innovator by imitating the discovery.

Mergers can increase the ability of firms to capture value from discoveries by reducing the number of rivals that can profit from imitating a discovery. Furthermore, technological spillovers that occur within the walls of the merged firm can increase the value of an innovation or increase the efficiency with which the firm can conduct R&D. Chapter 5 describes these merger-specific issues in more detail.

Patent races

One of the earliest approaches to adding dynamics to study the interaction between competition and innovation is the patent race. Firms compete by choosing R&D expenditures that generate a probability distribution for the discovery of a patentable invention. The first firm to realize the discovery wins the patent.

Simple models of patent races assume that patent protection excludes competition, so the winner of the race takes all or most of the available profit from an innovation. Patent protection is not essential for this property. Innovators in some industries can achieve temporary monopolies, or at least a very large increase in market share, in the absence of significant patent protection. The patent race model can be a reasonable description of R&D competition in markets without strong patent protection if firms have other ways to appropriate value, such as lead time, secrecy, or complementary factors of production.[4]

Generally, these models show that competition increases R&D effort and the probability of discovery, and reduces the expected time to patent an invention.[5] They also show that R&D competition can be socially excessive by incurring costs that exceed the benefits from the R&D, because firms ignore the negative effect of their R&D investments on the probability of discovery by others. This latter result has little policy significance for antitrust enforcement. The overwhelming empirical evidence is that R&D has large spillover benefits,[6] which suggests that, on average, firms invest too little rather than too much, in R&D.[7]

These early models are insightful and analytically convenient, but they offer only limited insight for strategic behavior that can have important consequences for competition and innovation. In particular, because these models assume that future success is independent of R&D that occurred in the past, they have no scope for strategic conduct that depends on a firm's technological lead compared to its rivals in the R&D competition.

For example, a firm that is first to make an intermediate discovery may gain an advantage that causes rivals to reduce their R&D efforts.[8] Under some conditions, the advantage can be so large that lagging firms cannot profitably leapfrog or even catch up to a firm that has made the intermediate discovery. In that case, lagging firms may drop out of the R&D competition entirely, leaving the technological leader as the only surviving investor. For firms that lag a technological leader, spending money on R&D can amount to pouring good money after bad if the leader is sure to maintain its technological edge. With its rivals out of the race, the leader can be content to invest in R&D at a profitable rate that can be much less than the industry would sustain if rival firms achieved technological parity.[9]

In other words, the effect of rivalry on R&D competition depends on the relative technological capabilities of the rivals, not just their number, which is a result that is confirmed in the empirical literature (see chapter 6).[10] R&D competition can be more intense in an industry with two firms that have similar technologies than in an industry with four firms if one firm has a substantial technological lead over its rivals. The next subsection describes a class of models that focus on innovation competition between firms that can occupy different positions in a technology space.

Stepwise innovation

The stepwise innovation model emerged from collaborative efforts by Philippe Aghion, Christopher Harris, Peter Howitt, and John Vickers.[11] In their model, firms differ in the quality or cost of their product. Innovation is a discrete improvement in quality or cost, similar to climbing steps on a ladder. Firms can move higher up the ladder by investing to improve their qualities or lower their production costs, but first they must catch up to the current industry leader before they can become the next leader. These assumptions rule out innovations that allow a firm to leapfrog the current leader.

The stepwise innovation model illustrates incentives to invest to improve a product that depend on the product's position on the technological ladder relative to rivals' products. When competitors have similar qualities, there is an incentive to create a better product to escape competition and enjoy a temporary claim to higher profits. The greater the competition among firms that have technological parity, the greater is the incentive to innovate to escape this competition. This is the opposite side of the replacement effect in Arrow's static model. Greater competition lowers the profits from existing products. This implies a smaller replacement effect, and a corresponding greater incentive to create new products.

In Arrow's model, a successful innovator wins a patent that excludes competition for the life of the patent. Instead, the stepwise model plausibly assumes that a firm that catches up to a market leader has to share the profits from innovation with its equally efficient rivals, and may have its profits squeezed by competition from products that are lower down on the innovation ladder. Consequently, competition has a countervailing effect for innovation in the stepwise model. The greater the extent of competition in the market, the less a lagging firm can earn if it catches up to a market leader. Aghion and his colleagues characterize this as the *Schumpeterian effect*, in which competitive pressures dull innovation incentives for technological laggards.

Competition has countervailing effects for innovation in the stepwise model. Competition spurs firms to innovate, but competition also lowers the profit earned by a firm that innovates and catches up to the market leader. Some competition is desirable to stimulate innovation in the stepwise model. If products do not compete, there is little benefit from moving up the innovation ladder. The incentive to innovate also can be low if competition is intense because competition reduces the payoff from catching up to the market leader, although this has to be balanced against the stimulus from escaping competition.

The stepwise innovation model highlights the interdependence between industry structure, competition, and innovation incentives. An industry may appear to have little competition because firms have highly differentiated products, but that differentiation could be the result of innovation driven by intense competition. The effect of competition on the incentive to innovate differs for firms that are at the innovation frontier compared to those who are behind the frontier, and the overall effect of an increase in competition depends on the composition of firms in an industry, as measured by their technological prowess, which in turn is driven by innovation incentives.

The stepwise model demonstrates that intermediate levels of competition can best promote competition. This is the inverted-U relationship between competition and innovation. The stepwise model also demonstrates the importance of technological dissimilarities for innovation competition. Empirical researchers have attempted to verify the inverted-U relationship, but with limited success. Empirical support is more robust for the effect of technological capabilities on innovation. Several empirical studies show that competition has more pronounced benefits for innovation by firms that are technological leaders in an industry.

A limitation of the original stepwise model is its restriction to duopolies, with competition measured by the intensity of competition within each duopoly.[12] Antitrust authorities have limited scope to influence the degree of competition in an industry. They can prohibit certain exclusionary conduct and impose conditions such as compulsory licensing, supply obligations, or nondiscrimination obligations, but they cannot force firms to compete more or less.

Antitrust authorities can influence the number of firms that compete in an industry by choosing to block or allow a proposed merger. With the important caveat that a merger is not the same as an industry with one less rival, Gilbert, Riis, and Riis (2018) explore how results from the stepwise innovation model change when competition is measured by the number of industry rivals rather than the intensity of competition in a duopoly. In some circumstances, the rate of industry innovation reaches a peak level with a moderate number of rivals, while in other circumstances, greater rivalry increases the rate of industry innovation. These results depend on firm behavior, consumer demand, and the firms' technologies. The authors emphasize the importance of these factors for qualitative inferences about the effects of competition on industry innovation.

Bidding for R&D assets

Many entrepreneurs invest in R&D with the intent to sell or license their results to an established company. For some, the strategy is a logical division of effort. Research organizations in the pharmaceutical and biotechnology industries specialize in the discovery of promising molecules or methodologies, and sell or license results to pharmaceutical companies that have the capabilities to administer and evaluate clinical trials, obtain regulatory approvals, and promote successful drugs. In the biotechnology sector, there were 68 acquisitions in 2014 alone, with a total reported value of $49 billion. In that year, biotech companies also entered into 152 licensing deals valued at an additional $47 billion.[13] For others, "innovate to sell" is the only practical strategy because established firms are protected by high barriers to competition and would crush upstarts if they try to compete. Facebook, Alphabet (Google), Amazon, and other tech goliaths have copied or threatened to copy promising new products.[14]

Competition in the acquiring market affects the incentive to innovate, in two ways. First, competition from existing products affects the maximum amount that a buyer would be willing to pay for an exclusive right to a new product. (I use the term "product" generically to describe any innovation, including R&D projects or new production technologies.) This maximum amount is the difference between the buyer's profit as the only supplier of the new product and the profit the firm can earn if it remains with an old product. I call the effect of competition on this profit difference the *value effect* of competition in the acquiring market.

A second effect arises from competition among potential buyers to acquire the innovation. Buyer competition for an exclusive right to a new product resembles an auction. Conditioned on the value of the new product, the winning bid is likely to be higher if more firms are interested in acquiring it. I call the effect of competition among buyers on the acquisition price the *auction effect*. For example, suppose that potential buyers of a new product are uncertain whether it is worth 80, 100, or something in between. With only two bidders, the acquisition price may be close to 80, while adding more bidders would increase the likelihood that the winning bid is closer to 100.

The auction effect is generally positive. The value effect is absent for a drastic innovation as defined in chapter 3 because competition from existing products does not affect the profit that a firm can earn from a drastic innovation. For a nondrastic innovation, the value effect is also positive if competition in the old product increases the profit that a firm can earn with the new product relative to what it would earn if the firm remained with the old product. In that case, both the value and auction effects imply that competition in the acquiring market increases the likely price paid to an innovator. Conversely, a merger of firms in the acquiring market lowers the likely acquisition price for an innovation and therefore lowers R&D incentives when the value effect is positive.

However, the value effect can be negative in some circumstances: An increase in competition in the acquiring market can lower the profit from a new product or technology relative to what a firm could earn if it remained with an existing product or technology.[15] In that case, the value and auction effects have opposite signs, and it is possible that the maximum return to an innovator would occur at an intermediate level of competition in the acquiring market. An implication is that a merger of firms in the acquiring market could increase the likely prices paid to acquire innovations and consequently increase R&D incentives when the value effect is negative.

Guillermo Marshall and Álvaro Parra describe innovation competition in a more detailed analysis that captures the auction and value effects for innovation incentives.[16] In their model, research laboratories engage in an innovate-to-sell strategy, while established firms compete with each other in the product market and with research labs in R&D. Consistent with the discussion in this section, they find that a decrease in the number of firms in the product market can lower the innovation rate if the value effect is negative, although that depends on an offsetting effect from the lessening of competition in R&D by firms in the product market that also compete in R&D.

3 Cumulative Innovation

The accumulation of knowledge enables innovation in the present and facilitates future innovations. In the words of Sir Isaac Newton, "If I have seen farther, it is by standing on the shoulders of giants."[17] Discoveries rarely spring de novo from the minds of inventors. Instead, they build on previous discoveries and add to the store of knowledge. The transistor and integrated circuit are typical examples of the evolution of cumulative innovation in the high-technology economy. The transistor implemented vacuum tube technology in a solid-state device. The integrated circuits that power numerous applications in the digital economy rely on the transistor and would not have been feasible without developments in photolithography to print circuit designs at microscopic dimensions.

Unfortunately, many important and influential theories of competition and innovation fail to address the trade-off between first-generation and subsequent innovation incentives. Arrow's replacement model treats innovation as a one-time event, as do simple models of patent races. In Schumpeter's theory of creative destruction, new innovations are waves that obliterate rather than build on prior discoveries. Rewards encourage innovation, but market rewards for discoveries can occur at the expense of rewards for innovations that build on those discoveries. Competition policies can balance these rewards, in part by preventing innovators from impeding subsequent innovators.

The history of many technological breakthroughs has dark chapters in which early inventors erected barriers to future innovation.[18] Wilbur and Orville Wright deserve accolades for their pioneering achievements in human flight. Among their accomplishments was the use of "wing warping" to steer an airplane, much as a bird bends its wings to change direction, for which they were awarded US Patent No. 821,393 in 1906. A few years later, Glenn Curtiss developed the use of ailerons to control flight, which is essentially the same technology used in modern aircraft. But the Wright brothers claimed that their patent also covered ailerons, and after years of litigation, they prevailed in court. Curtiss also won patents and resisted licensing new aviation entrants. The aircraft industry did not open itself to competition until 1917, when the US government compelled the industry to form a patent pool and license its patents. Also in response to concerns about the foreclosure of new technologies, in 1919, the government urged General Electric (GE), Marconi, AT&T, Telefunken, and Westinghouse to pool and cross-license their patents on radio technologies.[19]

Much of the debate about how to reward early and follow-on innovators emphasizes the design of intellectual property rights, which includes the scope of rights, their duration, and the novelty threshold for patentability.[20] Edmund Kitch argued that early inventors should have broad property rights that cover follow-on innovations, similar to the way that mineral rights for public land provide claimants with the exclusive right to prospect for additional discoveries within a defined region.[21] Broad and lengthy patents give more protection to initial inventors but, as a consequence of the Arrow replacement effect, they provide little incentive for follow-on innovations by the patentee that would replace its existing profits. A patentee can delegate subsequent innovation to others, but a broad patent scope allows the early innovator to require compensation from follow-on innovators because their innovations would infringe the original invention, absent a license. The compensation is a cost that reduces the profit from follow-on innovation.

Suzanne Scotchmer observed that, in general, it is not possible to design a system that provides efficient R&D incentives for both early and follow-on inventors without resorting to external subsidies, if they do not participate in an integrated R&D venture.[22] Greater rewards to one party necessarily occur at the expense of fewer rewards to the other party. Bargaining over licensing fees is problematic because licensing negotiations typically occur after firms have incurred R&D costs, which allows the early innovator to appropriate the benefits from these investments without adequately compensating follow-on innovators for their R&D expenses. Allowing innovators to bargain before R&D costs are sunk can solve the incentive problem, but that is a fiction if follow-on innovators are not known until they realize the results of their R&D investments.

Competition and intellectual property policies create trade-offs for incentives to invest in cumulative innovation. Potential inventors will not have financial incentives to incur costs to make a new discovery if they cannot anticipate receiving a healthy reward. However, rewards that take the form of exclusive rights can discourage future innovation by giving the first innovator the right to block future innovators or to impose a tax in the form of demands for large royalties in exchange for the right to use the first innovator's technology.

Figure 4.1 illustrates this trade-off for a frontier or first-generation invention that is also an input for follow-on innovation. The graph depicts a situation in which the first discovery can earn a per-unit royalty, which compensates the first inventor but is a cost for follow-on innovation. An example is the CRISPR technology for genome editing, which can be used to bioengineer new pharmaceuticals and agricultural crop protection products.[23]

In the graph, greater protection for first-generation innovation corresponds to a higher royalty. With perfect protection from imitation, the first-generation innovator can charge a royalty that maximizes its profit. Higher royalties, up to the profit-maximizing level r^*, encourage greater investment in the first-generation technology. However, higher royalties also lower the profit that follow-on innovators can earn. The royalty burden decreases the incentive to invest in follow-on technologies that require the right to use the innovation made in the first generation.

The optimal royalty strikes a balance between the reward for investment in first-generation innovation and the cost imposed for investment in follow-on innovation. Because the first-generation innovator achieves a maximum profit at r^*, a small reduction in the royalty below

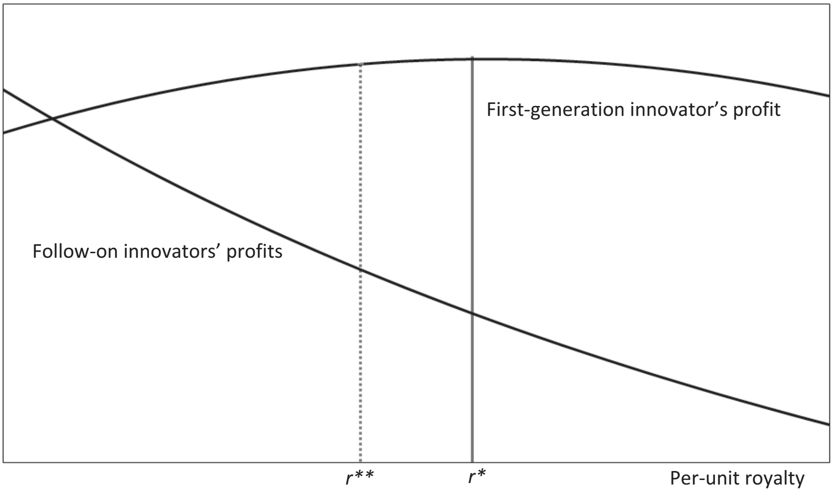

Figure 4.1
A small reduction in the royalty below r^* increases the follow-on innovators' profits and has only a small effect on the first-generation innovator's profit.

r^* would have little effect on the first-generation innovator's expected profit. The first-generation innovator's profit is almost independent of the royalty when it is close to r^*. In contrast, a small reduction in the royalty would have a large positive effect on profits and R&D investment for follow-on innovators. Consequently, total R&D investment for both first-generation and follow-on innovation would be higher at a royalty r^{**}, which is somewhat below r^*, than at the first-generation profit-maximizing level r^*. The total economic welfare also would be higher at r^{**}, for two reasons: Total investment would be higher and a lower royalty implies lower prices for follow-on innovations.[24]

A royalty that is below the level r^* that maximizes the profits of the first-generation innovator can be achieved by patent policies that limit intellectual property protection for the first-generation technology, such as by narrowing the scope of follow-on inventions that would infringe the first-generation patent. Alternatively, antitrust policies can prohibit conduct by the first-generation inventor that may exclude or raise costs for subsequent innovators. These policies may include limiting the use of exclusive territories for practicing follow-on innovations or prohibiting contractual terms, such as obligations for follow-on innovators to grant licenses to their innovations for use by the first-generation innovator.

Policies that allow an original inventor to exclude competitors or raise their costs provide more compensation for early innovators, at a cost of lower compensation for new competitors.[25] In principle, antitrust enforcement should account for the cumulative nature of innovation and the interaction between current R&D investment and future innovation incentives, but in actuality, that is a tall order. It is hard enough to predict the consequences of a merger for current R&D investment decisions. It is even more difficult to predict the consequences of these decisions for the future evolution of an industry and future innovation incentives.

There is at least one respect in which the cumulative nature of innovation has clear implications for antitrust enforcement. This has to do with policies that affect the cost or access to existing technologies that interoperate with other technologies. Firms that develop innovative system products can exclude future innovators by designing their systems to be incompatible with complementary products. Antitrust enforcement has a role to play by facilitating access to existing technologies for follow-on innovators.

The US Department of Justice and the Federal Trade Commission settled some merger cases by requiring the parties to grant nonexclusive licenses that cover a broad range of technology. While compulsory licenses at noncompensatory rates may undermine innovation incentives for products that are easily copied, they can also facilitate innovation by reducing the cost of assembling necessary intellectual property rights, eliminating costly infringement litigation, and neutralizing efforts by firms that use intellectual property rights to block new competition.

A requirement to license intellectual property at low royalties can harm innovation by depressing incentives to create first-generation innovation by more than it promotes follow-on innovation, but that need not be the case. James Bessen and Eric Maskin develop a model of cumulative innovation in which discoveries have spillover benefits that promote future discoveries.[26] They find that innovation and consumer welfare can be higher if discoveries have no patent protection (which is equivalent to compulsory licensing at zero royalties) than in an alternative scenario in which discoveries have strong patent protection. Case studies discussed in chapter 7 also suggest that compulsory licensing can promote economic welfare in some circumstances.

4 Managerial Incentives and Organizational Failures

Up to this point, we have equated innovation incentives to the profits that a firm can earn by investing in R&D. These profits capture economic forces such as the Arrow replacement effect and Schumpeterian limitations on the ability of firms to appropriate the benefits from their innovations. But most R&D effort is exerted in hierarchical organizations by managers and staff who have compensation schedules and other motivations, such as security of employment, that are only indirectly related to the economic returns from innovation outcomes. In addition, cognitive biases can shape the R&D paths chosen by the managers of these organizations.

The next subsection reviews theories that address information flows within an organization to see how they complicate the relationship between competition and innovation incentives. While it is useful to review these theories, they do not converge on predictions that justify a substantial departure from the lessons learned from profit-based theories of innovation incentives.

Managerial incentives

Economists have explored managerial incentives by treating the firm as a two-level hierarchy, with a principal (the firm owner) who has a claim on the firm's profits and an agent (the firm manager) who is paid by the owner and performs tasks on her or his behalf. There is a separation between ownership and control. With perfect information, the owner would simply direct the manager to perform the desired actions. In that case, the relationship between competition and innovation incentives does not depend on whether the owner of the firm performs R&D or whether she or he delegates that task to the manager.

In more realistic organizational settings, employees have better information than owners about the costs and payoffs from different projects, firm owners can only imperfectly monitor and enforce the activities that they would like their employees to do, and employees care more about their own welfare than owners' profits. Employees may want to work less, enjoy perks, avoid or take too much risk, or build empires that are unrelated to profitability. The owner's problem is to design a compensation schedule that incentivizes the manager to create value for the owner, but the separation of ownership and control implies that incentive schemes are imperfect in these situations. A key question is whether competition makes it easier or more difficult for owners to align managerial incentives with the owner's objectives.

Studies that have explored these agency theories reach different conclusions. Several theoretical studies find that competition reduces, but does not necessarily eliminate, the ability of managers to pursue their own objectives. Managerial choices that do not maximize profits increase the risk of bankruptcy, which harms managers as well as owners. Competition increases the risk of bankruptcy and should make managers less inclined to engage in unproductive activities.[27] Furthermore, when market outcomes are correlated, competition provides the owner (or owners) with information about firm performance, which she or he can use to benchmark managers.[28] These results do not imply that an increase in competition necessarily increases R&D incentives. Rather, they imply that the relationship between competition and innovation derived for owner-managed firms is more likely to hold for managerial firms when product markets are competitive.

Unfortunately, the concept that product market competition mitigates the problems caused by the separation of ownership and control is not robust to other reasonable assumptions.[29] Product market competition lowers profits, which squeezes the ability of owners to offer managerial compensation that induces desired efforts. If competition increases the risk of bankruptcy, with adverse consequences for employees, managers may avoid projects that have risky payoffs. Alternatively, managers may choose risky R&D projects if they profit handsomely when the project succeeds but bear only a fraction of the cost of failure.

It is useful to note that markets provide contrasting and generally ambiguous incentives for investment in risky R&D, even in the absence of complications from the separation of ownership and control. On the one hand, firms have incentives to invest in risky R&D projects because they profit most when they are the first to invent, but that has little social benefit if a second innovator would have soon followed.[30] This suggests too much incentive to invest in risky R&D. On the other hand, a firm, but not society, benefits when it innovates and takes sales from another innovator. This business-stealing effect is larger for incremental innovations, which implies a market bias in favor of low-risk projects and against risky R&D.[31]

Managerial theories of R&D investment have contrasting implications for R&D investment incentives that do not necessarily differ from the predictions of simpler economic models. For these reasons, I do not pursue managerial explanations for R&D investment further. A different theory, which has captured the attention of many corporate executives, stresses imperfections in the communication of information within an organization about the value of R&D decisions. This theory is the subject of the next subsection along with related theories of organization adjustment costs that affect innovation incentives.

Organizational failures and adjustment costs

A large body of literature in the tradition of business strategy stresses how organizational limitations affect firms' R&D decisions. Clayton Christensen, a Professor at the Harvard Business School, sent shock waves through the world of corporate executives when he advanced the thesis that entrenched firms fail to make important innovations because they focus *too much* on the immediate needs of their customers.[32] According to Christensen, incumbents are not complacent firms that rest on their past accomplishments and ignore their customers' needs. Instead, he argued that incumbents are attentive to customers' desires to make their existing products better, but at the cost of missing the wave of disruptive new technology.[33]

Christensen's "innovator's dilemma" is that leading incumbents fail when faced by disruptive technologies because they continue to make the same kinds of decisions that made them successful. By focusing on the technologies that brought past success, they fail to encourage the adoption and use of new disruptive technologies, which ultimately leads to their demise. It is a type of managerial failure, although not one that results from a separation of ownership and control. Rather, it presumes that managers of established firms have beliefs based on years of industry experience reinforced by communication flows with customers and each other, and cognitive biases that cause them to focus on incremental improvements to existing technologies and to overlook potentially disruptive innovations. These new technologies begin as inferior alternatives to existing products but have the potential to improve and displace those products. When that happens, it is too late for the entrenched firms to catch up, and they are replaced by new market leaders.

Christensen supported his thesis with anecdotal evidence from industries such as computer hard disk drives (HDDs), where new generations of disk drives were sponsored by an industry outsider that ultimately displaced the former market leader. The main quality attributes of HDDs are the disk diameter (also called the form factor), the capacity, the durability, and the speed at which data can be recorded on and retrieved from the drive. Initially, smaller HDDs were inferior to larger ones. Established companies focused on making improvements to drives with existing form factors rather than pursuing these smaller, but initially inferior designs. New competitors focused on drives with smaller form factors, and they ultimately displaced incumbents as the industry gradually transitioned from 14-inch to 2.5-inch form factors.

There are many other examples of established firms that missed out on new technologies. Xerox PARC, a division of the Xerox Corporation, was responsible during the 1970s for important innovations such as the Ethernet, the graphical user interface (GUI) for personal computing, the computer mouse, and the laser printer. Xerox commercialized the laser printer but not these other innovations, presumably because the parent company was more focused on serving its existing base of office equipment customers.

Christensen's *The Innovator's Dilemma* is one of the best-selling business books of all time. The thesis is provocative and may explain the failure of leading firms to manage disruptive innovations in some circumstances. Yet, the theory does not explain why some companies fail in the face of destructive innovation, while others prosper. Some have questioned whether his theory is a good description of innovation patterns in many industries, including the hard disk drive industry.[34]

Moreover, there are alternative explanations for the failure of firms to catch the wave of creative destruction, which are based on more conventional economic thinking. The Arrow replacement effect tells us that earnings from existing products are a deterrent for innovation. For example, Polaroid did not lag in digital technology. Consistent with the science-driven philosophy of its founder, Edwin Land, the company invested heavily in R&D and held a strong patent position in digital imagery in the 1980s. But Polaroid was reluctant to give up a business model that produced a reliable profit flow from sales of instant film, which it would lose if it led a transition to digital imagery. By the time the company realized that digital cameras would replace instant film photography, the competitive tide had turned and Polaroid was left behind.[35]

Polaroid's demise is not a good example of a company that missed a technological opportunity because it was too focused on the immediate needs of its customers. Polaroid did not want to give up the profits that the company earned from its sales of instant film. Although organizational failure may have been a contributing factor, its conduct is consistent with the economic force of the Arrow replacement effect.

It is not necessary to call on organizational failures to explain why established firms often fail to catch the wave of creative destruction. The Arrow replacement effect explains why dominant firms may be reluctant to embrace innovation opportunities that would destroy their existing profits. In other circumstances, economic theory tells us that dominant firms can have incentives to employ preemptive strategies, such as aggressive patenting to avoid competition and maintain their monopoly profits. These preemption incentives apply only to incremental innovations and are absent for drastic innovations that would disrupt an industry.[36] Unless incumbents derive some unique benefit from a drastic innovation, or high barriers protect them from competition, it is not surprising that drastic innovations will emerge from new competitors, particularly in technological fields where R&D does not require highly specialized and expensive assets, thereby allowing many potential innovators.

Christensen's theory of demand-driven management failure also overlooks the positive role that incumbents can play as innovation incubators. Many successful new competitors are led by researchers and entrepreneurs who gained experience working for established companies. Intel was founded by engineers who worked at Fairchild Semiconductor International, which in turn was founded by dropouts from the Shockley Semiconductor Laboratory, a division of Beckman Instruments, and before that, Bell Laboratories. April Franco and Darren Filson demonstrate that the movement of experienced top executives from established firms contributed to the adoption of some of the more important innovations in the hard disk drive industry.[37]

Acquisitions also shaped the course of the industry. Seagate, one of the more successful suppliers of hard disk drives, purchased Control Data, Maxtor, and Conner, among others. Complementary assets were also a significant factor in Seagate's success. Seagate survived in part because it invested heavily in developing a supply chain that was capable of responding to the explosive growth of the computer market in the 1990s and 2000s.[38]

Rebecca Henderson and Kim Clark advanced an alternative organizational hypothesis to explain innovation successes and failures.[39] In contrast to Christensen's theory that attention to consumer demand causes dominant firms to miss out on disruptive innovations, Henderson and Clark focus on a firm's existing capabilities and how they match the requirements of new products or services. Organizational capabilities are costly to acquire and adjust. For Henderson and Clark, a radical innovation is one that requires very different organizational capabilities than the capabilities that a firm currently has. This supply-side definition of a *radical* innovation contrasts with Arrow's definition of a *drastic* innovation, which makes the technologies it replaces obsolete. According to Arrow, whether an innovation is drastic or incremental depends on consumer demand and the characteristics of new and old technologies. From the perspective of Henderson and Clark, whether an invention is radical or incremental depends on whether the innovation requires new technical and commercial skills or new approaches to problem-solving for the innovating firm.

Henderson and Clark's theory of radical innovation can explain why some dominant firms fail to innovate even in circumstances in which they have incentives to preempt potential rivals. For instance, RCA possessed the core capabilities to develop the first commercial portable transistor radio and demonstrated a prototype as early as 1952, but it was Sony, not RCA, that first commercialized the transistor radio—using technology licensed from RCA.[40] RCA had expertise in the components necessary to produce a commercial portable transistor radio, but innovation required the company to think differently about how the components could be integrated to produce a commercial product.[41] Intel curtailed research in reduced instruction set computing (RISC) microprocessors because they would not be backwardly compatible with its existing complex instruction set computing (CISC) microprocessor architecture.[42] New competitors, such as ARM, introduced RISC microprocessors for mobile technologies. The movement from CISC to RISC was a radical innovation given Intel's commitment to backward compatibility.

Rebecca Henderson tested alternative explanations for innovation using data on forty-nine photolithography projects undertaken by nineteen firms.[43] Semiconductor photolithography is an optical process used to transfer device and circuit patterns onto a semiconductor substrate for the manufacture of integrated circuits. She found support in the data for the economic theory that incumbent firms have incentives to invest in R&D to stay ahead of rivals when innovations are incremental and not radical in the organizational sense. However, for some innovations that involved radical organizational adjustments, incumbents failed to deter competition, even though they invested heavily in R&D for these new technologies.

Radical innovations do not have to doom incumbent firms. Canon and Fuji successfully navigated the transition from film to digital photography, while Kodak and Polaroid failed. Apple has successfully managed waves of potentially destructive innovations, perhaps because the company has not focused on consumer marketing research.[44] Other firms have attempted to minimize institutional inertia by isolating R&D teams from existing organizational distractions.[45]

5 Concluding Remarks

Innovation is fundamentally a human endeavor, but the power of economic incentives, even for the creative act of innovation, should not be dismissed. Arrow's replacement effect identifies an important economic force where profits that are put at risk from innovation create a drag on innovation incentives. Preemption incentives can turn this around in some circumstances.

Agency theories that account for the separation between corporate ownership and managerial control complicate the effects of profitability on firms' R&D decisions. Presently, there is insufficient convergence among these theories to reject more conventional profit-maximizing behavior. Organizational theories that emphasize cognitive limitations and the confines of existing corporate structures offer useful perspectives on the ability of firms to maintain technological leadership, but they often can be consistent with purely economic incentives for R&D investment.

A corporate decision not to pursue a new technological opportunity might be explained by information limitations or bureaucratic inertia. It is also a rational economic decision if the expected return from the new technology is not large enough to cover the cost of transitioning the company to the organizational structure that is more effective at supporting investment in the new technology. Demand-side cognitive influences (which Christiansen calls the "innovator's dilemma") suggest that established firms are more likely to pursue incremental than drastic innovations. Economic theories suggest that dominant firms can have incentives to invest more than rivals in R&D, but only for incremental innovations that do not make existing products obsolete. It is costly for organizations to adapt to radical innovations. Similarly, the Arrow replacement effect implies that profits from existing products that are at risk from innovation create opportunity costs for R&D. All of these theories imply that it is critical that established firms do not erect artificial barriers to new competition, because disruptive innovations are often likely to come from rivals that are new to the industry.

Competition policy for innovation should account for the types of innovations at issue, the technological capabilities of the merging firms, and the nature of competition in the affected industry. Although the next chapter shows that mergers can have adverse consequences for innovation incentives, a merger can promote innovation in some circumstances by allowing the merged firm to appropriate greater value from innovations or by facilitating technology transfer between firms with different technological capabilities. The ability of merger partners to benefit from their respective technological capabilities depends on whether innovations require organizations to adapt to radical change. Mergers are unlikely to have appropriation benefits for new technologies that are radical as defined by Henderson and Clark.

The so-called right model of innovation incentives depends on multiple factors, including market structures, technological opportunities, organizational capabilities, information spillovers, and other dimensions such as whether innovations are products or processes, or create durable goods. The details matter. A merger that reduces competition may dull incentives for innovation in one industry, and yet increase incentives for innovation or have no effect at all on another industry.

"We have made progress by finding partial theories.… The possibility that there is an infinite sequence of more and more refined theories is in agreement with our experience so far." This quote is from the eminent theoretical physicist and cosmologist, Stephen Hawking, writing in *The Theory of Everything*.[46] Although innovation economics is far removed from cosmology, his statement applies to theories of innovation incentives. Our existing models are partial theories that work in limited circumstances, and it is likely that more refined theories will follow.

The survey of the economics of innovation in this chapter and the preceding chapter may disappoint those who are looking for a general theory to guide antitrust enforcement for innovation. The empirical studies described in chapter 6 do not entirely resolve this dilemma because they find that competition often—but not always—promotes innovation. But first, in chapter 5, we turn to enforcement by antitrust authorities for mergers and acquisitions that affect innovation and future price competition.

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

ARCADIA
A charitable fund of Lisbet Rausing and Peter Baldwin

This book was set in Palatino by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Gilbert, Richard J., 1945- author.
Title: Innovation matters : competition policy for the high-technology economy / Richard J. Gilbert.
Description: Cambridge, Massachusetts : MIT Press, [2020] | Includes bibliographical references and index.
Identifiers: LCCN 2019039525 | ISBN 9780262044042 (hardcover)
Subjects: LCSH: High technology industries. | Competition. | Antitrust law--Economic aspects. | Consolidation and merger of corporations--Law and legislation--Economic aspects.
Classification: LCC HC79.H53 G56 2020 | DDC 338.8/2--dc23
LC record available at https://lccn.loc.gov/2019039525