



Working Paper Series No. 16

Profiting from Innovation in the Digital Economy: Standards, Complementary Assets, and Business Models In the Wireless World

David J. Teece¹

23 August 2016

¹ Thomas Tusher Professor of Global Business & Director, Tusher Center on Intellectual Capital, Haas School of Business, University of California, Berkeley; Chairman, Berkeley Research Group, LLC. Email: dteece@thinkbrg.com.

I wish to thank Asish Arora, Gary Pisano, Greg Linden, and Benoit Reillier for helpful comments and valuable insights.

TABLE OF CONTENTS

| | |
|--|-----------|
| I. Introduction | 3 |
| II. PFI Revisited..... | 3 |
| III. Expanding the PFI Framework | 5 |
| A. General Purpose Technologies, Enabling Technologies, & the Appropriability Challenge.. | 5 |
| 1. General Purpose Technologies | 5 |
| 2. Enabling Technologies | 6 |
| 4. Enabling technologies in mobile communication ecosystems | 10 |
| B. Complementary Assets | 12 |
| 1. Varieties of complementarity | 12 |
| 2. PFI and complementarity reconsidered | 16 |
| 3. Platforms and ecosystems | 18 |
| C. Multi-level PFI..... | 20 |
| D. Modularity and Industry Performance | 23 |
| E. Business Model Configuration: Licensing and Alternatives | 25 |
| F. Standards | 27 |
| 1. Standard setting and standards development | 27 |
| 2. Standard essential patents (SEPs) and royalties | 29 |
| IV. Management and Policy Conclusions | 30 |
| References | 33 |
| Appendix: The Digital Revolution and its Effects..... | 37 |

I. Introduction

In this paper, I look anew at the Profiting from Innovation (PFI) framework laid out in Teece (1986, 1988a, 2006). The questions addressed in the earlier treatments—what determines which firms profit from an innovation, and which firms earn only meager (and possibly negative) returns—have enduring relevance for both management and public policy.² If anything, the importance of the issues is amplified as “digital disruption” gains pace and as the manner of innovation changes. In particular, the mega-convergence of certain industries that is being driven by the merging of wireless and Internet technologies requires that one open the aperture of business and economic inquiry from the innovation of individual products and processes to innovation within ecosystems and across the upstream and downstream levels of competition in an industrial system.

The wider aperture of inquiry in this paper turns out to reinforce the key elements of PFI. Intellectual property, the nature of knowledge, complementary assets, standards, and timing all remain center stage. What is brought into sharper focus is additional granularity with respect to different types of complementary assets and the ways they impact the capture of value from innovation. Additional clarity is also provided with respect to how a collaborative standards development process enables downstream innovation, and how standards-based business ecosystems emerge and compete. In particular, the challenges of capturing value from enabling technology are recognized, and policy implications are explored. This essay considers the impacts of changes wrought by the digital revolution, the activity of standards organizations, the presence of enabling technologies, and the growing importance of complementary assets and technologies in the information and communication technology (ICT) sectors and beyond. These phenomena have enhanced salience in the wake of the digital revolution and the associated convergence of industries that is described in the appendix.

II. PFI Revisited

The Profiting from Innovation (PFI) framework (Teece, 1986) was launched thirty years ago in a very different technological and business environment than most companies face today. Since PFI has become important to the technology management and business management literatures, it is worth revisiting periodically in order to see what revisions might be necessary. I first did so ten years ago (Teece, 2006), when I sketched a number of elaborations and extensions in response to shifts in the environment since 1986. These included further development of the multi-invention context for innovation, the incorporation of a richer understanding of network effects, and the consideration of business models engendered by the launch of the Internet. I also discussed the growing importance of complementary technologies and business ecosystems and their impact on PFI.

In the intervening decade, as mentioned earlier and elaborated in the appendix, the techno-business environment has shifted still further. The Internet is no longer a utility consulted just from user desktops. It is increasingly pervasive, accessed interactively by users on the go and extended to sensor-equipped terminals anywhere and everywhere. Means of communication have also evolved, from phone and email toward messaging apps that also serve as portals for shopping and a host of other services.

² Before PFI, capturing value was often considered just a matter of pricing strategy. Setting prices remains important, but value capture depends on more fundamental considerations.

PFI addressed a puzzle that had not been well explained in the previous literature, namely: why do highly creative, pioneering firms often fail to capture much of the economic returns from innovation? Apple's iPod was not the first MP3 player, but it has a commanding position in the category today. Merck was a pioneer in cholesterol-lowering drugs (Zocor), but Pfizer, a late entrant, secured a superior market position with Lipitor. At first glance, it is tempting to say that these examples reflect the result of Schumpeterian gales of creative destruction where winners are constantly challenged and overturned by entrants. But the cited cases and countless others involve mostly incremental/imitative entrants rather than the radical breakthroughs of Schumpeterian competition. And, of course, there are many cases where first or early movers captured and sustained significant competitive advantage over time. Genentech was a pioneer in using biotechnology to discover and develop drugs, and 30 years later was the second largest biotechnology firm (and, the most productive in its use of research and development dollars) right up to its acquisition by Hoffmann-La Roche in 2009. Intel invented the microprocessor and still has a leading market position more than 30 years later.

The focus of the 1986 article was on a single, autonomous innovation that was commercially viable. The paper thus ducked one question—why inventions so often fail to succeed in (or even reach) the market—and focused instead on how the spoils are divided once positive net value is within sight. The paper also focused on product innovations rather than process improvements, creative output, or other valuable intellectual property. These and other simplifications made the analysis tractable.

The PFI framework provides an explanation as to why some innovators win in the marketplace while others lose out—often to rank imitators—and why it is not inevitable that the pioneers will lose. The fundamental imperative for profiting from an innovation is that unless the inventor/innovator moves down an improvement path and enjoys strong natural protection against imitation and/or strong intellectual property protection, then potential future streams of income are at risk. The relevant appropriability regime is thus critical to shaping the possible outcomes.

Appropriability regimes can be "weak" (innovations are difficult to protect because they can be easily codified and legal protection of intellectual property is ineffective) or "strong" (innovations are easy to protect because knowledge about them is tacit and/or they are well protected legally). Appropriability regimes for many digital businesses are often quite weak. Business models cannot rely heavily on intellectual property (IP) to capture value because IP is generally not self-enforcing; of instances of patent infringement and trade secret misappropriation must be identified and negotiated or litigated, often at great expense. Even patents rarely, if ever, confer strong appropriability, outside of special cases such as new drugs, chemical products, and rather simple mechanical inventions (Levin et al., 1987). Many patents can be "invented around" at modest costs (Mansfield, 1985; Mansfield et al., 1981). Often patents provide little protection because the legal and financial requirements for upholding their validity or for proving their infringement are high, or because, in many countries, law enforcement for intellectual property is weak or nonexistent. In terms of preventing imitation or bringing infringers into licensing arrangements, a patent is merely a passport to a journey down the road to enforcement and potential royalty streams. Of course, a large portfolio of patents can prove valuable for cross-licensing deals with rivals that help reduce the likelihood of costly litigation. To help with appropriability, the inventor of a core technology can also seek complementary patents on new features and/or processes, and, in some cases, on designs.

The original PFI article also developed the linkages between complementary assets and profitability. This included not only the supply chain inputs needed to embody the innovation in a product or service, but also other products, technologies, and services needed to produce value in consumption. In some cases, such as mobile phones needing cellular networks, the complement is an infrastructure element that is available to all firms. In other cases, such as hardware needing software, the complement (e.g., an operating system) might be so specific to the innovation that the innovator had to provide it in-house or risk becoming over-reliant on a complementor. In a few other cases, such as the need of operating systems for applications, the platform concept, in which success involves bringing a large number of complementors into the fold, was already present. A successful open platform avoids the “bottleneck” problem of allowing other entities to control required elements of the system (e.g., a key productivity application, or a hit videogame) by encouraging competition among the complementors.

However, the PFI framework now needs a wider aperture to look at issues beyond the single product, single level (of an industry) innovation that was the focus in 1986. This is notwithstanding the robustness of the original framework. The reason is that additional color and explanatory power can be added through highlighting issues that digital convergence is presenting.

The additions I will sketch out do not invalidate the logic of the PFI model. The core independent variables in the PFI model—the strength of the appropriability regime complementary assets, complementary technologies,, standards (and associated installed base effects), and timing—are more relevant than ever for a world where firms increasingly offer platforms to which complementors can add incremental value with their own innovations.

III. Expanding the PFI Framework

The brief description advanced earlier (and expanded in the appendix) of the digital revolution has highlighted elements of PFI that need further elaboration, and suggested the need to also look at the level of industry at which innovation is taking place. The framework can be extended to speak to the special difficulty associated with appropriability when enabling technologies and/or general-purpose technologies are being developed and deployed. A more granular view of standards and complementary assets is also required, along with the analysis of related business ecosystem and business model issues.

I. A. General Purpose Technologies, Enabling Technologies, & the Appropriability Challenge

1. General Purpose Technologies

Bresnahan and Trajtenberg (1995) identify a GPT by the following three characteristics: (1) Pervasiveness: a GPT is used in—and impacts—many downstream sectors; (2) Technological Dynamism: the GPT has potential for improvement over time; and (3) Research Productivity: the GPT should make it easier to invent new products or processes in downstream applications. In other words, GPTs have economy-wide effects, get even better, and spawn other innovations. While some non-GPT technologies may possess each of the characteristics to some extent, a GPT will be distinguished by its cumulative economic impact (Jovanovic and Rousseau, 2005). They can be either endogenous or exogenous to the economic system (Lipsey, Bekar, and Carlaw, 1998). A GPT can be a product, a process, or an organizational system. Examples that Lipsey et al. (1998) consider “clear and dramatic” are limited to

printing, bronze, made-to-order materials, the waterwheel, steam power, electricity, the internal combustion engine, railways, motor vehicles, lasers, and the Internet.

The aggregate effect of a GPT occurs over a protracted length of time. Its invention is often the result of collaboration between individuals with disparate skills. The inventions may not be immediately identifiable as GPTs; they might be enabling technologies or less until improvements occur and new uses are discovered. For example, when the laser was invented around 1960, it was of scientific interest but had no obvious application. Today, lasers are ubiquitous, being implemented in applications ranging from supermarket checkout stands to military battlefields, with other uses in medicine, telecommunications, manufacturing, entertainment, and more.

The widespread applicability of a GPT will often require a decade or more to become evident, but over time the cumulative effects can be significant. There have been pioneering cliometric studies on the social benefits of GPTs such as steam engines (von Tunzelman, 1978), British railways (Hawke, 1970), and American railroads (Fogel, 1964; Fishlow, 1965). Hawke (1970), showed that the railroads saved between 7% and 11% of British national income in 1865.

2. Enabling Technologies

The threshold for a GPT is very high, but there's a similar category, enabling technologies (present but not well defined in the literature), that can be thought of as junior GPTs, meeting criteria (2) and (3), above, but not necessarily having measurable economy-wide impacts.³ While the list of GPTs is relatively short, there are hundreds, if not thousands, of enabling technologies. Each one might not be thought of as a "game changer" by economic historians endeavoring to understand economic growth, but they are nevertheless important to particular firms and industries. They can often be disruptive to the status quo and generate very considerable economic benefit and social surplus.

In some cases, enabling technologies can have quite massive effects. Containerization of cargo shipping, which greatly reduced the transportation cost of many types of products, had a large economic impact, even though container technology has not advanced much technologically over the years, which possibly disqualifies it as a GPT but not necessarily as an enabling technology. An econometric study by Bernhofen, El-Sahli, and Kneller (2016) found, after controlling for the effect of trade liberalization agreements, that the adoption of container shipping over the period 1967-1982 by a group of 22 OECD countries raised the value of trade by 1,240%.

An enabling technology can be used to drive technological change in an industry. The utility of the enabling technology is not exhausted by embedding it in a single product or even a single system; it can be licensed to a range of downstream customers for their own products and services. Enabling technologies were not the focus of PFI, which looked at commercially viable product innovations and implicitly assumed that only a narrow range of inventions was incorporated into a single product.

³ Enabling technology is closely related to "generic technology," which, according to Keenan (2003: 132) can have any of three meaning: "(1) A class of closely related technologies; (2) A technology the development of which will have implications across a range of other technologies; or (3) A technology the exploitation of which will yield benefits for a wide range of sectors of the economy and/or society." These are distinct from "generic knowledge" (Nelson, 1989). Generic knowledge comes from hands-on experience and is often a byproduct, rather than a goal, of formal innovation; it is easily discovered by peer professionals and generally cannot be protected as intellectual property.

One enabling technology (which might eventually become a GPT) is artificial intelligence (AI). AI encompasses a range of software techniques employed to teach computers to sense, reason, interpret, communicate, and make decisions in a human-like manner. It gained currency in the 1950s, but needed cheaper and faster computing power, especially developments in the complementary hardware technologies of processors and memory, to become commercially interesting. AI-based technologies can already recognize faces, understand speech, and drive vehicles. Regulators use AI to sort Internet chatter and zero in on fraud, illegal activity, and terrorist plots. The applications for AI are limited only by economics and imagination, and its cost continues to fall as it becomes available as part of the capabilities offered for hire by the providers of “cloud” (online) computing.

A closely associated technology is machine learning, a category of software tools with a relatively narrow goal of learning known properties from data with a minimum of pre-programmed rules. One sub-type, deep learning, performs abstract tasks such as recognizing images, using neural networks and other tools that try to mimic the operation of the human brain. Deep learning helps Google improve the search results it provides users and drives the recommendation engines used by many online retailers. The most viable business models for capturing value from machine learning are likely to be firm-specific because an algorithm is likely to learn faster and produce more accurate outputs when it is narrowly focused (Norton, 2016).

Enabling technologies vary not only in type but in ownership. They may be owned or controlled (in part or in whole) by private companies, universities, or consortia. Many software-based enabling technologies, such as deep learning, are often available as open-source software. The use of technology from the open source community, where appropriability is inherently limited, places a premium on the ability of the licensee to apply PFI principles, such as by assembling valuable complements. In the case of deep learning, for example, this requires access to large quantities of high-quality data for training along with a distinct product or service. In the remainder of this section, however, I will focus primarily on cases where enabling technology is owned, at least in part, by the focal organization.

A characteristic of enabling technologies and GPTs is that there are large positive spillover effects of two kinds: static and dynamic (Carlaw and Lipsey, 2002). Static spillovers are standard externalities that don't lead to any change in behavior by other economic agents, either at the time or in the future. Dynamic externalities occur when the innovation alters the current and future value of existing technologies and also enables further technological opportunities for other agents.

Spillovers, as discussed in the next section, imply that appropriability is especially weak for enabling technologies and GPTs. This implies that private enterprise will tend to underinvest in creating them, absent government support. It also implies that the designers and enforcers of intellectual property regimes need to be sympathetic to the appropriability challenges that developers of enabling technologies and GPTs face.

3. Appropriability

Even though the social contribution of GPTs and enabling technologies is very high, the pioneers can typically extract only a tiny fraction of the value. Despite the flagging of this problem by Bresnahan and Trajtenberg (1995), there has been little effort to understand it or develop policies to mitigate it. Bresnahan and Trajtenberg (1995) identified two ways in which the social benefits of a GPT may be curtailed by appropriability issues. First, the capture of a GPT's value from downstream sectors may be

limited by business or regulatory barriers. Second, the uptake of the GPT by multiple downstream sectors may be too low, in the sense that if these sectors were coordinating their activity, they would see that faster uptake of the technology would support its more rapid improvement by the upstream innovator. The practical significance of this second point is that GPTs develop faster when there's a large, demanding application sector (avoiding the need to coordinate smaller ones), such as the U.S. Defense Department's early purchases of transistors and microprocessors. Helpman (1998, p.4) characterized the undersupply problems as follows:

GPTs introduce two-types of externalities: one between the GPT and the application sectors; another across the application sectors. The former stems from the difficulties that a GPT inventor may have in appropriating the fruits of her invention. When institutional conditions prevent full appropriation, the GPT is effectively underpriced and therefore, undersupplied. The latter stems from the fact that, since the application sectors are not coordinated, each one conditions its expansion on the available general-purpose technology. But if they coordinated a joint expansion, they would raise the profitability of the GPT and encourage its improvement. A better GPT benefits them all.

A key (policy) implication is that inventors of GPTs have great difficulty appropriating the fruits of their investment, and management will be challenged in finding workable revenue architectures. Owners of enabling technologies face similar challenges. As a result, underinvestment in GPTs and enabling technologies is a near certainty. Public policy needs to redress this in some fashion.

Several factors conspire to restrict the ability of owners of GPTs and enabling technologies to capture value from downstream implementers. First, designing a business model and/or controlling the necessary complementary assets and technologies to internalize more of the spillovers is typically beyond the scope of what a single firm can do. A case in point is Pilkington Glass and its invention of float glass, which is, arguably, an important enabling technology. The float process, introduced in 1959, had reduced production cost versus the dominant plate glass process by more than a third by the time the first license was granted in 1962, and by about 75% by 1974 (Teece, 2000, Appendix Table B.2). It has also facilitated downstream innovations such as flat panel displays. Despite Pilkington's strong intellectual property position and the evident and quite substantial value of its technology, Pilkington licensed the float process on an exclusive territorial basis to all comers at a mere 6% of revenues. This reflects the fact that the invention was so universal in its attractiveness, and family-owned Pilkington so unprepared and unable to implement the technology by itself on a worldwide basis, that widespread licensing seemed the best alternative. As a result, a modest royalty for Pilkington effectively passed most of the benefits to manufacturers and consumers; Pilkington was able to garner only a tiny share of the total return to the technology for itself.

Second, the future value of a GPT or enabling technology may not be clear at the outset; and even if it is, there may be regulatory constraints that limit value capture. The transistor, developed by AT&T's Bell Labs, was protected by patents, but AT&T licensed the technology widely and almost royalty free. AT&T's logic appeared to be that, as a large-scale buyer of telecom equipment, it would benefit from the competition among its licensees, a strategic perspective that, at minimum, missed capturing value from the technology beyond the telecom industry (Levin, 1982). AT&T was also highly regulated, which raised additional barriers to commercialization and value capture.

Third, enabling technologies, by their nature, are intermediate inputs in the value chain, which are sold or licensed to downstream firms. Designing a business model to capture value more completely is far easier in consumer final-product sales than in interfirm intermediate product markets. Although there is theoretically only one monopoly rent to capture in the value chain, the chance of the innovator being able to design and implement an airtight business model to do so is low. For example, suppose there is a patented technology that when placed on a communications chip in a mobile phone will create \$100 of incremental value in terms of average consumer willingness to pay. The ability to extract this as “rent” from the seller of the enhanced component will be compromised by a licensing model and the inability to effectively price discriminate. Moreover, patents are not self-enforcing, so the value that can be captured by a license is affected by the costs associated with launching a licensing program, including the litigation costs from challenging the patent infringers and trade secret misappropriators. Allowing a larger slice of economic value to flow to the licensee may help launch a licensing program, but this in turn will reduce the innovator’s share of the profits.

Fourth, the bargaining position of the innovator is undermined when it lacks a credible threat to exclusively commercialize the technology and practice the patent(s) itself because it lacks the relevant assets and capabilities. This can be true for any technology, but it is especially true for enabling technology with applications in several different industries. Hence, the broader the applicability of an enabling technology across industries, the less complete the rent appropriation is likely to be. The original PFI framework applied to a discrete innovation that was assumed to be commercially viable. Licensing was the recommended strategy if there was strong intellectual property protection, although this was said to be a rare circumstance. With weak appropriability, the basic idea was to attempt to capture value by acquiring or controlling the required complementary assets for in-house commercialization of the innovation. Otherwise, the innovator would have to either (1) build them, which is slow and expensive, or (2) form partnerships. Either strategy would drain some of the value away from the innovator.

The very nature of enabling technologies today is that they are likely to be embedded in multi-invention innovations for which other firms also hold relevant patents, so appropriability is inherently challenging. Capturing more value requires not only applying the technology but also driving the technology’s path forward and into derivative applications. This inherently involves engaging with partners because, as discussed in the four points above, the pursuit of full vertical integration and/or horizontal diversification is unlikely to be viable given existing capabilities and resources. For example, to make 3D printing (also called “additive manufacturing”) viable for industrial manufacturing, complementary advances in design software, metallurgy, machine tools, and more will be needed. For most innovations, relevant capabilities are often already available to tap externally. Outsourcing can shorten the path to successful commercialization, but will also increase the loss of value for the innovator. A start-up innovator will almost certainly need to rely on partners, and even most incumbent firms will lack some of the relevant capabilities to derive full value from the enabling technology.

In short, capturing value from enabling technologies is more challenging from a business model standpoint than capturing it from modest discrete innovations. Teece (1986) was very clear that if the innovator has to rely on partners to commercialize the technology, the profits from innovation would need to be shared. This is quintessentially the case with enabling technologies and GPTs.

4. Enabling technologies in mobile communication ecosystems

The PFI framework has to be up to the task of addressing the distribution of profits within complex ecosystems if it is to be a general framework. While PFI tells us that profits will go to the bottleneck assets, these are not always easy to identify, and they may shift over time. For example, an increase in data rates with a new generation of technology may increase profits to providers of video content (and to advertising vendors) while imposing greater capital and depreciation costs on network providers. And the fortunes of specific types of videos will face different appropriability regimes, such as whether they are treated as copyright protected.

In the standard PFI model, the drivers are (1) appropriability regime (i.e., strength of IP protection and difficulty of imitation), (2) complementary assets and technologies and related business model issues, (3) standards and installed base effects, and (4) timing. For the mobile ecosystem, however, one must also include, in recognition of the multiple ecosystems at work, (5) ecosystem strength. As a matter of theory, and assuming that there are no artificial (regulatory) barriers to entry, then the industry level (i.e., the location along the value chain, such as components, equipment, etc.) need not be critical and the above drivers remain the most powerful. The fates of individual firms and ecosystems will vary with the success of their efforts to exploit the opportunities created by each step-change in the enabling technology.

Within ecosystems, value captured by individual firms will depend on the scarcity of each firm's resources and the business models they adopt. A well-functioning ecosystem will allow multiple avenues for earning profits. In Google's ecosystem, Google subsidizes the ecosystem's technology but profits handsomely from advertising linked to user searches. In Apple's ecosystem, Apple profits from selling attractive and easy-to-use devices and from taking a share of the revenue from apps and content sold by their providers.

The wireless communications sector affords numerous examples that illuminate the issues of value capture for all types of participants, including the providers of enabling technologies. The key advances behind the digital communications revolution began as proprietary technologies that were subsequently codified in a series of wireless standards, each of which provides a step change improvement in communication performance. The improvements in mobile telephony standards from 2G in the early 1990s through the current 4G have been more than incremental, with each generation providing dramatic performance improvements in transmission capacity, service quality, congestion management, cell handover, and signal quality. 5G is on a path to be rolled out in 2020 with yet further boosts to bandwidth and speed that will see the emergence of new, wireless-based business models in industries dealing with massive quantities of data or mission-critical processing. Each generational advance makes new types of data services possible, and wireless is already posing a competitive challenge to fixed wireline services, which have historically been a cash cow for phone companies. The inverse is also true; if enabling technology fails to advance, or if new enabling technologies are not developed, then new products and services are less likely to be developed. Indeed, the most significant technological innovation undergirding the "grand convergence" is this dramatic improvement in mobile broadband capacity, which has helped feed growing demand as more people have gone mobile and new, bandwidth-hungry services like live video streaming are developed.

A key enabling technology supporting 3G was "CDMA," a 2G technology that was pioneered and promoted initially only by a single firm, Qualcomm. After spirited competition, engineers (from many

companies) who were participating in establishing a 3G standard at SDOs selected CDMA-based technology because it offered a far higher increase in bandwidth efficiency than was possible with the successor to GSM, which was the dominant 2G technology. In a similar vein, “OFDMA” (a newer, non-Qualcomm technology) was selected by standard-setting bodies as part of the 4G standard, allowing further increases in speed and bandwidth for mobile communications. These technical decisions involve a great deal of discussion, testing and negotiation. Each new generation of communication technology requires billions of R&D dollars and over a decade of elapsed time to develop and formalize in standards.

The technological advances across the generations of these wireless enabling technologies have been historically enabled by armies of engineers and billions of dollars in R&D investment at numerous companies, including AT&T, IBM, TI, Motorola, Siemens, and Ericsson. The other current major wireless technology developers are Qualcomm, which sells chips using its technology, Nokia, which is now almost exclusively a telecom equipment supplier, and InterDigital, a pure licensing company.⁴ Standards are thus a prime example of a multi-invention context in which any application uses innovations from numerous firms spanning multiple domains of knowledge.

The complexity and interdependence of the communication technologies underlying the mobile data revolution have caused multiple, connected business ecosystems to emerge (Teece, 2012). Specialized firms and even vertically integrated companies are no longer islands (if they ever were) connected to others only by market transactions. Each is now part of one or more constellations of business entities and organizations whose fates are tied technologically and/or competitively. Ecosystems have sprung up for a range of new services including streaming media, cloud computing, the Internet of things, and mobile payment systems. Each of these areas has mobile communications at its core, enabling the parts of the system to work together and enhancing its availability for end users.

In some cases, technologies specific to a particular function are designed into the wireless standard at an early stage. This requires that companies contemplating major new services form alliances and participate in standards development. But each generation of the wireless standard also serves as an enabling technology for services that have yet to be devised.

Today the mobile industry includes (1) dedicated technology and component firms like Qualcomm; (2) implementers of standards-compliant infrastructure equipment like Ericsson; (3) network service providers, like AT&T; (4) downstream implementers of standards-compliant mobile devices like Samsung; (5) software vendors like Apple and Google developing operating systems and apps; and (6) content providers, from individuals to media conglomerates. Many companies span two or more of these links in the value chain.

The total size of the mobile ecosystem in 2014 was, according to the Boston Consulting Group, around \$3.3 trillion. Table 1 shows how this revenue breaks down across eight categories, including technology, networks, hardware, software, and services. It is more difficult, however, to determine how profits are distributed among these groups, and the factors driving those returns.

⁴ The primary consumer-side implementers of the technology, Apple and Samsung, have not historically been major technology contributors to wireless standards.

Table 1: The Value of the Mobile Ecosystem, 2014

| <i>Category</i> | <i>2014 Revenue</i> |
|------------------------------|---------------------|
| Core Comms. Technologies | \$40 billion |
| Mobile Infrastructure Eqpmt. | \$220 billion |
| Mobile Site Operations | \$230 billion |
| Mobile Operators | \$1,010 billion |
| Components for Devices | \$260 billion |
| Device Manufacturers | \$450 billion |
| Device Retail | \$520 billion |
| Mobile Content and Apps | \$530 billion |
| Total | \$3,260 billion |

Source: BCG (2015), p.11

II. B. Complementary Assets

While complements were highlighted in the original PFI formulation, the analysis was broad brush. For instance, time was ignored. With the introduction of time, co-evolution becomes important. PFI implicitly recognized “rents” (profits) as flowing to the bottleneck assets, and that could be different complementary assets at different points in time.

While bottleneck assets loom large in PFI as a magnet for profits, rents also arise from combining different technologies in unique, value-enhancing ways that lead to gains exceeding the additive nature of standard complements, where the demand for one product (e.g., automobiles) can drive demand for another (e.g., gasoline). Even though the combination in a gasoline automobile engine involves complex technical interactions, it does not involve complex economic interactions. In many instances, technologies that were not obviously complementary, such as microchips and biological materials, can be combined (“orchestrated”) to produce entirely new products that are unique and valuable to users. This is a value-creating form of complementarity; standard notions of complements and scope economies fall far short in the innovation context. The refinement of the PFI framework, and a better understanding of ecosystems, begins by more closely examining the multiple types of complementarity.

1. Varieties of complementarity

Complements are pervasive throughout the economic system, and particularly in technology development and business transformation. Nevertheless, they are frequently ignored. However, they are central in PFI. Perhaps their neglect can be blamed on Schumpeter (1942), who stressed that “new combinations” of artifacts organized by the entrepreneur brought gales of creative destruction; but he never mentioned that, with complements, a rising tide can lift many boats. In the PFI framework, complements need to be considered with more granularity in order to illuminate value capture issues, particularly the ramifications of digital convergence.

At the heart of economic notions of complementarity is the notion, due to Edgeworth, that the marginal value of a variable increases with another variable. As simple as it sounds, there is much complexity to the concept of complementarity, prompting Nobel Laureate Paul Samuelson to say in 1974 that:

The time is ripe for a fresh, modern look at the concept of complementarity ... the last word has not yet been said on this ancient preoccupation of literary and mathematical

*economists. The simplest things are often the most complicated to understand fully.
(Samuelson, 1974, p.1255)*

The literature on complements has remained confused, and this section will make an effort to bring some clarity. It should be noted at the outset that it is common for two or more technologies often turn out to produce much more when practiced together.⁵ The first steam trains emerged when high-pressure steam engines were yoked to coal cars running on coal-mining hand cart rails. The lawnmower came from coupling a small gasoline engine to a miniature mechanical reaper. The laser and the computer were capable of much together (e.g., CD players); and of still more when optical fiber was added to the mix.

Economists tend to think of complementarity in terms of its effect on factor prices or on value from use (Carlaw and Lipsey, 2002). Innovation studies look instead at technological relatedness and the impact of new combinations of existing technologies. However, the full economic significance of complements lies not just with their ability to improve and support appropriability. Absent complementary technologies, many products simply won't get developed and launched. This was the case, for example, in the U.S. electrical supply industry at the end of the 19th century. The industry had a killer app—lighting—but was mired in a “war of the currents” between alternating and direct current, each of which had certain deficiencies. It was only with the development of rotary converters that one system (alternating current) was able to develop a dominant position and spur rapid deployment (David, 1992). Next I briefly describe multiple (and complementary) types of complementarity, summarized in Table 2:

- (1) **Hicksian (Production) Complementarity:** Factors of production are Hicksian complements when a decrease in the price of one factor leads to an increase in the quantity used of its complements in production. An innovation that reduces the cost of a factor is equivalent to a decline in the factor's price. This of course presupposes an existing process utilizing the two factors, which makes it somewhat irrelevant for studying innovation. Yet, it generalizes to the insight included in the original PFI framework that the successful commercialization of an innovation will affect demand for complementary goods and services.
- (2) **Edgeworth/Pareto (Consumer) Complementarity:** Two goods, X and Y, are Edgeworth complements in consumption if the utility of consuming them together is greater than that of consuming them in isolation. As a result, the quantity demanded of either good is affected by a change in the quantity demanded of the

⁵ If a piece of technology is independent of others and offers little scope for improvement, we have what Mokyr (2002, p.19) calls a “singleton technique” that might be discovered by chance, with little understanding of the underlying mechanism. While this was common before 1800, it is rare in today's complex knowledge domains.

other. If demand is downward sloping, then this is the consumption-side equivalent of the production-side Hicksian case. In Edgeworth's broader conception, complements can include non-priced items such as government policies, or organizational structures. Milgrom and Roberts (1994) formalized this for a group of activities, labeling them as Edgeworth complements if doing more of any subset increases the returns to doing more of any subset of the remaining activities.

- (3) **Hirshleifer (Asset Price) Complementarity:** This is a financial perspective on the Hicksian insight that can, at least in theory, be used to profit from an innovation. If innovation is likely to move prices as described above, then the initial invention creates foreknowledge about how asset prices might move in the future. An economically rational inventor with financial means and this type of foreknowledge could gain additional value by going long in asset markets that would be positively impacted by the invention and/or short in markets that would be negatively impacted. This model of anticipated complementarity was in fact invoked in the original PFI article by noting that the innovator could speculate, where futures markets were available, on complementary assets that were likely to increase in value (Teece, 1986, p.295). Hirshleifer (1971) illustrates this principle with an analysis of Eli Whitney, who received a patent for the cotton gin in 1794 yet died poor 30 years later after dissipating his profits in litigation over patent infringement. Hirshleifer points out that Whitney could have foreseen the negative effect his invention would have on the price of cotton and the positive effect on the price of land and then speculated accordingly on either or both.
- (4) **Cournot (Input Oligopoly) Complementarity:** Two products are Cournot complements if they are used together but sold by separate companies with monopoly power. Consider, for example, the case where each firm has a monopoly over one of two inputs used to make a product. If the companies are unable to collude, they will charge prices that maximize their individual profits considered in isolation yet fail to maximize their collective profit. This problem arises in PFI when there are two or more bottleneck assets needed to produce an innovative product, and they are owned by separate parties. Whereas common ownership would lead to charging a monopoly price, the separate firms, absent collusion, might charge a higher price and lower their total profit, harming consumers as well. This is an

interesting theoretical puzzle but not likely to be of practical significance because markets are seldom in equilibrium long enough to determine the “right” price. A real-world example would be Intel and Microsoft in the PC industry; their dominance of intermediate goods did not harm consumers because competition in the downstream market drove down end-user prices. The Cournot model has been misapplied in the context of patents where economic theorists have postulated that owners of standard essential patents will pursue self-interested pricing strategies that result in high cumulative royalties (Galetovic and Gupta, 2016) that are destructive of the market. In reality, we tend to observe the opposite, in part because patents are not self-enforcing and in part because licensing as a business model has historically been weak in its ability to capture anything but a tiny fraction of the social benefits from enabling technology. FRAND licensing regimes have also limited the bargaining power of patent holders still further.⁶

- (5) **Teecian (Technological) Complementarity:** Technological complementarity occurs when the value of an innovation depends on altering the nature of one or more existing technologies and/or on creating new ones. This type of complementarity has nothing to do with prices and quantities. It applies when the full benefit (or even any benefit) of the innovation cannot be achieved until some other, complementary technology (which has independent, lower value uses on its own) has been created or re-engineered. This is in fact typical of enabling technologies. For example, realizing the full value from the introduction of electricity required the creation of electric motors that could be attached to machines. The introduction of a new generation of cellular network requires new microchips and handsets to exploit it. Technological complementarity can pose a PFI challenge for an innovator if the complement is created by a separate company and becomes a bottleneck asset. One solution is to create the complement in-house, but the innovator may lack the capabilities to do so.
- (6) **Bresnahan and Trajtenberg (Innovational) Complementarity:** Improvements in a general-purpose technology will increase the productivity in downstream sectors. For instance, the improvement of a cellular network opens new innovation

⁶ FRAND is a legal term that stands for “fair, reasonable, and non-discriminatory.”

opportunities for firms making wireless data devices. This is a special case of (5) above.

Table 2: Types of Complementarity: Summary

| <i>Type</i> | <i>Representative Authors</i> | <i>Description</i> |
|-----------------|--------------------------------|--|
| Production | Hicks (1970) | A decrease in price of X leads to an increase in the quantity of Y |
| Consumption | Edgeworth (1897/1925) | An increase in the quantity demanded of X leads to increased demand for Y |
| Asset Price | Hirshleifer (1971) | Financial arbitrage opportunities are created by foreknowledge of the probable impact of an innovation. |
| Input Oligopoly | Cournot (1838/1960) | Inputs X and Y will be sold for less if the companies can collude to maximize profits. |
| Technological | Teece (1986, 1988b, 2006) | Unlocking the full value of an innovation requires additional innovation in one or more lateral or vertical complements. |
| Innovational | Bresnahan & Trajtenberg (1995) | Improvements in a GPT increases the productivity of goods in downstream applications. |

While complements were highlighted in the original PFI formulation, the analysis was broadbrush. Time was ignored. With the introduction of time, co-evolution becomes important. PFI implicitly recognized “rents” (profits) as flowing to the bottleneck assets, and that could be different complementary assets at different points in time. However, rents are generated not only from exploiting bottleneck assets; they also arise from combining different technologies in unique, value-enhancing ways that lead to gains exceeding the additive nature of standard Edgeworth complements. With Edgeworth complements, the demand for one product, e.g., automobiles can drive demand for another (e.g., gasoline). Even though the combination in a gasoline automobile engine involves complex technical interactions, it does not involve complex economic interactions. Edgeworth complements and Cournot complements are largely irrelevant to the innovation context. Systems integration is one point where information sharing can occur without collusion on price. In reality, of course, complementary technologies can be combined (“orchestrated”) to produce new products that are unique and valuable to users.

2. PFI and complementarity reconsidered

The lines between these various forms of complementarity are in some cases blurred, since two goods can fit several of these categories at the same time. PFI (implicitly) embraced all of them except Cournot. What these types of complementarity have in common is that they all raise complex coordination issues, which in turn render appropriability more difficult. An innovator has no guarantee that the downstream users of the technology (and the providers of complements) will make the

investments needed to generate the largest value. Pessimistic expectations can become self-fulfilling, which reduces the innovator's incentives for investing in future research and development.⁷

In most cases, technological and innovational complementarities impose coordination and control challenges that impair the alignment of activities across firms in a market economy.⁸ Such issues were raised decades ago by Malmgren (1961) and Richardson (1972), were echoed by Williamson (1975), remarked upon by Teece (1984), explored empirically by Armour and Teece (1980) and Helfat and Teece (1987), but then dropped. The economics literature tends to assume that upstream and downstream investment expectations will converge, which seems unlikely given the proprietary (and hence secret) nature of much of the required innovation activity.⁹

Moreover, the imposition of "reasonable" royalty rates in the case of standard essential patents (SEPs), if interpreted by the courts to mean "low" rates, undermines the innovator's incentive to invest in improving the technology further, compounding the problem still further. There is no market mechanism short of vertical (and horizontal) integration that can ensure socially optimal innovation and adoption of enabling (and GPT) technologies that permit further "add in" innovation by downstream implementers. The exploitation of Hirshleifer complementarities can potentially ease the problem by employing financial engineering strategies to help capture spillover benefits. However, these will not be feasible in many settings.

From a managerial perspective, some of these concerns are addressed under the heading of business model issues (Teece, 2010b), leaving it to entrepreneurs to design creative organizational arrangements to help solve the coordination and associated appropriability challenges. Private ordering (contractual) solutions are possible in many cases. However, traditional responses to the coordination and appropriability issues are likely to be less efficacious the greater the generality of the technology and the smaller the entity that owns it. The reasons relate to the sheer scope and complexity of the economic applications of a significant enabling technology. When the application opportunities are numerous and diverse, joint activity places heavy demands on the management resources of the innovator, leading it to fall back on the licensing option. Licensing, however, is often inadequate, especially in instances where courts are reluctant (or unable) to issue injunctions. These issues are explored in more detail in Section E below.

⁷ These points have been given only cursory treatment in the economics literature. Malmgren (1961) and Richardson (1972) had noted the issue in general but the profession has ignored it, believing that expectations with respect to complementary investments somehow converge.

⁸ Technological complementarities are largely absent from economic analysis. In fact, they completely vitiate the concept of a production function, which assumes that a fixed list of inputs is used to practice a technology known to all firms. In reality, production functions, even in the absence of a major innovation, are often firm-specific and quite proprietary. Schumpeter (1934) observed nearly a century ago that the very essence of innovation is typically "new combinations." However, his theory brought no granularity to the analysis. Nor did he consider the appropriability issues around new combinations because his main focus was on the ability of new products and processes to displace existing ones. This spoke to substitution, not complementarities.

⁹ Vertical integration can partially mitigate coordination problems. Armour and Teece (1980) established that R&D levels in the petroleum industry were sensitive to the extent of vertical integration in a direction suggesting that integration can ease the coordination issues when new technology is developed and deployed. Helfat and Teece (1987) showed that vertical integration reduced risk, which can include the uncertainty that accompanies commercialization of new technology.

In the original PFI paper's discussion of business model issues, the greatest emphasis was given to questions around value chain issues, namely, whether the firm ready and able to assemble the specific and generic assets needed to bring the innovation to market. Since then, I and others have expanded the notion of business models to encompass the architecture of the business, including customers, costs, and likely competitor responses (Chesbrough and Rosenbloom, 2002; Teece, 2010b). Moreover, business model design is an iterative process. This point is particularly relevant as more and more businesses are built around software. Software-based business models are relatively easy to modify in response to customer feedback, changes in the user base, or other evidence of missed opportunities. My subsequent work on PFI has emphasized technological as well as value chain complementarity (Somaya and Teece, 2006; Teece, 2006; Somaya, Teece, and Wakeman, 2011). Whereas Hicksian and Edgeworth complementarities present the innovator with challenges and opportunities that can be mitigated or exploited by thinking ahead strategically, technological and innovational complementarities place a large, short-term burden on the innovator to coordinate with all owners of relevant intellectual property. In the original PFI framework, the implied advice to the innovator was to own (or control, or neutralize) any bottleneck asset(s). In a multi-invention context the bottleneck could be a technology rather than a conventional asset. The innovator must either acquire, license, or ally with the owner of relevant technologies in order to ensure a predictable path to follow both for the initial commercialization of the innovation and for its future development (Chesbrough and Teece, 1996). A thorough understanding of profit distribution in the value chain also requires examination of relevant ecosystems, which are inherently composed of various types of complementarity. Ecosystems were absent from PFI in 1986 and only mentioned in passing in PFI-2006.

3. Platforms and ecosystems

Complementarity is the handmaiden of platforms, and platforms power ecosystems. Because of the progress and diffusion of digital technologies, platforms are becoming pervasive. Digital platforms use common standards to integrate products and services—and companies—using the Internet or private networks. Integrated digital platforms concern multiple business functions and enable business ecosystems. While various types of platforms can be recognized (Gawer, 2014; Hazlett, Teece, and Waverman, 2011), there is nonetheless considerable ambiguity about what constitutes a platform—or an ecosystem.

Consider, for instance, the auto industry. It has for decades worked with “platforms”; but these are of a different kind. An auto “platform” is a shared set of common designs, parts, and production efforts arranged around a number of (what appeared externally to be) different models. It perhaps began at General Motors in the 1960s when GM used the same platform (chassis) for the Buick Skylark, Chevy Chevelle, Pontiac Le Mans, and Oldsmobile Cutlass. The platform approach lowers costs by reducing the need to design parts that are functionally identical but different in each model.

In the digital world, platforms can be software-only, like Alphabet's Android operating system (OS), or they can be linked to hardware, like Apple's iPhone and iPad, which are tightly integrated with Apple's proprietary iOS software. Platforms may impact consumers directly, as in the case of these smartphone technologies, or they can be behind the scenes, like the software that a manufacturing company uses to coordinate and monitor its suppliers. Either way, platforms in the ICT world are more attractive when they allow users to tap into a rich array of complements and add-ons, i.e., when they have strong network effects.

Consider further the iOS (Apple) ecosystem. Apple has, to a large extent, “solved” the coordination problem by integrating most of the innovation-intensive elements of its value chain—the microprocessor and the handset hardware. As a consumer-facing company, it does not have to worry about coordinating with downstream partners. The sale of apps in its app store and the sale of music in the iTunes Store can potentially offset some of the weak appropriability for the operating system, which must compete with Google’s free alternative, Android. Apple’s content revenue has been steadily increasing in importance, and it finally began breaking out “Services” revenue (mostly app sales) in January 2016. In the first quarter of the calendar year, Services were \$5.99 billion, nearly 12% of total revenue, making it the second largest segment after the (much larger) iPhone and slightly larger than revenue from Mac sales. Alphabet’s rival offering, Google Play, the content portal for the Android ecosystem, has also thrived.

There are two basic types of digital platform, with numerous hybrid combinations (Evans and Gawer, 2016). A *transaction platform* facilitates exchanges by otherwise fragmented groups of consumers and/or firms. The paradigm here is eBay, which allows huge numbers of individual sellers and buyers located anywhere in the world to find one another with an ease that was previously unimaginable. While digitization has enabled transaction platforms in a growing range of industries, this transactional type of platform is not entirely new. For example, the credit card industry has long provided a viable payment option that merchants will accept, that banks will join by issuing cards and processing transactions, and that cardholders find of value. An *innovation platform* provides a base technology and distribution system to which other companies can add their own innovations, increasing the value for the system as a whole. Apple’s “app” ecosystem is the paradigmatic example of this. Innovation platforms fit perfectly within the original PFI framework, which emphasized the need to access complementors. While the relationship is less obvious in the case of transaction platforms, successful examples such as Amazon need to attract transactors, e.g., Amazon Marketplace vendors, in much the same way they must attract key partners, e.g., the delivery services, by providing a sufficiently attractive platform whose participants will reach a critical mass of buyers and sellers.

In essence, a digital platform provides a hub around which companies and users can, jointly or separately, innovate and attract users far more productively than if they were to try to achieve the same goals in the absence of the platform. Owning or controlling a successful platform upon which other firms erect their business model can provide a commanding position from which to enhance an ecosystem and capture value from it. The dynamic capabilities framework (Teece, 2007, 2014) helps explain why some firms successfully create platform-powered ecosystems that combine multiple business models. Firms like Amazon, Apple, Facebook, Google, and Microsoft have managed to sense opportunities outside of their original business, seize them by mobilizing relevant resources and, most importantly, transforming their organizations by adding platform capabilities, particularly the management of complements in the ecosystem. Other firms that had many of the assets to succeed (e.g., telecoms operators, discussed below) may have lacked the strong dynamic capabilities that would have been required for them to create a competitive alternative.¹⁰

It is common to assume that platforms are owned by a single firm, but this often is not the case. The company or companies at the center must provide coordinating mechanisms, rules, intellectual

¹⁰ I thank Benoit Reillier (email correspondence, August 3, 2016) for this insightful observation.

property, and financial capital to create structure and momentum for the market or ecosystem it seeks to create. Standards help define platforms, and these standards need not be proprietary. For example, the Internet itself is a platform and Internet protocols are not proprietary.

In the ICT industry, many underlying software technologies are stacked one upon another, with devices such as smartphones on top of the stack. Although some firms try to control an entire vertical stack, they generally do not succeed. Because of this, vertical cooperation (and horizontal competition) is necessary and common.

Platform leaders take responsibility for guiding the ongoing technological evolution of the system (Gawer and Cusumano, 2002). Creating and capturing value therefore requires a mix of openness (to attract complementors) and a degree of control (to create a good user experience). Once created, there may be opportunities to capture value through device sales and other mechanisms. When there is platform-to-platform competition, adoption and commercial success is likely a function of who can recruit the most (and the best) complementors. Over time, the advantage belongs to the platform leaders that set the rules in the manner most likely to benefit the system as a whole and not just their own short-term interests.

III. C. Multi-level PFI¹¹

In platform-based ecosystems, competition occurs at three levels: (1) between one platform and another, as was the case for VHS versus Betamax in the fight for the VCR market, or the current case of Apple's iOS versus Google's Android in the mobile sector; (2) between a platform and its partners, as occurred with Microsoft capturing a part of the value from browsers, streaming media, and instant messaging applications that worked on its Windows operating system; and (3) among complementors, each seeking a position within a platform-based ecosystem, as in the case of two mobile games each chasing the same consumer segment. With regard to capturing value, "[s]trategy becomes vastly more complex as firms consider dynamic interactions of a multi-layered ecosystem" (Parker and Van Alstyne, 2015: 5).

Notwithstanding this complexity, one can distill a few generalizations from the literature and from recent history. First, competition between platforms tends to produce winner-take-all outcomes when there are large demand- or supply-side scale economies, multi-homing costs, and no benefit from niche specialization. Platform-to-platform competition also leads to openness that results from each platform endeavoring to recruit more developers; the greater the openness, the less the opportunity for the provider to capture value directly. A good value capture strategy is to court well-known brands/partners who can bring large blocks of customers to the platform (Cusumano and Gawer, 2002).

Expanding the perspective to multiple levels, managing partner-to-platform competition and partner-to-partner competition requires striking a balance between cooperation and competition. Throughout any analysis of multi-level PFI, the fundamental question is the following: is there likely to be a competitive bottleneck? If so, where; and when will it become apparent if it is not currently so? The astute identification of a future bottleneck is a potential opportunity to build or buy the necessary resources to benefit (or not suffer) from it. This takes strong dynamic capabilities, including organizational skills in sensing and seizing (Teece, 2007). In the case of two-sided and multi-sided platforms, the platform

¹¹ This section is based in part on Hazlett, Teece, and Waverman (2011).

owner will try to capture value by internalizing the (usually positive) externalities. Generally, it is desirable to work things out through what Oliver Williamson (1985) calls “private ordering” arrangements in order to reduce the likelihood of a state-mandated arrangement, which is likely to impose an arbitrary solution that limits outcomes and destroys value for consumers, too.

In order to explore some of the PFI considerations within and across the levels of an industry, I will use examples from digital cellular telephony, an industry that did not exist in 1986, the year of the first PFI article. The massive transformations that it has enabled have rapidly taken hold during the first decades of the new millennium. Mobile telephony provides useful illustrations of the additional granularity provided by the broader aperture of the expanded PFI model.

The original Teece (1986) model applied well enough to feature phones. Prior to the introduction of the iPhone, the cellular handset market was dominated by a handful of firms with prodigious technological and marketing (complementary) assets. The Top-5 lists of the early 2000s were consistently dominated by Nokia, Motorola, Samsung, Sony-Ericsson, and LG Electronics, which collectively accounted for 70% to 80% of global market share.¹² Their collective dominance appeared unshakable; Nokia alone continuously accounted for a third or more of worldwide mobile phone sales as late as 2010.

Apple’s iPhone provided a more PC-like experience than the proto-smartphones then on the market, and it was supported by a growing market for content through the iTunes Store that Apple had established for its iPod. Within a year of the iPhone’s launch, Apple introduced the App Store to make (Apple-vetted) third-party software available.

In addition to its strong positioning with regard to complements, the iPhone benefitted from Apple’s (path-dependent) learning with the iPod, its positive consumer image, and its ability to integrate hardware, software, and services.¹³ Success was rapid and durable. By the end of 2008, the iPhone dominated the market for high-end phones and set the dominant design for rivals to adopt. The introduction of smartphones and their associated platforms changed not only the hardware but also the business model. In 2011, Steve Elop at Nokia told his employees “Our competitors aren’t taking our market share with devices; they are taking our market share with an entire ecosystem.”¹⁴

To account for this change, PFI must increasingly be applied across multiple levels (upstream and downstream) of an industry, and cellular telephony is a case in point. Much of the enabling technology gets built onto (communications) chips, which are provided by a handful of firms competing on cost and quality. The chips must go into infrastructure equipment and terminal devices (primarily handsets). The quality of the user experience is dependent on the skill with which the technology is integrated into a system by mobile carriers and handset makers because the performance that one can achieve out of cell networks is intimately tied to network design and handset antenna design. Each of these levels can be viewed as an ecosystem, they can be grouped to incorporate interactions. Indeed, the Internet itself can be thought of as an ecosystem. Hazlett, Teece, and Waverman (2011) show that, at least in the early

¹² Compiled from press reports of the time on annual market share.

¹³ As Tim Cook, a long-time executive at Apple and its current CEO, said in 2013: “Apple has the ability to innovate in [hardware, software and services] and create magic... This isn’t something you can just write a check for. This is something you build over decades” (AFP, 2013).

¹⁴ The leaked Nokia memo was widely reproduced online. See <http://www.engadget.com/2011/02/08/nokia-ceo-stephen-elop-rallies-troops-in-brutally-honest-burnin/> (accessed June 7, 2016).

years of the smartphone era in the U.S. market, the network providers were the losers and the handset providers the winners (see Table 3). This is consistent with Dedrick, Kraemer, and Linden (2011). Profits are made and lost at many layers in the system. PFI predicts that the bottleneck assets, if any, will earn the lion's share of the rents. In the PC ecosystem, Microsoft's ownership and control of the Windows operating system gave it (along with Intel as the main microprocessor provider) a privileged position with respect to capturing PC ecosystem profits. Microsoft was aided in its value capture mission by the beneficial linkages it created between the Windows operating system and personal productivity software such as Excel and Outlook.

Table 3. Q Ratios for Selected Wireless Ecosystem Players

| <i>Company</i> | <i>2008</i> | <i>1Q2009</i> | <i>2Q2009</i> | <i>Ratio to SP500 (2Q2009)</i> | <i>Enterprise Value (\$bill. as of 5.4.10)</i> |
|----------------|-------------|---------------|---------------|------------------------------------|--|
| Sprint | .6 | .6 | .5 | 0.68 | 28.9 |
| Apple | 2.6 | 2.5 | 3.1 | 4.25 | 212.2 |
| RIM | 5.0 | 4.6 | 6.2 | 8.49 | ~38.5 |
| Nokia | 1.0 | .9 | 1.2 | 1.64 | ~41.5 |
| Motorola | .7 | .5 | .3 | 0.41 | 11.7 |
| Qualcomm | 2.5 | 2.7 | 3.2 | 4.38 | 50.7 |
| S&P 500 | .55 | .61 | .73 | 1.0 | |

Source: Manual of Ideas (Sept. 21, 2009).

Network providers such as Verizon, AT&T, and Vodafone face multiple challenges. First, consumers see differentiation at the device level, not so much at the network level (although service quality varies by location). Second, these carriers must make massive investment in network infrastructure with each new generation of technology yet compete fiercely on price. In the United States market, with multiple nationwide networks in operation, the network is not a bottleneck asset. For several years after the introduction of the iPhone, AT&T profited from its exclusive arrangement with Apple, but that ended in 2011 when Apple introduced a phone that worked on the network of AT&T's chief rival, Verizon. The value to Apple of accessing a large pool of potential new customers outweighed the benefits it was receiving through the AT&T partnership. The end of the partnership occurred at a time of Apple's choosing. Third, Google has invested in providing a certain amount of optical fiber for transport, further reducing the bottleneck potential of the operators, who also suffered a blow from the rise of Internet telephony.

As this suggests, the true bottlenecks, i.e., the assets that are in high demand but not competitively supplied, are the mobile operating systems of Apple and Google. In addition, Apple's iPhone has product/feature patents that have proven formidable. Although the two OS-based ecosystems compete, they are not viewed by users as direct substitutes, in part because Apple's hardware and brand image are exclusive to the iOS ecosystem but also because users face switching costs when moving between the two platforms. For several years, the lion's share of profits in the cellular industry has gone to these two firms, which are currently (June 2016) the two most valuable publicly listed companies in the U.S. While the operating system bottleneck in smartphones is similar to Microsoft's position in PCs, there is no bottleneck for cell phone processors equivalent to that of Intel. This has to do with the success of the open ecosystem created by the UK's ARM Ltd., in which many licensees, competing fiercely against each

other, are able to access the key intellectual property to design the chips that run the phone's software. Nearly all smartphones today use processor chips built around one of ARM's designs.¹⁵ ARM receives a royalty from each processor, which is typically the second most expensive component after the display. ARM's extensive portfolio of intellectual property around power-efficient hardware and software could be considered a bottleneck because its licensees, such as Samsung and Qualcomm, would need a great deal of time and expense to replace it. However, ARM has kept its royalty rates relatively low, recognizing the price pressure faced by its licensees.

In the first years after the iPhone appeared, other handset suppliers like Nokia and RIM/BlackBerry still earned profits. This changed as the business models of Apple and Google took hold. Their business models are quite different and reflect the capabilities that each firm brought to the cellular telephony field, in which neither one had previously competed.

Google's business model in the online world has been based since near its beginning on ad-supported technology development. It has no technical or market lock-in, just the attractiveness of easy access to large quantities of well-organized data and content. If better search, mapping and email systems emerged, it is not hard for users to migrate. Yahoo, an early leader in the same space, suffered from this lack of any significant hold on its users, and was left behind by Google's superior ability to steadily improve and expand its offering. For cellular handsets, Google structured the Android ecosystem so that Google captures the rents from mobile ads to fund its continued value creation through improvement of its Android OS, while the makers of Android devices dissipate their profits by competing against each other. Google focuses on software and leaves most hardware development to partner firms.

Apple, on the other hand, had a long history developing hardware as well as software prior to entering the cell phone industry. Apple's profits from the iOS ecosystem flow mostly from device sales, which is possible because Apple is their exclusive manufacturer. Apple is still largely the integrated firm it has always been, in which hardware and software integration is a source of advantage.

One important commonality (pioneered by Apple) is that both Apple and Google offer a platform (a digital storefront for applications) for independent software developers who add value (which they must share with Apple and Google) to the iOS and Android ecosystems. In the first instance, this was primarily about the provision of complementary assets to enhance the network effects that attract and bind users to each system. As mentioned earlier, profits from third-party sales of content are only recently starting to make a significant contribution to the firm's bottom line. Both companies have also expanded to the point that they benefit from the traditional economies of scale and scope.

IV. D. Modularity and Industry Performance

While PFI is primarily a tool for thinking about profitability within individual firms, or in the ecosystems in which they participate, the logic also applies to the industry as a whole. Just as an innovator must design a profitable business model to ensure the ability to improve the technology going forward, an industry, although lacking a central coordination mechanism, must do the same. The future progress may be in danger due to changes in the way applied research in these enabling technologies is funded.

¹⁵ Intel has tried to extend its reach into the cell phone, at one point even designing with an ARM license it had acquired, but its capabilities in low-power computing, where ARM excels, were limited due to its historical focus on high performance. ARM is starting to face competition from several open-source hardware initiatives, but these are more likely to be important, if at all, in less-established markets, such as the so-called Internet of Things.

The sources of invention in the electronics and communications industry was once highly consolidated. AT&T's Bell Labs in particular was the wellspring of much of the new technology for the telecom and electronics industry. It had long pursued economically significant research, as represented by the invention of the transistor—the basis for the entire computer industry—in 1947. NTT has fulfilled a somewhat related role, providing research, equipment manufacture, and telecom services in Japan. Sweden's Ericsson, a private firm, was also integrated from research through to equipment design and manufacture.

Unfortunately, Bell Labs has virtually ceased to exist in all but name. In place of “blue-sky” research into technologies with no immediate use that later gave rise to innovations of great value to the telecommunications industry, research today is small scale and very product development focused. The decline began with the 1982 consent decree pursuant to which AT&T was broken up in 1984 and Bell Labs made vulnerable. Bell Labs has since passed through a series of successor companies and is currently part of Nokia. The loss of Bell Labs' massive contributions to global science in general—and to communications technology in particular—is collateral damage from the breakup of AT&T. Downstream sectors have experienced massive changes as well. In earlier generations of the technology, companies like Motorola and Ericsson made components, infrastructure equipment, and consumer devices, following a model that had been established long before by the state-run monopolies in charge of fixed-line telephony. The breakdown of this industry structure into vertically specialized groups of firms resembles the trajectory of the computer industry, where vertically integrated companies like DEC that made midrange and minicomputers gave way to the specialist companies that populate the modular PC industry.

The modular business model, when present at the industry level, squeezes profit out of each module and “ownership of complementary assets” no longer works as a core appropriability mechanism. This was the case for personal computers, where profit for the modules (hard drives, mother boards, final assembly, etc.) became vanishingly small. Yet the microprocessor and operating system proved resistant to imitation and competition for a variety of reasons, including technological prowess, intellectual property, branding and, especially in Microsoft's case, network effects.

The modularization of the mobile telephony production system limits the further progress of the industry's enabling technologies. Companies such as Ericsson are finding themselves forced to de-verticalize and specialize in one or two modules to survive. However, modules are inherently competitive, making it hard to earn sufficient profits to fund R&D that is likely to benefit the industry as a whole. This effect is exacerbated if the providers of enabling technology (including owners of standards essential patents) are unable to reap a reasonable return for their technological contributors to standards development.

In digital electronic systems, architectural innovation is often desirable and especially difficult. Modularization, meanwhile, enables autonomous innovation to continue at a rapid pace, but it eventually has limited impact unless systemic or architecture innovation (involving new standards) happens too because, in many complex systems, it is innovation in the system (and/or in the underlying enabling technology) that affords the greatest benefits. It is also system innovation that enable changes in business models.

V. E. Business Model Configuration: Licensing and Alternatives

As already noted, the original PFI framework explored only a limited repertoire of business models; the main choices recognized were (a) to license and (b) to invest in one or more complementary assets and/or technologies. Joint venture arrangements were also analyzed, as some reward was considered better than none. The framework implicitly relied on ideas from bargaining theory, which generally showed how rents would be shared.

It is common for economists to assume that technology licensing is a straightforward way for innovators to capture value. Indeed, it is not uncommon for them to assume that a patent conveys an automatic monopoly, which is far from the truth. Not only is it rare for a patent, or even a patent portfolio, to coincide with a relevant (in antitrust terms) market, it is also unlikely that a court will automatically grant an injunction upon the mere request of a patent owner. This was not the case in the past, and, since the 2006 Supreme Court decision in *eBay Inc. v. MercExchange, L.L.C.*, it is at most a remote possibility. These observations help explain the conundrum that Nobel laureate Ken Arrow outlined half a century ago, to which scholars have yet to respond at all, let alone comprehensively:

Patent royalties are generally so low that the profits from exploiting one's own invention are not appreciably greater than those derived from the use of others' knowledge. It really calls for some explanation, why the firm that has developed the knowledge cannot demand a greater share of the resulting profits" (Arrow, 1962, p 355)

The explanation that Arrow was looking for goes beyond the fact the courts have often sided with implementers over inventors. It relates, as discussed earlier, to the inherent inability of firms that invent enabling technologies and GPTs to attract and harness the financial resources needed to invest in the complementary assets required to implement appropriability-enhancing business models. As noted, timing and standards matter, along with what Helpman (1998, p.4) called "institutional conditions." Strategy should also be added to the list, along with whatever talent the inventor can muster to simultaneously invent, be entrepreneurial, innovate business model architectures, and orchestrate complementary technologies.

As mentioned earlier, the identification and control of key complements both inside and outside of ecosystems play a central role in the PFI story and helps provide the explanation that Arrow was searching for. If the complements are not owned by, or cannot easily be replicated by, the inventor, then rents will likely be drained away toward the complement owner(s), and the innovator will be in no better position than an imitator who controls relevant complements.

PFI could easily have collapsed into a discourse on bargaining theory to explain likely rent division, with appeals to game-theoretic models developed by Bertrand, Nash, Rubinstein and others. Rather than doing so I have chosen to focus on framing the broader set of issues rather than narrowing the focus to the division of rents between complementor and innovator. This focus is chosen because the context is not static (as game-theoretic models assume), and the innovator sometimes has the time to build and/or acquire relevant capabilities.

However, there are a plethora of other issues to consider besides complements. For instance, strategy analysis is an essential step in designing a competitively substantive business model unless the business

mode design passes through the filters of strategy analysis, it is unlikely to be viable. Getting the business model design right is critical to capturing value. However, we must recognize that it is sometimes impossible to design a business model to capture value even when value is being generated. A textbook example is the provision of lighthouse services. It was not practical to limit which ships would and would not benefit; if some paid, the others could free ride. Society's solution everywhere is to have a government agency (in the U.S., the Coast Guard) provide the services and try to recoup the cost through taxation or imposition of mandatory fees. Such services/goods are referred to as public goods in recognition of the difficulty (or near impossibility) of constructing a viable business model.

There are of course exceptions to the requirement to innovate in business models. For example, small improvements in a manufacturing process (even if cumulatively large) will usually not require business model innovation. But the more enabling the innovation, and the more challenging the revenue architecture, the greater the changes likely to be needed in standard business models (Teece 2010, p. 186). The life and fortunes of innovations are compromised on account of this, particularly when inventors lack the wherewithal to access the necessary resources. Moreover, designing a new business model itself requires creativity, insight, and deep customer knowledge. Technological DNA needs to be married to business/entrepreneurial DNA for inventors to succeed.¹⁶ This is a tall order for most inventors.

Even Thomas Edison with his portfolio of 1,000+ patents failed commercially on many fronts, despite the fundamental nature of many of his technological contributions. He abandoned the recording business after arguably failing to get its business model right (by offering a closed platform in which only Edison discs could work on Edison phonographs). His lackluster commercial results with the phonograph were also due to deficiencies in the product (poor sound reproduction, recordings that could only survive a few plays). Others moved in to fill the void, including Alexander Graham Bell, by using wax in the place of tin foil (for the recording cylinder) and a floating stylus instead of a rigid needle that would incise. In addition to being a great inventor, Edison was also very good at judging the needs of the market. This meant that he and the other inventors he worked with were able to focus on inventions that were truly useful and had a chance to be commercially viable. He created electricity as a system and pioneered business methods such as critical path techniques, engineering storyboards, and interdisciplinary product development, central storerooms, and corporate libraries.

While Edison did not die poor (apparently he was worth \$12 million in 1932), his net worth was paltry in relation to the value he created for society. Remsen Crawford, Edison's biographer and close friend, has endeavored to explain Edison's own interpretation of his modest financial success (Crawford, 1932). Edison by no means relied on patent licensing as his business model. He founded commercial manufacturing ventures, including several that merged to form the company now known as General Electric. He was well aware that the patent system was often quite ineffectual and once related how he had fought for 14 years for his patent on the incandescent bulb against infringers, despairing (consistent with the Arrow observation cited earlier) that: "my opponents were able to keep me out of profits of that particular patent until the rights to it were well-nigh useless" (cited in Crawford, 1932, p. 56). He

¹⁶ This paraphrases a conversation with the author and Andy Bryant, Chairman of Intel, 2015.

had similar problems monetizing his patents on devices for capturing and viewing moving pictures as well.¹⁷

Edison's experience is not uncommon today. The appropriability problem for innovators is almost always severe. As a consequence, technological innovation needs to be matched with business model innovation to help the inventor capture more value. Naked licensing will rarely suffice. This is all the more true if the courts favor infringers, which they seem more and more inclined to do. The PFI framework helps explain why intellectual property law and the courts should be more inventor-friendly and intellectual property rights strengthened. Holdouts (patent infringers refusing to take a license even after the creation of a licensing program suggests that patent validity is likely high) need to be curbed. While it has long been recognized that basic research is, by its nature, a "public good" in the economic sense, it is perhaps necessary to recognize that other types of new technology have a substantial public good component too. As a practical matter, GPTs and enabling technologies have appropriability characteristics not unlike basic research for somewhat different reasons. The lighthouse example (of a public good) is much more general (and also applicable to innovation) than previous scholars have perhaps recognized. It is not just that free riders (e.g., patent holdouts) are common; the phenomenon is amplified by the fact that many inventions, especially of GPTs and enabling technologies, have such broad applicability that direct investment business models designed to exhaust the full range of possibilities are too expensive for markets to finance. Furthermore, they often take too long. The result is that implementers in different industries often achieve windfalls from the inventor's travails. Underinvestment in inventive and innovative activity is a likely consequence. While it is hard to prove that innovation would be more plentiful under a stronger appropriability regime, weak innovation could be one contributor to the declining rate of U.S. productivity growth over the past decade.

VI. F. Standards

As shown in Table 1, technology licensing accounted for a modest 1.2 percent share of total wireless revenue. There are roughly a dozen major holders of "essential" patents that any implementation of the technology is likely to require. Despite theoretical fantasies about patent hold-up (i.e., patent holders changing terms to extract more value from licensees), it is never observed (Galetovic, Haber, and Levine, 2015). Patents are not self-enforcing, so that appropriability is far from perfect and patent hold-out (i.e., infringers failing to pay royalties) is a relatively frequent occurrence.

1. Standard setting and standards development

Technological standards provide the foundation on top of which rival firms build their product and service offerings. The timely creation and development of standards is important for the ability of non-integrated firms to offer steadily improving hardware and software.

Standards appeared in the original PFI framework primarily in the form of dominant designs. Dominant designs (e.g., Henry Ford's Model T, Apple's iPhone) emerge via market competition and become *de facto* standards, unlike the wireless telecom and other types of complex standards that are

¹⁷ More interestingly, Edison saw the usurpation of profits by pirates as "particularly apt to result in the case of some extraordinary patents. I could invent a new monkey wrench which might go without infringement, but the moment I take certain forces and work out a moving picture for the first time in the history of the world, like that produced by the Kinetoscope, mark you how the pirates rose up and call it their own" (Crawford, 1932, p.57). This is consistent with the perspective advanced earlier.

cooperatively developed. Standard setting and standards development are worthy of further consideration.

De facto standards will continue to play a role. As mobile telephony moves ahead to 5G, the field of competition will widen from smartphones to the so-called Internet of Things (IoT). The “things” will be vast networks of sensor-enabled smart devices that will provide data and control to consumers and companies, enabling a vast range of new applications, including smart homes, smart cities, dynamic utility pricing, mobile health, and autonomous vehicles. While the underlying 5G communication standards are being developed through a formal collaborative mechanism, the standards for the communication protocols among devices will likely be settled in the market. Competing consortia, including the Allseen Alliance and the Open Connectivity Foundation, have published open source code for device interoperability at the application layer, and Alphabet’s Nest Labs decided to follow suit with its equivalent software in order not to be stuck in a niche (Merritt, 2016).

Standards for complex enabling technologies such as 5G communications require collaborative development if they are to provide full interoperability across ecosystems and industries. Many firms have a stake in contributing to the performance of the technology, which they will use as the basis for launching their own products in the market.

Among other benefits, standards enable modularization, which allows firms to specialize within the ecosystem, developing complements to the platform that are certain to work through the standard interface (see Langlois and Robertson, 1992). Modularity also facilitates competition and entry. For instance, Wi-Fi standards have led to the development of a competitive market for small-network wireless gear. Of course standard setting was important in the old industrial world, but the challenge (and the opportunity) from standards in old-line industry is relatively minor compared to the digital world.

The process by which a technology is selected for inclusion in a standard typically involves inquiry and negotiation within a standard setting organization (SSO) such as SAE International, a U.S.-based global SSO for the transport industries. Many standards promulgated by SSOs, such as the shape of an electrical outlet or the side of the road on which to drive, contain little in the way of new technology.

In other cases, standards assemble and anoint new technologies with strong implications for downstream innovation. In such cases, the complex technical details of a standard are likely to be hammered out in a standard development organization (SDO), a forum that allows participating firms to contribute technology and to shape the standard. While standards development involves quite a lot of technology exchange, actual R&D is done not by the SDO but rather back at the participating companies. Once the standard is set, SDO participants are free (within certain limits discussed below) to pursue their own business models around their patents. Much of the enabling technology gets built onto (communications) chips, which are provided by a handful of firms competing on cost and quality. The chips must go into infrastructure equipment and terminal devices (primarily phones). The microprocessor chips that enable the software applications on smartphones may be separate or integrated with the semiconductor intellectual property for the standards-based communication protocols.

While some scholars may treat standard setting and standards development interchangeably, the preceding descriptions make clear that they are very different. Table 4 compares the SSO and SDO models.

Table 4: Standard Setting versus Standard Development

| | <i>Standard Setting (SSO) (Model One)</i> | <i>Standard Development (SDO) (Model Two)</i> |
|-----------------|--|---|
| Process | Selection amongst known alternatives offered by contributors; choices serendipitous... no clear winner | New technologies developed, often at great expense to contributors. Standard adopted because it's of superior performance |
| Outcomes | Uniformity, compatibility | Innovation, uniformity, compatibility |
| Pricing | Usually zero (patents & trade secrets only rarely implicated) | FRAND (fair, reasonable and non-discriminatory) |
| Examples | Left- v Right-hand drive autos, SAE component; British v American electrical outlets | 3G, 4G, LTE; 802.11 Wi-Fi (IEEE, ETSI) |

A subset of patents relevant to the standard developed in an SDO will be deemed “essential” for any implementation of the standard. A formal declaration of essentiality is a special type of complementary asset for an innovation that helps ensure that it will earn a royalty stream over the lifetime of the standard. A “standard essential” designation need not imply the harnessing of an enabling technology, but sometimes it does.

2. Standard essential patents (SEPs) and royalties

It is important to recognize that selecting standards is not always a matter of picking amongst equally good alternatives. In some cases, it involves private parties investing hundreds of millions of dollars in R&D in order to put forth technology that they hope will be adopted industry-wide. In addition to the potential for a royalty stream from its innovations, a participating firm may also hope to leverage complementary intellectual property that is separate from the standard and to leverage the advanced familiarity of its engineers to stay ahead of its rivals.

Once the candidate technologies have been submitted, experts at the SDO collaborate to evaluate competing possibilities with an eye toward producing the highest-performance result possible. Most SDOs consist of numerous working groups each striving to find the ways to achieve the goals that they have been set. Nevertheless, what’s animating public policy in the U.S. is the Department of Justice’s concern “that sometimes technology acquires value only because it becomes embedded in a standard,”

i.e., that the contents of a standard might be arbitrary rather than objectively superior to existing alternatives.¹⁸

Applying this “arbitrary standard” assumption when the “enabling innovation” paradigm is at work harms the voluntary research cooperation that has emerged in recent years. The likely consequences of policies/rules that assume the world is Model One (SSO) when it’s Model Two (SDO) are deleterious. Standards development organizations have made architectural innovation possible so that industry and consumers get the benefit of both modularization and architectural innovation. This is a noteworthy development.

Generally, standard setting bodies recognize that (patented) contributions to a standard ought to receive recognition by being eligible for royalties to be paid by those that implement the standard. The main reason of course is to create incentives for contributions to standards setting bodies and for future improvements to the underlying technologies.

Yet it is also important that patent holders do not over-claim, just as it is important that users take a license. Most SDOs have policies that advance and support the idea of fair, reasonable, and non-discriminatory (FRAND) licensing terms, particularly for standard essential patents (SEPs). The technology owners and implementers must still negotiate to determine fees and/or royalties for each licensor based on projected volumes, creditworthiness, etc. Although the definition of “reasonable” does not have an engineer’s precision attached to it—with disagreements occasionally leading to litigation—the FRAND system has worked very well for over half a century.¹⁹

IV. Management and Policy Conclusions

Teece (1986) endeavored to explain how the profits flowing from preselected, commercially viable innovations would likely be divided, based on structural, behavioral, and strategic factors. The core insight of the PFI model is that the distribution of the profits (flowing from an innovation) among the innovators, rivals, complementors, suppliers, and consumers depends primarily on the features of the innovation relative to possible substitutes, the ease of imitation, the control of standards, the degree to which essential complementary assets are controlled by others, and the value-capture business models effectively available to the pioneer and/or followers. The value-maximizing strategy and business model ought to take into account the appropriability regime (the nature of the innovation; the strength of the corresponding intellectual property rights), complementary assets (especially any bottleneck assets that are, or could be, in limited supply), and industry evolution (including whether standards have been established and the presence or absence of network effects).

While the Profiting From Innovation framework is loosely generalizable to other contexts, the implicit “one technology, one patent, one product” assumption in Teece (1986) has become less and less realistic. Thirty years ago, R&D excellence and commercialization capability resided in maybe 50-100 private corporations worldwide (e.g., IBM, TI, GE, Siemens, Roche, Exxon, Shell). Today we have 100 plus strong technology firms in almost every field and hundreds of new enterprises. Public institution and universities in the US, Europe and Asia play big roles too. This dramatic shift to a rich plethora of

¹⁸ David Meyer, Deputy Assistant Attorney General, DOJ, March 26th 2008.

¹⁹ The IEEE, a standards organization for the electronics industry, has made policy errors in recent years such as adopting bylaws in 2015 that potentially allow the discussion of licensing terms within working groups that consist largely of future licensees, essentially opening the door to buyer cartels amongst implementers.

partners—some necessary many optional—has changed the management landscape, requiring specific embellishment to the PFI framework. In particular, complementary assets and complementary technologies are more significant than ever in a world of competing and intersecting platforms, likely dwarfing the installed base effects that drove profits in the PC era. While scale, scope, and installed base/network effects still matter, complements are center stage in the wireless-enabled world of digital convergence. There are 6 groups of embellishments added into the framework:

- 1) Because of the importance of complementary assets, additional granularity has been added to complements. It is no longer possible to have meaningful discussions about complementary assets without a clear understanding of the nature of the complementarity at issue.
- 2) Business ecosystems are increasingly the relevant competitive unit. When platform-powered ecosystems are properly matched to business models, they can help create and capture value.
- 3) Multi-invention contexts, in which individual products draw on multiple internal and external sources of technology (patented and unpatented), are pervasive, and now assumed to be the norm in PFI.
- 4) Business model choices to aid in value creation and capture are much more complex than the simplified “licensing versus in-house production” business model of PFI ‘86. Building platforms and strategically matching business models to them requires strong dynamic capabilities.
- 5) By implication, PFI has much to do with dynamic capabilities (Teece, 2014), because of the need to orchestrate complementary assets, design business models and match strategy and capabilities.
- 6) It is very hard for innovating enterprises to design a business model or create an investment strategy to capture more than a modest amount of the upside associated with enabling and general purpose technologies. Such technologies/innovations are difficult for individual inventors to capture value from

unless the intellectual property system is working very well. Public policy should not only favor innovation, it should especially favor those companies and individuals that create enabling technologies (and GPTs, once they are evident). The economic system needs to support them in terms of both value capture and value creation.

With these embellishments, the framework responds to the quandaries advanced by the two Nobel Laureates mentioned earlier: (1) Samuelson's call for a fresh look at complementarities (at least in the context of innovation), and (2) Arrow's concern that imitation is often as profitable as innovation. The success of the original PFI model stemmed from its rather narrow frame—wider than most exercises in neoclassical economics, but narrower than an ecosystem or industry-level perspective. While the narrower frame kept the problem tractable, it limited the model's applicability in more complex settings. In this paper, I have opened the aperture a notch.

Ways that the aperture has been widened include analyses of enabling technology, multi-level issues, and the differing types of complementarity. Already in Teece (1986), complementarity was recognized as vertical, lateral, and horizontal, and either technological or value chain-based, far beyond what's envisaged by standard economic analysis. For example, the bringing together of electronics and mechanical/servo technologies led to the creation of robotics. Combinations like these go well beyond economies of scope or other economic notions of improved efficiency.

Another potentially useful form of complementarity is that of Hirshleifer complements. The ability to recognize how an innovation will affect the value of complementary goods can be quite important for capturing value and can be exploited at the same time as traditional Edgeworth complementarity. For instance, Apple's decision to acquire processor design capabilities in 2008 (a year after the release of the first iPhone, which used a Samsung processor) was not only a reflection of the potential benefits it could draw from tight hardware integration but also a recognition that reliance on an outside chip supplier such as Samsung or Intel for the electronic brain in future generations of a hit product would disperse a lot of the value that Apple had generated.

Complementarity lies at the core of ecosystems. Today, most ecosystems are structured around one or more platforms. Platforms allow modularization. Modularization in turn allows contributors of autonomous innovation from component and subassembly (i.e., module) suppliers without requiring them to coordinate with the rest of the value chain beyond meeting the standards for their modular interface, which can potentially accelerate innovation. On the other hand, modularity can undermine the ability of the ecosystem to generate systemic (or architectural) innovation (Teece, 1984; Henderson and Clark, 1990). This may account why Apple, a vertically integrated product development company, has excelled at creating breakthrough product innovations, while the fully modular ecosystem that has grown up around Google's Android OS has excelled at imitation and incremental innovation.

The mobile phone case also shows that profits can be earned in different ways. As discussed above, the platform leaders Apple and Google have found completely different business models for extracting value from the ecosystem. In Apple's ecosystem in particular, complementors don't get a free ride; they must share the revenue they earn from iOS apps with Apple.

Long-term ecosystem prosperity requires continued development of enabling technologies. This in turn requires that developers of the technology be adequately compensated. In the past, when the developers were vertically integrated and/or diversified, they were able to design business models that kept the technology affordable while they profited in ways other than licensing, although this was still a small fraction of the total social returns. Now, when most developers are narrowly specialized, the incentive design challenge is even greater. Profiting from enabling technologies is difficult outside of a full vertical integration model because it implicitly involves a plethora of transactions and licensing or joint venture activities that are inherently difficult to structure, establish, and manage.

After some recent U.S. court decisions, I question whether the FRAND licensing rules for standard essential patents provide sufficient compensation for contributions to standards development. The licensing regime has worked well in the past for mobile telephony, balancing the interests of technology contributors and technology implementers. However, the IEEE patent rule changes adopted in 2015 are a step in the wrong direction and will likely cripple innovation, the IEEE, or both (Teece and Sherry, 2016).

The SDOs in which important enabling technologies are co-developed are a relatively new phenomenon and require understanding. Licensing regimes with respect to SEPs are somewhat fragile. Large-scale cooperation among multiple participants with quite different interests is always hard to achieve. This is particularly true when there are short-term differences in incentives and the players are heterogeneous. Policy makers and the courts need to be keenly aware of the appropriability challenges faced by the developers of enabling technology. Otherwise, society will deny itself the benefits of critical inventions and innovations.

Opening the aperture of PFI from an individual innovation to the broader industrial ecosystem brings in complications. But the basic story is the same. Appropriability is always a challenge; conquering it requires good (strategic) management and good business model design along with a supportive policy environment. At the business ecosystem level, it may require building and controlling a platform. Additional advice embedded in PFI is for managers to also shape the business environment to the extent possible, in order to profit from it. Managers must realize that value capture should be part of every exercise in strategy, business model design, and innovation. Merely being pioneer is not the road to riches.

Policy makers must recognize the challenges that innovators face with respect to capturing enough value to continue innovating in the future. The problems are amplified for inventors of enabling technologies. To keep a society's innovation engine fueled, the government needs to judiciously support value capture, not just value creation. Otherwise, incentives to innovate will be compromised and future citizens will suffer (Teece, Sherry, 2016a). The more fundamental and enabling the technology, the greater the problem.

References

Abernathy, W.J., Utterback, J.M. 1978. Patterns of industrial innovation. *Technology Review*, 80(7), 40–47.

- AFP. 2013. Apple reiterates its ability to “create magic.” *Taipei Times*, 14 February.
<http://www.taipetimes.com/News/biz/archives/2013/02/14/2003554839> (accessed February 15, 2013).
- Anderson, P., Tushman, M. L. 1990. Technological discontinuities and dominant design. *Administrative Science Quarterly*, 35(4), 604–633.
- Armour, H. O., and Teece, D. J. 1980. Vertical integration and technological innovation. *Review of Economics and Statistics*, 62(3), 470-474.
- Arrow, K. J., 1962. Comment, in R. R. Nelson (ed.), *The Rate and the Direction of Inventive Activity: Economic and Social Factors*. Princeton, NJ: Princeton University Press, 353-358.
- BCG 2015. *The Mobile Revolution: How Mobile Technologies Drive a Trillion-Dollar Impact*. Boston: Boston Consulting Group.
https://www.bcgperspectives.com/content/articles/telecommunications_technology_business_transformation_mobile_revolution/ (accessed June 16, 2016).
- Bernhofen, D. M., El-Sahli, Z., and Kneller, R. 2016. Estimating the effects of the container revolution on world trade. *Journal of International Economics*, 98, 36-50.
- Bresnahan, T., and Trajtenberg, M. 1995. General purpose technologies: ‘engines of growth’? *Journal of Econometrics*, 65(1), 83-108.
- Carlaw, K. I., and Lipsey, R. G. 2002. Externalities, technological complementarities and sustained economic growth. *Research Policy*, 31(8), 1305-1315.
- Chesbrough, H., and Rosenbloom, R. S. 2002. The role of the business model in capturing value from innovation: evidence from Xerox Corporation’s technology. *Industrial and Corporate Change*, 11(3), 529-555.
- Chesbrough, H., and Teece, D. J. 1996. When is virtual virtuous: organizing for innovation. *Harvard Business Review*, 74(1), 65–73.
- Christensen, C.M. 1997. *The Innovator’s Dilemma*. Boston, MA: Harvard Business School Press.
- Clark, K.B. 1985. The interaction of design hierarchies and market concepts in technological evolution. *Research Policy*, 14(5), 235–251.
- Cournot, A. A. 1960. *Recherches sur les Principes Mathematiques de la Theorie des Richesses*. Translated into English by N. Bacon as *Researches in the Mathematical Principles of the Theory of Wealth*. London: Haffner.
- Crawford, R. 1932. Did Thomas Edison die a poor man? *Modern Mechanics and Inventions*, 7(3), 54-57, 177.
- Cusumano, M. A., and Gawer, A. 2002. The elements of platform leadership. *MIT Sloan Management Review*, 43(3), 51-58.
- David, P. A. 1992. Heroes, herds and hysteresis in technological history: Thomas Edison and ‘The Battle of the Systems’ reconsidered. *Industrial and Corporate Change*, 1(1), 129-180.
- Dedrick, J., Kraemer, K. L., and Linden, G. 2011. The distribution of value in the mobile phone supply chain. *Telecommunications Policy*, 35(6), 505-521.
- Edgeworth, F. Y. 1925. The pure theory of monopoly, in *Papers Relating to Political Economy, Vol. I*. London: Macmillan and Co., 111-142.

- Evans, P. C., and Gawer, A. 2016. *The Rise of the Platform Enterprise A Global Survey*. New York: Center for Global Enterprise. http://www.thecge.net/wp-content/uploads/2016/01/PDF-WEB-Platform-Survey_01_12.pdf (accessed February 16, 2016).
- Fishlow, A. 1965. *American Railroads and the Transformation of the Ante-Bellum Economy*. Cambridge, MA: Harvard University Press.
- Fogel, R. W. 1964. *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore: Johns Hopkins University Press.
- Galetovic, A., and Gupta, K. 2016. Royalty stacking and standard essential patents: theory and evidence from the world mobile wireless industry. Draft dated June 2016. <http://ssrn.com/abstract=2790347> (accessed, August 1, 2016).
- Galetovic, A., Haber, S., and Levine, R. 2015. An empirical examination of patent holdup. *Journal of Competition Law and Economics*, 11(3), 549-578.
- Gawer, A. 2014. Bridging differing perspectives on technological platforms: toward an integrative framework. *Research Policy*, 43(7), 1239-1249.
- Gawer, A., and Cusumano, M. A. 2002. *Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation*. Boston: Harvard Business School Press.
- Harris, M. W. 2014. Snapping up Kodak. *IEEE Spectrum*, 51(2), 30-62.
- Hawke, G. R. 1970. *Railways and Economic Growth in England and Wales, 1840-1870*. Oxford: Clarendon Press.
- Hazlett, T., Teece, D., Waverman, L. 2011. Walled Garden Rivalry: The Creation of Mobile Network Ecosystems. George Mason University Law and Economics Research Paper Series #11-50. http://www.law.gmu.edu/assets/files/publications/working_papers/1150WalledGardenRivalry.pdf (accessed August 1, 2016).
- Helfat, C. E., and Teece, D. J. 1987. Vertical integration and risk reduction. *Journal of Law, Economics, and Organization*, 3(1), 47-67.
- Helpman, E. (ed.) 1998. *General Purpose Technologies and Economic Growth*. Cambridge, MA: MIT Press.
- Henderson, R.M., Clark, K.B. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, 35(1), 9–30.
- Hicks, J. 1970. Elasticity of substitution again: substitutes and complements. *Oxford Economic Papers*, 22(3), 289-296.
- Hirshleifer, J. 1971. The private and social value of information and the reward to inventive activity. *American Economic Review*, 61(4), 561-574.
- Jovanovic, B., and Rousseau, P. L. 2005. General purpose technologies, in P. Aghion and S. N. Durlauf (eds.), *Handbook of Economic Growth*, 1. Amsterdam: Elsevier, 1181-1224.
- Keenan, M. 2003. Identifying emerging generic technologies at the national level: the UK experience. *Journal of Forecasting*, 22(2-3), 129-160.
- Kushida, K. E. 2011. Leading without followers: how politics and market dynamics trapped innovations in Japan's domestic "Galapagos" telecommunications sector. *Journal of Industry, Competition and Trade*, 11(3), 279-307.
- Langlois, R. N., and Robertson, P. L. 1992. Networks and innovation in a modular system: Lessons from the microcomputer and stereo component industries. *Research Policy*, 21(4), 297-313.

- Levin, R. C. 1982. The semiconductor industry, in R. R. Nelson (ed.), *Government and Technical Progress: A Cross-Industry Analysis*. New York: Pergamon Press, 9-100.
- Levin, R.C., Klevorick, A.K., Nelson, R.R., Winter, S.G., Gilbert, R., and Griliches, Z. 1987. Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity*, 1987(3), 783–831.
- Lipsey, R. G., Bekar, C., and Carlaw, K. 1998. What requires explanation?, in Helpman, E. (ed.), *General Purpose Technologies and Economic Growth*. Cambridge, MA: MIT Press.
- Malmgren, H. B. 1961. Information, expectations and the theory of the firm. *Quarterly Journal of Economics*, 75(3), 399-421.
- Mansfield, E. 1985. How rapidly does new industrial technology leak out? *Journal of Industrial Economics*, 34(2), 217–223.
- Mansfield, E., Schwartz, M., Wagner, S. 1981. Imitation costs and patents: an empirical study. *Economic Journal*, 91(364), 907–918.
- Merritt, R. 2016. "Google's Nest Shares Thread Code." EETimes.com, May 11.
http://www.eetimes.com/document.asp?doc_id=1329656& (accessed May 11, 2016).
- Milgrom, P., and Roberts, J. 1994. Complementarities and systems: understanding Japanese economic organization. *Estudios Economicos*, 9(1), 3-42.
- Nelson, R. R. 1989. What is private and what is public about technology? *Science, Technology & Human Values*, 14(3), 229–241.
- Norton, S. CIO Explainer: What is Artificial Intelligence? *wsj.com*, July 18, 2016.
<http://blogs.wsj.com/cio/2016/07/18/cio-explainer-what-is-artificial-intelligence/> (accessed August 1, 2016).
- Parker, G., and Van Alstyne, M. 2015. Platform strategy, in M. Augier and D. J. Teece (eds.), *The Palgrave Encyclopedia of Strategic Management*.
<http://www.palgraveconnect.com/esm/doi/10.1057/9781137294678.0794> (accessed August 1, 2016).
- Richardson, G. 1972. The organisation of industry. *Economic Journal*, 82(326), 883–96.
- Samuelson, P. A. 1974. Complementarity: an essay on the 40th anniversary of the Hicks-Allen revolution in demand theory. *Journal of Economic Literature*, 12(4), 1255-1289.
- Schumpeter, J. A. 1934. *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.
- Schumpeter, J. A. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper & Brothers.
- Somaya, D., Teece, D. J., 2006. Patents, licensing and entrepreneurship: effectuating innovation in multi-invention contexts. In *Entrepreneurship, Innovation, and the Growth of Free-Market Economies*, E. Sheshinski, R. J. Strom, and W. J. Baumol (eds.). Princeton: Princeton University Press, 185-212.
- Somaya, D., Teece, D. J., and Wakeman, S. 2011. Innovation in multi-invention contexts: mapping solutions to technological and intellectual property complexity. *California Management Review*, 53(4), 47–79.
- Teece, D. J. 1984. Economic analysis and strategic management. *California Management Review*, 26(3), 87-110.
- Teece, D. J. 1986. Profiting from technological innovation. *Research Policy*, 15(6), 285–305.

- Teece, D. J. 1988a. Capturing value from technological innovation: integration, strategic partnering, and licensing decisions. *Interfaces*, 18(3), 46-61.
- Teece, D. J. 1988b. Technological Change and the Nature of the Firm," in G. Dosi, C. Freeman, R. Nelson, G. Silverberg, and L. Soete (eds.), *Technical Change and Economic Theory*. London: Pinter, 256–281.
- Teece, D. J. 2000. *Managing Intellectual Capital: Organizational, Strategic, and Policy Dimensions*. Oxford: Oxford University Press.
- Teece, D. J. 2006. Reflections on profiting from innovation. *Research Policy*, 35(8), 1131–1146.
- Teece, D. J. 2007. Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28(13), 1319-1350.
- Teece, D. J. 2010a. Technological innovation and the theory of the firm: the role of enterprise-level knowledge, complementarities, and (dynamic) capabilities, in N. Rosenberg and B. Hall (eds.), *Handbook of the Economics of Innovation*, Amsterdam: Elsevier, 1–15.
- Teece, D. J., 2010b. Business models, business strategy and innovation. *Long Range Planning*, 43(2), 172-194.
- Teece, D. J. 2012. Next generation competition: new concepts for understanding how innovation shapes competition and policy in the digital economy. *Journal of Law, Economics and Policy*, 9(1), 97–118.
- Teece, D. J. 2014. The foundations of enterprise performance: dynamic and ordinary capabilities in an (economic) theory of firms. *Academy of Management Perspectives*, 28(4), 328–352.
- Teece, D. J., and Sherry, E. F. 2016. The IEEE’s new IPR policy: did the IEEE shoot itself in the foot and harm innovation? Manuscript dated August 3, 2016.
- von Tunzelmann, G. N. 1978. *Steam Power and British Industrialization to 1860*. Oxford: Clarendon Press.
- Williamson, O. E. 1975. *Markets and Hierarchies*. New York: Free Press.
- Williamson, O. E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.
- Zott, C., Amit, R., and Massa, L. 2011. The business model: recent developments and future research. *Journal of Management*, 37(4), 1019-1042.

Appendix: The Digital Revolution and its Effects

This appendix provides background information on digital technology and its impact on firms, business ecosystems, and industries. The mass production, ongoing enhancement, and widespread use of digital logic circuits has led to revolutionary improvements in the technologies they enable, including computers, telecommunications gear, the Internet, and, most recently, wireless networks that can, for most practical purposes, rival wired networks for speed. Advances in wireless broadband technology have enabled not only smartphones but now the emerging “internet of things” (IoT), in which an array of objects embedded with sensors, microprocessors, and wireless communication can be remotely monitored and controlled, from nannycams to driverless cars to factories.

The digital revolution has steadily converted more and more analog information into digital formats, making it more amenable to automated processing. Once information is digitized, it can easily move

between media, and across vast distances, with reliable reproduction. This has had enormous engineering and organizational consequences. For instance, the widespread implantation of digital sensors has helped blur the lines between manufacturing and services. Rolls Royce no longer sells engines; it sells hours of thrust applied to the plane in the sky (and on the ground). The implications not just for the performance of separate products and systems but for connectedness and complementarities are palpable.

The old way of manufacturing and marketing involved making several parts and welding, screwing, or riveting them together. Now a growing number of products can be designed on a digital computer and printed on a 3D printer that creates a solid object by adding successive layers of material. Because these factories need fewer skilled manual workers, they can be anywhere, affecting the distribution of returns to innovation. The relentless miniaturization of microchip circuitry allows once unthinkable computing power to be embedded into a large range of objects from cameras to industrial robots. Industries are being digitized too, most notably retail, led by young brands like eBay and Amazon, while technology is more slowly integrated by the most agile of the older merchants (while the less agile fade away). The digital revolution creates a virtuous cycle. The engineering of next-generation components is aided by the advances of the previous generation. Even fundamental scientific research is benefiting from digitally enabled research tools, such as those used in genomics.

Digital technologies allow easy inter-operability conditional on common standards. As a result, products that were once separate are more easily integrated. The digital revolution has abolished borders between telephones, music players, the web, TVs, cameras, and more. Convergence in the smartphone, for example, benefits end users by allowing access to multiple services from a single handset. This offers conveniences and simplifications along with cost saving. At the same time, it pushes providers of content and services towards horizontal and vertical partnerships to leverage the power of business ecosystems.

Digital convergence has its roots in the 1980s when telecoms and computing first converged with the introduction of digital switching in the central office. The 2G cellular standard, which was rolled out worldwide in the 1990s, brought the benefits of digital telephony to consumer devices. While all digital handsets were in some sense computing devices, it was not until the 2007 introduction of the iPhone that PC-like functionality gained widespread popularity in a mobile phone. The capabilities of smartphones have continued to advance to the point that annual PC shipments have been declining slightly since 2012, after decades of fairly steady growth.

In the computing era that predated wireless-driven convergence, industry boundaries were somewhat distinct. Even today, when an increasing number of firms are positioned across industry lines, not all industries are seeing their boundaries breached. For instance, even though digital technologies are being deployed in farming, the agricultural sector is not converging (yet) with any other (although biotech is a possible candidate).

The drivers of convergence are (1) a common (0,1) base for handling diverse types of information, including words, sounds, and images; (2) widespread use of common standards, which allows connectivity between diverse information devices; and (3) the advance of enabling technologies, including computers, data storage, batteries, and wireless communications. These forces are pushing multiple industries toward a “Grand Convergence” in a fully digitized and integrated swath of the economy that encompasses banking, computing, advertising, social media, print, broadcast, camera,

timekeeping, mapping, insurance, and even, to some extent, education. Digital technologies will be not just a storefront in these industries but their engine. Mobile wireless will integrate these sectors by providing ubiquitous and continuous connectivity with consumers through devices that will go beyond phones to include wearables and, perhaps one day, implants.

Convergence-driven ecosystems have drastically enlarged the partnership networks with which companies must engage. ARM is a company that provides a form of intellectual property for the design of a processing unit that is incorporated by its customers into their microchips, which in turn go into electronic devices in a growing array of industries (“verticals”) that are adding digital intelligence to the equipment they rely on. In the words of ARM’s CEO, Simon Segars:

... Way back, our ecosystem was the semiconductor companies we were working with. We focused on working with our licensees.... Over time, we started working with more and more people up and down the supply chain. All that time we’ve been expanding that ecosystem as the use cases evolve and the range of the technology evolves. So we talk to more and more people, and when you get into IoT, there seems to be an almost limitless ecosystem you need to develop in these different verticals.²⁰

The disastrous effects of a failure to connect with ecosystem partners is demonstrated by the problems that a Japanese wireless carrier, NTT Docomo, had when it tried to take its domestic success overseas. Docomo’s i-mode system, launched in 1999, became one of the first successful wireless data services. The i-mode service, limited by the 2G cellular technology of the time, allowed keypad phones to access email and certain specially redesigned web pages. It also included a simplified version of an app store through which third parties provided i-mode users with paid services and content and shared the profits with NTT. Although i-mode was wildly successful—and profitable—in Japan, efforts to export it failed. NTT invested heavily in overseas partnerships with firms like AT&T Wireless, but failed to convince them to adopt the integrated i-mode business model (Kushida, 2011). NTT also faced an equipment problem in export markets because the Japanese companies making i-mode phones had no presence outside Japan, where the wireless standards at the time were incompatible with those of most other countries. NTT had difficulty convincing the leading non-Japanese phone manufacturers, particularly then-dominant Nokia, to develop i-mode compatible devices. Another element of the ecosystem, i-mode compatible content, was also in short supply. In 2002, NTT’s partners in the US and Europe began to roll out i-mode-based services, but the uptake by consumers was poor and Docomo took a big write-down on its overseas investments. It was not until 2007 that Apple would crack the wireless data ecosystem challenge outside Japan, in large part by internalizing the key elements of hardware and software and by relying on standard web formats rather than requiring specially designed content.

An important case demonstrating how digital convergence can impact the nexus of profit in and/or across industries is cameras. Kodak invented a mass-market camera (the Box Brownie) in the early 1900s and refined its designs incrementally over decades as it made most of its money from selling photographic film. Although a Kodak engineer demonstrated a digital camera prototype in 1975, the company’s film business was so profitable that it only slowly explored the digital camera opportunity, failing to recognize the technology’s disruptive potential. It continued to invest in R&D, with results that included the first megapixel image sensor. However, Kodak’s management had no sense of urgency and

²⁰ Quoted in “Executive Insight: Simon Segars,” June 8, 2016, <http://semiengineering.com/executive-insight-simon-segars-2/>

digital camera technologies were commercialized most effectively by companies other than Kodak. Digital cameras came to market in the 1990s and thereafter followed the same “smaller, better, cheaper” trajectory as personal computers. In the 2000s, as Kodak responded to the faster-than-expected decline of its film business with entries in the digital camera and color printer markets, it looked to its patent portfolio as a source of additional revenue and was able to earn more than \$3 billion in licensing royalties (Harris, 2014). By the early 2010s, all the world’s leading film-based camera companies were either in decline, or, in the case of both Kodak and Polaroid, bankrupt. Canon and Nikon, which were both already diversified beyond photography, survived as camera vendors by retreating to the professional and semi-pro niche markets. Kodak eventually attempted to sell a share of its patents that had been valued at well over \$1 billion. Although other digital patent portfolios sold around this time for far more than their initial valuations, Kodak’s went for \$94 million, a mere fraction of the pre-auction estimate, in part because the company was clearly teetering on the edge of bankruptcy (Harris, 2014).

The Kodak story clearly shows the danger of underestimating the risk of digital disruption. For reasons most likely associated with managerial myopia, Kodak was slow to take the digital threat to its film business seriously, and then was slow to monetize its intellectual property. It was also slow to recognize that a converged device, the camera-phone, would undermine the digital camera and color printer markets. In short, timing remains a vital element of the PFI toolkit.

The music industry tells a similar story of fast-moving disruption. Starting in the mid-1990s, music could be digitized and compressed by consumers using a formal standard that was illegally reverse engineered and spread widely without charge. Illegal downloading devastated record stores, musical artists, and record labels, which were slow to offer legal alternatives. Apple improved the situation with the introduction of the iPod, followed by the iTunes Music Store, but the ability to buy individual songs instead of full albums shattered the old business model for the record industry. The same devastation was visited on the map industry, once dominated by Rand McNally, Shell Oil, and National Geographic. Maps and routing are now provided free by Google and others.

Wireless communications are at the core of the current phase of digital convergence. It is not just that mobile phones have evolved from feature phones to smartphones (a merged mobile communication and computing device), it is also that applications are moving to the cloud. The convergence is coming together in the data center as well as in mobile devices. For example, the vending industry is being animated by mobile applications that allow payment and delivery by Uber drivers of items from high-end automated vending machines, linking together a series of cloud-based transactions to provide consumer convenience. As industries converge around powerful digital platforms, the development and ownership and/or control of complementary assets/technologies will be central to competitive outcomes and the distribution of profits from innovation.