Intangible Assets and a Theory of Heterogeneous Firm

David J. Teece
Tusher Center for Intellectual Capital
Haas School of Business
University of California, Berkeley
email: teece@haas.berkeley.edu

January 21, 2014

The is the pre-proof version of a chapter in A. Bounfour and T. Miyagawa (eds.), 2015. Intangibles, Market Failure and Innovation Performance. New York: Springer, pp. 217-239. The final version can be found at http://dx.doi.org/10.1007/978-3-319-07533-4_9

Abstract

This article outlines a capabilities-enriched economic theory of the firm and its sources of competitive advantage. The nature and key categories of intangibles are discussed, with an emphasis on their suitability for providing differentiation in an era when so many services and tangible goods are readily available on a global basis. The linkages in the conversion of intangibles into profits are analyzed, including the frequent need for co-specialized complements. Among the key categories of intangibles are organizational capabilities, which can be either ordinary or dynamic. Ordinary capabilities are, generally, those that can be measured against best practice and with some effort, imitated by rivals. Dynamic capabilities, which reside in both signature processes and management skills, allow the enterprise and its top management to develop conjectures about the evolution of consumer preferences, business problems, markets, and technology; validate them; and realign assets and competences to enable continuous innovation for the creation of competitive advantage. The key concepts of complementarity, entrepreneurial management, and dynamic capabilities are then applied to deepening the economic theory of the firm, combining with the dominant transaction cost approach to provide a richer understanding of why firms are needed in the economic system.

Key Words: asset orchestration, competitive advantage, complements, dynamic capabilities, entrepreneurial management, intellectual property, know-how, resources, transaction cost economics, theory of the firm
In the nineteenth and twentieth centuries, the assets that economists saw as sources of value were the traditional factors of production—land, labor, and capital—which were scarce and/or stayed within national boundaries. While these factors remain important for national economies, their mere ownership by firms does not guarantee that the firm will generate profits, because labor and capital have become highly mobile.

In fact, in today’s global economy, most intermediate goods and a great deal of the world’s information are so widely available that some say the world is “flat” (Friedman, 2007), i.e., uniformly globalized. It is well recognized that a consequence of efficient factor markets is that it will be hard for any firm to earn better than a competitive return (Barney, 1986).

The notion of “flatness” is, however, an extreme simplification. In reality, the integration of markets for products, people, and ideas are far from complete and the world remains “semiglobalized” (Ghemawat, 2003). In particular, the (dynamic) capabilities required for business enterprises to learn and orchestrate (coordinate and control) resources globally remain scarce, and many types of intangibles do not “travel” easily.

The nexus of reduced barriers to global trade and investment and continued limits to the transfer of capabilities and know-how shapes competitive advantage. As a result, the development and astute management of intangible resources are central to sustained enterprise competitiveness. There are obvious implications for national economic growth and development, too.
The new global reality necessitates the development of new conceptual frameworks for business and economic analysis. As former U.S. Federal Reserve Chairman Alan Greenspan remarked nearly a decade ago, “we must begin the important work of developing a framework capable of analyzing the growth of an economy increasingly dominated by conceptual products” (Greenspan, 2004).

Perhaps surprisingly, mainstream economic theory has almost completely failed to come to grips with the role of intangibles, including the intuition and skills of top management, in creating value. Economists, from Adam Smith on, have never had much to say about the role of managers in coordinating the tangible and intangible assets of business enterprises in ways that both create and capture value. Perhaps the reason is that the task of understanding these issues is daunting. Indeed, figuring out the foundations, at a deep level, of enterprise-generated cash flow continues to be one of the greatest conundrums in economic and financial theory. Even management scholars struggle to arrive at an answer with any generality.

To lay out a capabilities-enriched economic theory, the paper proceeds as follows. It begins with a discussion of the nature and key categories of intangibles, emphasizing their importance for the generation of competitive advantage. The links between intangibles and profits are laid out, with an emphasis on the role of co-specialized complements. Special attention is then given to delineating a category of intangibles known as organizational capabilities and to analyzing the critical role of dynamic
capabilities in the creation and maintenance of competitive advantage. The key concepts of complementarity, entrepreneurial management, and dynamic capabilities are then applied to the enrichment of the economic theory of the firm, combining with the dominant transaction cost approach to provide a deeper understanding of why firms are needed in the economic system.

1 Intangible Assets and the VRIN Criteria

In the field of strategic management, the “resource-based” theory of the firm puts great emphasis on the importance of VRIN assets, those that are valuable, rare, inimitable, and non-substitutable (Barney, 1991). The resource approach sees competitive advantage as flowing from a firm’s VRIN resources. As explained later in this paper, the resource-based approach is not an adequate theory of the sources of competitive advantage. One also needs (dynamic) capabilities and strategy (see Figure 1).

That said, VRIN resources are important building blocks for any theory of competitive advantage. Furthermore, the most important class of VRIN assets is that of intangibles, or what might be referred to as intellectual capital (Teece, 2000).
Ownership (or control) of intangibles and their complements allows innovating firms to differentiate and establish some degree of competitive advantage. The augmentation and orchestration of these assets helps (along with strategy) to generate longer-run enterprise competitive advantage.

This is true across virtually all industries. Consider petroleum extraction. At least as important as the ownership of oil and gas reserves are a company’s unique exploration and extraction technologies, the ability to deploy them effectively and safely, and relationships established over time with contract suppliers, regional authorities, and nation states.
In short, intangible assets are a very economically significant asset class, with powerful implications for building and maintaining competitive advantage for the enterprise (and for a nation). Yet most intangible assets are not even recorded on corporate balance sheets.

### 1.1 Characteristics

Table 1 summarizes the differences between intangible and physical assets along selected dimensions.

<table>
<thead>
<tr>
<th></th>
<th>Intangible Assets</th>
<th>Physical Assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rival in use?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Property rights</td>
<td>Narrow and often fuzzy</td>
<td>Broad and relatively clear</td>
</tr>
<tr>
<td>Measurement and valuation</td>
<td>Relatively difficult</td>
<td>Relatively straightforward</td>
</tr>
<tr>
<td>Tradability</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Recognized on balance sheets</td>
<td>Only in limited ways</td>
<td>Yes (at book value)</td>
</tr>
<tr>
<td>Potential strategic (VRIN) importance</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

First, intangibles differ from physical assets because they are not what economists call “rival in use,” i.e., consumption by one individual does not reduce the amount left for another. One engineer’s use of Newton’s laws of motion does not subtract from the ability of others to use the same laws. However, while the use of particular industrial
knowledge by multiple users will not reduce the availability of that knowledge, and, in some cases, will enhance it, the economic value of the knowledge may well decline, unless no users are direct competitors in the market.

Another important difference between intangible and physical assets is the availability and enforceability of property rights. Physical assets, such as plant, equipment, and land, are generally well protected, at least in developed countries.⁠¹ Ownership is relatively easy to define, and the “boundaries” of the property are usually obvious. Whether theft has occurred is relatively easy to ascertain, and enforcement of the property right is generally available. Intangible assets and intellectual capital, on the other hand, have relatively poor protection. Although there are some exceptions, intellectual property rights are narrowly delineated.

Whereas most physical assets can be bought and sold in “thick” (i.e., liquid) markets with relative ease (apart from equipment that has been highly customized), markets for most intangibles, if they exist at all, will be “thin.” This is in part because of the limited nature of property rights surrounding intangible assets. It’s also because the value of a knowledge asset often derives from the presence of complementary assets in a way that is context-dependent. Certain knowledge assets (such as technological capabilities) cannot be meaningfully secured without acquiring a company or business unit, then finding a way to retain key personnel. Furthermore, some knowledge assets can be costly to transfer following a purchase (Teece, 1981). The number of buyers who will be willing

---

¹ Needless to say, through land use and other controls, national governments and local authorities can dramatically impair the value of real estate by limiting its use.
and able to pay for a knowledge asset’s full potential strategic value (i.e., its value in use to the present owner) is generally limited.

This nontradability is central to the (strategic) importance of intangibles. In a world where most assets and services are available for hire, the ownership of unique, non-tradable assets like intangibles offers a potential source of advantage.

Whereas tangible assets are generally included on balance sheets, intangible assets are less readily measured, and their valuation remains too controversial for financial accounting standards boards to agree upon a methodology. As a result, intangibles are mostly absent from a firm’s financial statements. Under international accounting standards (IFRS, 2012), only non-physical, non-financial assets that are technically separable from the physical and human resources of the firm can be reported as assets. Examples include patents, copyrights, trademarks, customer lists, franchises, marketing rights, software, and digital content.

Intangibles for internal use, such as improved business processes and better-trained staff, are excluded from financial statements. Investments in the creation of new intangibles, including major categories such as marketing and R&D, continue to be mostly expensed rather than capitalized. The chief exception arises from mergers and acquisitions, where accounting rules treat any purchase premium over book value as “goodwill,” a non-separable intangible asset recognized by accountants.
Corporate balance sheets are thus poor proxies for the economic value of the assets of an enterprise. Moreover, management is often unaware of some of the firm’s knowledge assets and of the deeper sources of its competitive advantage. As the saying goes, you cannot manage what you cannot measure. Yet the effective management of intangibles is one of the most likely foundations for profitability.

The creation of intangibles can be very challenging, depending on the characteristics of the technology involved (Teece, 2005). The most common way of adding to the enterprise’s stock of intangibles is investment in learning activities, including formal R&D. One reward for this effort is that imitation and replication of intangibles by rivals is often hard, which provides some insulation against the dissipation of profits.

One common feature of both physical and intangible assets is depreciation, or obsolescence. While knowledge does not “wear out” as do most physical assets do, knowledge is frequently subject to rapid depreciation because of the creation of new inventions and innovations. If a firm’s own renewal process does not make its existing knowledge obsolete, then a competitor’s knowledge activities are likely to do so.

### 1.2 Types

There are many types of intangibles. The patent, a form of intellectual property, is perhaps the best known.
A valid patent theoretically provides rights for exclusive use of an invention by the patent owner, but reality is seldom so simple. The validity of a patent may need to be proved in court at considerable expense before it is accepted by rivals. Ascertaining whether infringement has occurred can be difficult. There can also be “holes” and “gaps” in intellectual property coverage. Moreover, patents (and copyrights) eventually expire.

Trade secrets, another class of intangible, can augment the value of a patent position. They do not provide rights of exclusion over a knowledge domain, but they protect covered secrets in perpetuity. Trade secret protection is possible, however, only if a firm can put its product before the public and still keep the underlying technology secret. This is most likely to be true of industrial processes.

Trade secrets are part of a broad and critical class of intangible called know-how. Know-how is often embedded in the organization as a whole, which can make it the most difficult element of a product’s value chain for rivals to imitate. Thus Dell’s direct sales and build-to-order business model was embodied in manufacturing, distribution, and IT systems that competitors found hard to imitate, at least for many years (Kraemer et al., 2000). Capabilities, about which more will be said later, are related to know-how.

Another intangible asset of central importance is the firm’s business model,2 i.e., the logic of how a business creates and delivers value to customers while earning a profit for itself (Chesbrough & Rosenbloom 2002; Teece, 2010). Business model innovations

---

2 Business models in their entirety are generally not protected by intellectual property rights. Certain elements of a model might qualify for patent or copyright protection.
are critical to success in unsettled markets where traditional revenue and pricing models are no longer applicable. The growth of the Internet is both allowing and requiring business model innovation in many industries ranging from music to insurance. In particular, the Internet requires new pricing structures for many products because users are now accustomed to getting information for free. In other industries, middlemen serving as information brokers are being disintermediated.

Other interesting classes of intangible assets include brands, customer and business relationships, and organizational culture.

2 Profiting from Intangibles

Markets are a great leveler. If assets or their services are traded in a market, they can be accessed by all who can pay. The range of domains in which competitive advantage can be built narrows as more and more activities become outsourceable. The Internet and other recent innovations have vastly expanded the number and type of goods and services that are readily accessed externally.

Intangible assets, perhaps the most important category of non-tradables, have the potential to form a basis for long-term profitability if the assets are astutely managed. However, intangible assets by themselves will not generally yield value; they must almost always be combined with other intangible and physical complements in a way that
yields value for a customer. Then the assets and complements must be managed in a way that maximizes appropriability.

2.1 Co-specialized Complements and Ecosystems

The aggregate economic value achieved by combining two or more complementary assets exceeds the value that would be achieved by using these factors in different activities. When complements are worth far more together than in any other separate uses, the complements are said to be co-specialized, and managerial “control” of the complements becomes critical. Whether that control is achieved through ownership or simply through setting the rules for a supporting ecosystem depends on the facts and circumstances. Complements that are not available from competitively-priced suppliers must generally be owned by the focal firm to avoid dissipating profits (Teece 1986, 2006).

Complementarity is not a new phenomenon. Rosenberg (1979: 26) notes: “Time and again in the history of American technology it has happened that the productivity of a given invention has turned on the availability of complementary technologies.” Furthermore, “the growing productivity of industrial economies is the complex outcome of large numbers of interlocking, mutually reinforcing technologies, the individual components of which are of very limited economic consequences by themselves. The smallest relevant unit of observation, therefore, is seldom a single innovation but, more typically, an interrelated clustering of innovations” (Rosenberg, 1979: 28-29).

---

3 The VRIN criteria discussed earlier tend to overlook this point, i.e., the V of VRIN is likely to be highly context dependent.
The ability to assemble interdependent configurations of co-specialized assets, as in the case of systemic innovation (Teece, 1984), can provide a unique value proposition. Common ownership of intangibles and certain complements will enable them to co-evolve in a coordinated way.

Co-specialization is becoming ubiquitous for devices and services that span multiple industries, such as smartphones that combine functions of computing, communication, and consumer entertainment products. As former Nokia CEO Stephen Elop said in his February 2011 (internal) “burning platform” memo, “The battle of devices has now become a war of ecosystems, where ecosystems include not only the hardware and software of the device, but developers, applications, ecommerce, advertising, search, social applications, location-based services, unified communications and many other things. Our competitors aren’t taking our market share with devices; they are taking our market share with an entire ecosystem.”

The smartphone is an example of a multi-invention context (Somaya et al., 2011). It is one manifestation of the increase in technological complementarities that has generated a growing need for taking account of external intellectual property rights. Complicated products—particularly those with many components, parts or functions—may “read on” hundreds, if not thousands, of patents. Innovation in one product or service often

---

4 The leaked Nokia memo was widely reproduced online. See, for example, http://www.engadget.com/2011/02/08/nokia-ceo-stephen-elope-rallies-troops-in-brutally-honest-burnin/ (accessed December 26, 2013).
increases the value of their complement(s) and may require the in-licensing of patent portfolios to facilitate design and operating freedom.

### 2.2 Appropriability

The appropriability of the income generated by (or with) a knowledge asset is a function of its (inherent) value, its nature (i.e., the type of knowledge), and its ease of imitation (particularly, the effectiveness of intellectual property rights as a barrier to imitation). 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). Table 2 shows this interaction of imitability and intellectual property rights.

<table>
<thead>
<tr>
<th>Inherent Potential for Imitation</th>
<th>Intellectual Property Rights</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Loose</td>
</tr>
<tr>
<td>Low</td>
<td>Tight</td>
</tr>
</tbody>
</table>

Note: Imitation potential depends on the difficulty and complexity of the relevant know-how.
Source: Teece (2005)
Factors that make imitation difficult enhance appropriability. Thus, the more tacit the firm’s productive knowledge, the harder is imitation by its competitors. When the tacit component is high, imitation may well be impossible, absent the hiring away of key individuals and the (possibly illegal) transfer of key organizational processes.

The tacitness of knowledge varies to some extent over the product cycle. New products and processes are often highly nuanced. Thus in the pre-paradigmatic phase of technological innovation (Abernathy & Utterback, 1978; Teece, 1986), the tacit component is likely to be high. Once a dominant design emerges, the rate of change of product design slows, and there is then the opportunity, if not the need, to codify technology. However, more rapid rates of innovation mean that there may be no time to codify (make explicit) new knowledge even when it is technically feasible to do so.

The observability of a technology also affects imitability. While insight into product technology can be obtained thorough strategies such as reverse engineering, this is generally not the case for process technology. Secrets are thus more protectable if there is no need to expose them in contexts where competitors can learn about them.

A technology becomes covered by intellectual property once it is legally recognized. In the case of patents, the conversion occurs when a particular country’s patent office recognizes the inventor’s application and grants a patent. That’s not the end of it, however. Patents can be (and often are) challenged by users/implementers. Hence, the value of a patent may evolve over time (Sherry & Teece, 2004).
Intellectual property rights vary across jurisdictions in terms of the types of inventions to which they apply, how long they last, and how well they are enforced. But the value of intellectual property also differs across fields of endeavor, not just across industries or countries.

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 et al., 1981; Mansfield, 1985). They are especially ineffective at protecting process innovation. 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.

The more fundamental the invention, the better the chances that a broad patent will be granted, and granted in multiple jurisdictions around the world. The inventor of a core technology can further strengthen appropriability by seeking complementary patents on new features and/or manufacturing processes, and possibly on designs. The way the claims in the patent are written also matters.

---

5 Mansfield et al. (1981) found that about 60 per cent of the patented innovations in their sample were imitated within four years. In a later study, Mansfield (1985) found that information concerning product and process development decisions was generally in the hands of at least several rivals within 12-18 months, on average, after that decision is made. Process development decisions tend to leak out more than product development decisions in practically all industries, but the difference on average was found to be less than six months.
While a patent is presumed to be valid in many jurisdictions, validity is never firmly established until a patent has been upheld in court. The most valuable patents are those that are broad in scope, have already been upheld in court, and cover a technology essential to the manufacture and sale of products in high demand.

The character of the appropriability regime (strong, weak, or in between) should shape strategy. Weak appropriability dictates reliance on other value capture mechanisms, such as developing complementary assets (e.g., an attractive brand image) that would earn a premium even if the intangible itself did not (Pisano & Teece, 2007).

3 Intangible Assets, Resources and Capabilities

For long-term profitability, management must make decisions and take actions to build, modify, and renew intangibles and other resources. A good understanding of how this works is missing from most economic and financial models. The dynamic capabilities framework, which has emerged over the past twenty years in the field of strategic management, provides a theoretical infrastructure in which intangible assets can be seen as jewels. The crown, the frame in which they are assembled, is composed of dynamic capabilities, as explained below. In practice, the crown as well as the jewels need to be constantly revamped in order to support durable value to the realm.

To put this in its larger context, resources are potentially valuable assets (tangible and intangible) and people that are semi-permanently attached to a firm. As discussed above, some of these resources will meet the VRIN criteria. In order to keep their VRIN status,
these resources must be constantly renewed. The need for renewal is amplified in fast-moving environments such as those characteristic of high-tech sectors. However, the need to renew resources also occurs in “low-tech” industries (e.g., life insurance).

How resources are used, and hence the value they generate, depends on the firm’s capabilities, the subject to which this paper now turns. In this regard, it is useful to distinguish between ordinary and dynamic capabilities.

### 3.1 Ordinary Capabilities

Ordinary capabilities can best be thought of as achieving technical efficiency and “doing things right” in basic business functions: operations, administration, and governance. If done to a very high level of performance, an ordinary capability can become known as a best practice. Such capabilities often have a high public domain component, and, even if not, they are readily imitable and can therefore generally be acquired. I don’t mean to denigrate their importance; they are often fundamental. But, on their own, they won’t bring long-run success.

Ordinary capabilities involve the performance of those administrative-, operational-, or governance-related functions that are (technically) necessary to complete currently-planned tasks. They are embedded in some combination of (1) skilled personnel, including, under certain circumstances, independent contractors; (2) facilities and equipment; and (3) processes and routines, including the administrative coordination needed to get the job done.
Ordinary capabilities are mostly technical in nature. Much of the knowledge behind them can be borrowed, or “bought,” through consultants or through a modest investment in training (Bloom et al., 2013). These capabilities can be measured against the requirements of specific tasks (such as good preventive maintenance, or proper supply chain management) and thus benchmarked to best practice. Strong ordinary capabilities are an indication that the firm has achieved best practices and owns or has access to skilled people and advanced equipment. Exercising them keeps people employed. But on their own they do not generate more than a competitive return—and possibly less—except when the competitive environment is very weak.

A recent demonstration of this was provided by a controlled study by Bloom et al. (2013), in which 14 Indian textile plants were taught a set of 38 well-known (in developed countries) management practices, resulting in a 17% increase in productivity in the first year. The apparent reason for the firms’ initial (avoidable) inefficiency was that the Indian managers had either not known about the superior practices or had been skeptical of what they had heard.

However, in an environment open to global competition where firms can look to similar benchmarks and have access to competitive off-the-shelf technologies and training, good and even “best” practices will diffuse rather quickly among at least some firms. The management consulting industry works hard to introduce clients to new and better—and
usually non-proprietary—“best” practices, which contributes to making best practices nearly universal.

But best practices can become a trap, as the relentless and single-minded pursuit of efficiency can drive out the capacity to effectuate change, and the organization becomes sclerotic. Efficiency is easiest to achieve if the set of tasks the organization is to perform remain fixed. Hence, there is often inertia imposed by efforts to achieve best practice.

3.2 Dynamic Capabilities

Whereas ordinary capabilities are about doing things right, dynamic capabilities are about doing the right things, at the right time, based on unique managerial orchestration processes, a strong and change-oriented organizational culture, and a prescient assessment of the business environment and technological opportunities. Ordinary capabilities contribute to a firm’s technical fitness in specific areas, but strong dynamic capabilities assist firms in achieving overall evolutionary fitness.

Strong dynamic capabilities help enable an enterprise to profitably build and orchestrate its competences and other assets that lie both within and beyond its boundaries, reconfiguring them as needed to innovate and respond to (or bring about) changes in the market and in the business environment more generally (Teece et al., 1997; Pisano & Teece, 2007). They allow the enterprise and its top management to develop conjectures about the evolution of consumer preferences, business problems, markets, and
technology; validate them; and realign assets and competences to enable continuous innovation and change.

Learning is central to such developments. The enterprise must learn (1) what customers want, (2) what new technologies might allow, (3) what aspects of the business model are working, and (4) whether the current strategy is effective and the company is on the path toward building a valuable business.

Dynamic capabilities reside, in part, with individual managers and the top management team. At certain key junctures, the ability of a CEO and the top management team to recognize a key development or trend, then delineate a response and guide the firm in its co-creation activities, may be the most important element of the firm’s dynamic capabilities. But the organization’s values, culture, and its collective ability to quickly implement a new business model or other changes are also integral to the strength or weakness of the firm’s dynamic capabilities.

To the extent that dynamic capabilities are routinized, particularly at upper management levels, these practices are likely to rely on “signature processes” (Gratton & Ghoshal, 2005). Signature processes, characterized in part by the methods and frequency with which top managers interact, arise from a company’s heritage, including its prior management actions, certain irreversible investments, and context-specific learning.
Because of their deep, enterprise-specific roots, signature processes are not so easily imitated by other firms that did not and cannot share this history and that may have a different, incompatible corporate culture as well. Moreover, the replicability of a process or business model is often confounded, particularly externally, by what Lippman and Rumelt (1982) call “uncertain imitability.” This, along with a high tacit component to the underlying knowledge, may keep a signature process effectively proprietary for considerable periods. As a result, signature processes themselves could satisfy the VRIN criteria. Hence, signature processes (and signature business models) are likely to be an important source of inter-firm heterogeneity, at least for a while (Jacobides & Winter, 2012).

Over longer periods of time, however, even signature processes may become imitable by others. This transformation occurred with Toyota’s lean production model, which is a tightly integrated set of processes that encompasses the entire value chain, from product design to customer relations (Womack et al., 1990). The “Toyota Production System” provided the automaker a source of competitive advantage for decades despite numerous and sustained attempts at imitation by rivals. However, it eventually diffused to other firms and even other industries. The multidivisional form (M-form) of business organization is another such example. Armour and Teece (1978) showed how early adopters of the M-form in the petroleum industry reaped significant profits from the new organizational structure, but the M-form-specific profits were competed away after about a dozen years.
Dynamic capabilities encompass how an enterprise obtains strengths, extends these strengths, innovates, synchronizes business processes and models with the business environment, and/or shapes the business environment in its favor. For applied purposes, they can usefully be broken down into three primary clusters: (1) identification, development, co-development, and assessment of technological opportunities in relationship to customer needs (*sensing*); (2) mobilization of resources to address needs and opportunities, and to capture value from doing so (*seizing*); and (3) continued renewal (*transforming*). Sensing, seizing and transforming are essential if the firm is to sustain itself as customers, competitors, and technologies change (Teece, 2007). Asset orchestration is a meta-process that envelops and engages all three clusters.

In firms with strong dynamic capabilities, many actions and activities will take place simultaneously: servicing existing customers, acquiring new ones, developing new products and services, hiring top talent, retaining talent, raising capital, introducing new processes, improving operations, transforming as circumstances change, and so on. This requires what O’Reilly and Tushman (2004) call “ambidexterity,” the ability to simultaneously keep an existing business in tune while actively exploring new opportunities. Ambidexterity is an example of a dynamic capability (O’Reilly & Tushman, 2008). It is especially critical when industries are in rapid transition.

In the modern parlance of Silicon Valley, firms must (and do) “pivot” (Ries, 2011) when inflection points occur in the ecosystem or when they discover that their strategy and/or business model is no longer working. While path dependence poses a constraint on the
future actions of all enterprises, for some firms the legacy of the past, in the form of the dynamic capabilities they have built, can also provide the foundation and fulcrum of future growth.

4 Toward a Capabilities-based Theory of the Firm

This framework of organizational capabilities, and of intangible resources more generally, can shed light on a fundamental question in economics, namely why the enterprise form of organization exists at all when price-based allocation via contracting is generally considered by economists to be efficient. A large literature has grown up addressing the issue of what types of assets and activities will be internal to the firm rather than allocated by the price system. The leading school of thought in this area concerns transaction costs, but transaction cost economics omits consideration of a number of variables that co-determine not only firm boundaries but also firm success or failure.

The dynamic capabilities framework, which posits that knowledge assets and their entrepreneurial management have become central to profit maximization in an era of globalized commerce and information, suggests a new theory of the firm. It combines the transaction-level understanding of the transaction cost framework with the enterprise-level understanding of management studies. In other words, transaction costs and capabilities are complementary, not competing, lenses for analyzing the business enterprise.
4.1 **Transaction Costs and the Boundaries of the Firm**

The primary contribution of transaction cost economics (Williamson, 1975, 1985) to the theory of the firm is in the area of firm boundaries, i.e., the governance modes that the firm will use to conduct its business. To accomplish this, the transaction cost economics framework holds “production” activity constant even though it may depend endogenously on governance modes, as well as on the managerial actions, strategy, and structures chosen. An activity is most likely to be internalized when the assets involved are highly specific to that activity alone. Market modes of governance are seen as likely to “fail” in such cases because contracts between legally separate entities would likely lead to opportunistic renegotiation.

In other words, the appropriability problem of the firm is couched entirely in terms of the risk from opportunistic behavior by potential partners. From such a diagnosis, internalization of the partner’s activity is the single and obvious solution. Co-creation activities, conducted within a strategic alliance, for example, are not considered.

The transaction cost conception of market failure is simply too narrow for some purposes. Williamson (1971), in his best-known statement on market failure, which he endorsed 28 years later (Williamson, 1999), restricted his attention to market failures that were “failures only in the limited sense that they involve transaction costs that can be attenuated by substituting internal organization for market exchange” (Williamson, 1971: 114). In the transaction cost view, entrepreneurial and managerial functions such as opportunity discovery, learning, and knowledge creation play almost no role.
Transaction cost economics, while helpful in many ways, nevertheless deflected attention away from more important issues around the very existence of markets. Market creation and co-creation functions are not merely a response to a market that has somehow failed to perform (relative to an ideal standard). Rather, it is often the case that the market has quite simply failed to emerge and/or needs to be created or co-created by entrepreneurially managed business enterprises (Pitelis and Teece, 2010).

In other words, the rationale for the business enterprise is not just to achieve efficiencies relative to a theoretical market-based benchmark, but also to create and manage co-specialization. This necessitates a deeper understanding of complementarity.

4.2 **Complementarities and Co-specialization**

The theory of the firm has benefited, and can benefit further, from a more rigorous exploration of the concepts of complementarities and co-specialization. Early applications in the innovation literature can be found in Rosenberg (1979, 1982) and Teece (1986). Work on complementarities in a strategic context includes Teece (1980), Miller (1988), and Milgrom and Roberts (1990a, 1990b).⁶

Teece (1980) pointed out that the complementarity of two assets or activities in and of itself has no direct implication for the boundaries of the firm because contractual

---

⁶ For a review of the literature on complementarity and the related mathematical concept of supermodularity, see Ennen and Richter (2010).
arrangements exist that, in theory, can enable joint activities to take place without common ownership of the parts. Assets that are co-specialized to each other, however, need to be employed jointly, usually inside the firm. In the case of innovation, Teece (1980, 1986, 2006) defined contexts in which directly owning complementary assets is important for capturing value.

A robust theory of complementarities that provides economic insight has yet to emerge. While there is little doubt that complementary relationships exist among heterogeneous factors inside the firm (and that these can impact firm performance), the contexts in which such interactions occur have not been adequately specified. However, some evidence has been assembled. Monteverde and Teece (1982), while testing for the importance of asset specificity in predicting outsourcing decisions by GM and Ford, also found that a “systems effect”—defined as “the degree to which any given component's design affects the performance or [system-level integration] of other components” (p.210)—was statistically significant in explaining GM and Ford’s outsourcing decisions.

It should be noted that the notion of complementarity can be applied at a high level of aggregation, as with the Toyota Production System. It can also be applied at a very fine level of specificity, such as the complementarity between the (integrated) design and manufacture of automobile components. Parmigiani and Mitchell (2009) use the example of automobile dashboards, which they note typically consist of multiple, interdependent, complementary components. Both levels of aggregation seem to provide insights, suggesting the power and generality of insights from the concept of complementarity.
Complementarities expressed through their mathematical corollary (supermodularity) represent a rupture with mainstream models of production in economics. With production functions of the standard kind, decision makers need only equate marginal revenues to marginal cost, and they will deliver the (global) maximum in output. Complementarity, modeled as supermodularity, enables some departures from this extreme caricature by recognizing the existence of local maxima, reducing the deterministic nature of the model. A complements-based model of production also implies that design choices are discrete rather than continuous.

There are many circumstances where internal organization is clearly a superior way to organize and orchestrate the innovative activity essential to the renewal of firm resources. The most important (and also the most under-researched) domain within which organization inside the firm is likely to be necessary is the creation, transfer, protection, and orchestration of know-how and other intangibles of multiple, complementary types and/or from multiple disciplines.

Building and assembling co-specialized intangibles inside the firm (rather than accessing them through a skein of contracts) is not done primarily to guard against opportunism and recontracting hazards. While those considerations matter, effective coordination and alignment of the assets is the critical point, and would be virtually impossible to achieve through the price system. The market failure in this type of case is more fundamental than the mere presence of “transaction costs that can be attenuated” by unitary ownership.
In a dynamic capabilities perspective, the entrepreneurial manager must be free to orchestrate highly co-specialized assets. When performed astutely and proactively, such orchestration can: (1) keep the assets in value-creating alignment, (2) identify new co-specialized assets to be developed through the investment process, (3) pursue new market opportunities to which the assets, combined or separated, are suited, and (4) divest assets that no longer yield special value. These goals cannot be readily achieved through contracting mechanisms in part because of dynamic transaction costs but also because there may not be a competent entity to build the assets that are needed. There is limited utility in labeling these business issues as a transactions cost problem.

Although opportunism surely exists and must be guarded against, the emphasis in dynamic capabilities is on creating the assets that in transaction cost economics become the object of rent appropriation. And effective asset creation depends as much on the talent and skill of entrepreneurial managers as on the capabilities embedded in the enterprise itself.

4.3 Managers

Transaction cost economics, and economic theory more generally, leaves us without an understanding of the distinctive role of the manager. Managers must not only choose among market arrangements, alliances, and internal organization; they must also

---

7 Langlois (1992) defines dynamic transaction costs as “the costs of persuading, negotiating, coordinating and teaching outside suppliers” (1992: 113).
understand how to design and implement different governance structures, to coordinate investment activities, to design and implement business models, and to craft appropriability strategies.

As both a theoretical and practical matter, it is important to ask how firms allocate resources so that they are in their first best use. How firms build, augment, and modify their resource base over time is also of critical importance. In other words, there are important resource allocation functions that (neoclassical) economic theory ignores: namely, how does the non-market coordination inside, between, and amongst firms actually take place? Who performs that role when the price mechanism is not available? Economic theory yields poor answers.

An economic theory of markets needs to somehow recognize that a good deal of resource allocation takes place inside firms and between and amongst firms as a result of entrepreneurial and managerial decisions, activated by managerially designed systems. When managers do take the stage in modern economics and finance, the focus is usually on the distribution, not the creation, of the spoils between managers and shareholders (Jensen, 2000). This begs the question of where the wealth of firms comes from in the first place.

Although management skills have long been recognized in practice as a source of value, the proposition is finding new empirical support. Google, for example set up a project to

---

8 See Bloom and Van Reenen (2007) for a notable exception to the virtual exclusion of firm-specific managerial practices from the economics literature.
test the impact of various management practices and found that “even ‘the smallest incremental increases in manager quality were quite powerful’” (Garvin, 2013: 77, citing Neal Patel, co-leader of Google’s study).

Yet managers are scarce even in some versions of the resource-based view of the firm (e.g., Wernerfelt, 1984). An idea first advanced by Penrose (1959) is that every firm has resources, including managerial skills, that can potentially be deployed into multiple product arenas. However, the resource-based view gives scant attention to the processes and skills needed for renewing the firm’s resources. Moreover, Penrose and those who followed never provided any granularity with respect to the skills that undergird the growth and diversification of the firm, particularly the critical entrepreneurial skills of sensing, seizing, and transforming.

Manager are integral to harnessing the hard-to-imitate practices that undergird the generation, ownership, and management of know-how and other intangible assets. The capabilities to build and astutely manage these intangibles and their related complements have come to overshadow production-related economies of scale and scope as determinants of competitive outcomes in many contexts.

Perhaps more importantly, entrepreneurial managers are needed to design organizations that can discover and create new knowledge and then commercialize market-relevant new technologies. Entrepreneurial managers learn about new opportunities and sometimes help create them, transferring technology as needed. The topic of entrepreneurship, in
both new ventures and existing firms, is sufficiently important that it merits separate attention.

4.4 *Entrepreneurship and market creation*\(^9\)

Entrepreneurship is too often left out of theories about how economies function and how enterprises evolve. Most economic theories of the firm, apart from a few based directly on entrepreneurship (e.g., Sautet, 2000), include an implicit assumption that all opportunities are known. And if they are not known, information costs are all that stand in the way of discovery.

But opportunity discovery is often far from straightforward. In globally competitive environments, consumer needs, technology, and competitor activity are constantly in flux. While the path ahead for some emerging marketplace trajectories is easily recognized, most emerging trajectories are hard to discern. Sensing new opportunities amid the noise is very much a learning, creative, and interpretive activity at which, by definition, entrepreneurs excel. Necessary complements to individual insight are research and related activities that draw on expert talent and organizational strengths.

A useful tripartite conception of entrepreneurship was proposed by Sarasvathy et al. (2003): (1) the recognition and arbitrage of pre-existing but as-yet-unmatched supply and demand; (2) the process of discovering and exploiting new uses for existing products, such as recognizing a latent demand for gourmet coffee, or of finding a new way to

\(^{9}\) This section draws on material in Al-Aali and Teece (2014).
supply an existing demand, such as a better cure for a disease; and (3) the creation and exploitation of new opportunities by conceiving of possible future demands and supplies that do not yet exist. The third, market creation form of entrepreneurship requires what Kirzner called “alertness,” which includes “awareness of the ways the human agent can, by imaginative, bold leaps of faith, and determination, in fact create the future for which his present acts are designed” (Kirzner, 1985: 56).

The notion that entrepreneurs must create each market before there are prices and consumer preferences that can lead to economic efficiency dates back to the work of Frank Knight (1921). However, this insight was largely eclipsed, particularly in the economics literature, by the contractual approaches to the firm put forward by Coase (1937), Williamson (1975), and others in which markets, technologies, and prices are simply assumed to exist (Boudreaux & Holcombe, 1989).

Entrepreneurship and contractual (transaction cost) approaches are not incompatible. Foss et al. (2007) combined them in a theory of the firm by positing that a significant reason for the formation of firms in a world of uncertainty is to allow entrepreneurs to experiment with different combinations of heterogeneous capital. Teece (2014) outlined an entrepreneurial theory of the multinational enterprise.

Over time, the coordination and further development of capital assets will render them more and more specific to their use and to each other. Common ownership within the firm is thus the efficient means of preventing the possibility of a future hold-up by an
external owner of one of the assets. It’s also a way to best exploit the complementarities, especially when time and place matter.

Thus entrepreneurial sensing and asset orchestration provide a more complete explanation for the existence of the firm than does transaction cost reasoning alone.

Entrepreneurial managers aren’t simply responding to market failure. They are mobilizing organizational and other resources to stimulate new economic activity. Entrepreneurs are vital to this process because of their ability to form judgments in the face of uncertainty about the conditions in markets that don’t yet exist.

Entrepreneurial activity demands a flexible, iterative approach to decision making (Alvarez & Barney, 2007). Performing the required tasks takes adaptive leadership, deep knowledge of markets, and a clear understanding of the technical, physical, and human constraints of the resources at hand.

Market creation, including co-creation that involves networks and alliances, is a categorically different process from a make-or-buy decision that can be explained by arguing that markets “fail” under certain conditions, such as where complex know-how transfers are involved. The market in this case has yet to emerge, and might never do so in the absence of the entrepreneur.

The view of the firm as fundamentally entrepreneurial and market-creating is markedly different from contractual and market failure approaches. However, while
entrepreneurship deepens an analysis of the existence of firms, it cannot, by itself, account for inter-firm heterogeneity and firm-level competitive advantage because it omits essential elements of environmental fit, strategy, and the need to respond to challenges as well as opportunities. Moreover, entrepreneurship, even in new ventures, is a social process, for the top management team and, ideally, for the whole organization (Foss et al., 2008). In short, dynamic capabilities, which include entrepreneurial management but also much more, must be included in the theory of the firm.

### 4.5 Capabilities

Dynamic capabilities, and organizational capabilities more generally, are all but absent from economic theories of the firm and of markets. The (neoclassical) economic model of market exchange takes for granted that somehow, somewhere new goods and services are being designed, developed, and produced by some method that will be technically efficient, conditional on factor costs. Moreover, it is often assumed that everyone knows all relevant information.

Transaction cost economics implicitly assumes what might be referred to as capabilities neutrality. In transaction cost economics, so called “production costs”—which might be thought of as a proxy for the firm’s level of operational capability—are assumed to be the same across organizational types so that the choice between market and non-market arrangements swings entirely on transaction/governance costs.
The introduction of capabilities to the theory not only helps inform the choice of transaction governance, but also brings the possibility of explaining differences between firms in productive efficiency and profitability as a function of managerial activity. The field of strategic management is built on the recognition that firms are different in ways that drive performance differences (Rumelt et al., 1991).

As noted earlier, the production theory of neoclassical economics (implicitly) assumes away numerous organizational problems, rendering firms more or less interchangeable. A production function (or production sets) assumes specified relationships between inputs and outputs and the existence of a global maximum in most states of the world. Inside the black box that is the firm, best practices are implicitly being followed by all.

An exception to the simplification of economics that all firms operate efficiently is Leibenstein’s (1966) concept of x-inefficiency, which refers to particular firms operating above their cost curves. This allows for firm-level heterogeneity. Leibenstein and others attributed x-inefficiency to a lack of competition; but the more fundamental reason is likely to be poor management and limited information. In any event, Leibenstein’s theory, despite being cited occasionally, has not really been embraced by economists.

The dynamic capabilities framework suggests a theory of the firm that not only accommodates firms with x-inefficiency (i.e., firms with costs above the technically efficient level); they can also suffer from what might be called “d-ineffectiveness” (i.e., weak dynamic capabilities). In other words, (1) not all firms are at the best practice
frontier and (2) even those that have adopted best practice may be producing the “wrong” products relative to current market requirements and technological opportunities.

Thus, a theory of “capability economics” allows for (and helps explain) heterogeneity amongst firms. The so-called Austrian School of economics allows for entrepreneurs and for differences between firms related to differential access to information, but it doesn’t have much room for the manager. There is a place for both the entrepreneur and the manager in capability economics—and in the dynamic capabilities framework more generally.

Firm-level heterogeneity can be, and has been, assessed empirically. Ordinary capabilities are generally measurable and therefore relatively straightforward to compare across firms. Although dynamic capabilities are complex and not always directly observable, researchers have had success assessing them through the use of surveys (e.g., Morgan et al., 2009), secondary sources of data about corporate decisions (e.g., Adner & Helfat, 2003), and the provision of advisory services (Feiler & Teece, forthcoming).

5 Conclusion

This paper has analyzed the central importance of intangibles for the generation of firm-level profits. It then showed how a specific category of intangibles called dynamic capabilities is able to account for persistent firm-level differences that mainstream economics often assumes away. The proposed capabilities-based theory opens up the
black box of the firm and injects into economic theory new considerations that are ignored in most microeconomic and transaction cost models.

A rich understanding of the existence and role of the firm must encompass what successful firms actually do. Concepts such as complementarity, entrepreneurial management, and dynamic capabilities need to be integrated more fully into mainstream models.

In the theory advanced here, management’s task is not just to overcome “failure” in the market for intangibles when structuring the firm; it must also build and leverage distinctive intangible resources, especially signature processes and signature business models, and then combine and orchestrate assets internally and externally, guided by a prescient strategy. Organizations must be designed for the flexibility to undertake periodic renewal and transformation.

In other words, the growth and survival of the enterprise is not just about working around market failures; it’s also about creating and implementing VRIN resources and managing complementarities to enable excellence in meeting (and sometimes even modifying) market demand in ways that are hard for competitors to imitate.

In the semi-globalized world economy, intangible assets are more apt than most physical assets to be VRIN. An organization that can bring good strategy and strong dynamic capabilities to intangible assets is likely to have durable competitive advantage.
Acknowledgement

I wish to thank Greg Linden for comments and help in editing.

References


