Market Frictions, Governance and Economic Rents: Taking Stock and Looking Ahead

Joseph T. Mahoney
University of Illinois at Urbana−Champaign, College of Business

Lihong Qian
University of Illinois at Urbana−Champaign, College of Business

Abstract

This paper develops a more unified organizational economics theory within Strategy. We begin with perfectly competitive markets derived from the first fundamental welfare theorem of economics, and develop a parsimonious typology of market frictions. We show how two primary questions in Strategy — why firms exist and why some firms outperform others — can be evaluated from this market frictions logic. Building on this logic enables more systematic explanations and predictions concerning governance structures and economic rents in Strategy research.
Market Frictions, Governance and Economic Rents: Taking Stock and Looking Ahead

Joseph T. Mahoney
Investors in Business Education Professor of Strategy, &
Director of Graduate Studies,
Department of Business Administration
College of Business
University of Illinois at Urbana-Champaign
140C Wohlers Hall
1206 South Sixth Street
Champaign, IL 61820
josephm@illinois.edu

Lihong Qian
Department of Business Administration
College of Business
University of Illinois at Urbana-Champaign
350 Wohlers Hall
1206 South Sixth Street
Champaign, IL 61820
(812) 391-0206
lqian3@illinois.edu
Abstract

This paper develops a more unified organizational economics theory within Strategy. We begin with perfectly competitive markets derived from the first fundamental welfare theorem of economics, and develop a parsimonious typology of market frictions. We show how two primary questions in Strategy -- why firms exist and why some firms outperform others -- can be evaluated from this market frictions logic. Building on this logic enables more systematic explanations and predictions concerning governance structures and economic rents in Strategy research.

Keywords: Market frictions; organizational economics; transaction costs theory; resource-based approach
Introduction

To address two primary questions in the strategic management field -- why firms exist (the boundary of the firm) and why some firms outperform others (as measured by economic rents and competitive advantage), the extant research literature provides various explanations derived from several theoretical perspectives and base disciplines. This paper, however, remains focused on contributions from organizational economics (Barney & Ouchi 1986; Mahoney 2005). Transaction costs theory regards the transaction as the unit of analysis to examine how to better achieve firm-level cost minimization through organizational boundary choices, including make-or-buy decisions and governance structure designs (Williamson 1993). Property rights theory emphasizes the concepts of ownership as residual claimancy (Alchian & Demsetz 1972) and as residual control rights (Grossman & Hart 1986) as well as the implications of property rights partitioning for the economic value creation of resources (Kim & Mahoney 2005). Agency theory characterizes the construct of the firm as a nexus of complete contracts (Holmstrom 1979; Jensen & Meckling 1976) and focuses on the economic incentive implications of various contractual specifications under different contexts, including shareholder-manager and employer-employee relationships. The resource-based approach emphasizes the search for economic rents and sustainable competitive advantages, treating (the bundle of) resources as the unit of analysis (Amit & Schoemaker 1993; Barney 1991; Peteraf 1993). The dynamic capabilities approach emphasizes individual skills, core competencies and organizational routines (Nelson & Winter 1982; Teece, Pisano & Shuen 1997). Finally, real options theory incorporates inter-project and inter-temporal dimensions to organizational boundary choices as well as the search for economic value creation (Bowman & Hurry 1993; Dixit & Pindyck 1994; Kogut 1991).
This brief review above reveals at least two points that motivate the current paper. First, each organizational economics theory has evolved with its own canonical problem and specialized language, which imposes difficulties for developing a more fully-integrative paradigm within the strategic management field. Second, these research literatures are somewhat fragmented and while successful efforts have been made to better join these theories, such connections have typically been made only in a pair-wise manner, such as: (i) transaction costs theory joined with property rights theory (Chi 1994); (ii) property rights theory joined with the resource-based approach (Kim & Mahoney 2002); (iii) the resource-based approach joined with transaction costs theory (Mahoney & Pandian 1992); and (iv) transaction costs theory joined with real options theory (Folta 1998).

The current paper develops a more unified organizational economics framework within strategic management by emphasizing first-order theoretical elements that are the common building blocks of these theories. In particular, we develop a parsimonious typology of market frictions and elucidate their connections to issues of governance structure choices and economic rents, which are considered within various organizational economics theories. We show that the market frictions logic is well developed for explaining firm boundary choices, and is also fruitful in moving towards an explanation of economic rents. Toward this objective, the next section considers a parsimonious typology of market frictions, which can be applied to the canonical problems of organizational boundaries in transaction costs theory, as well as the search for economic rents in the resource-based approach.
A Parsimonious Typology of Market Frictions

The first fundamental theorem of welfare economics demonstrates that competitive equilibrium leads to efficient resource allocation and Pareto efficiency (Arrow & Hahn 1970). Assumptions contained within this fundamental theorem include: independence in consumption and in production (e.g., there are neither externalities nor inter-project and inter-temporal spillovers), perfect information, perfect rationality, and complete markets (Debreu 1959). Under the neoclassical competitive equilibrium paradigm, firms are considered as a production function and earn zero economic rents in long-run equilibrium (Arrow & Hahn 1970; Cyert, Kumar & Williamson 1993).

Relaxation of those assumptions of the first fundamental welfare theorem within neoclassical economics has been developed in the extant strategic management research literature (Mahoney 2001; Yao 1988). Specifically -- and, in sharp contrast with neoclassical economic theory -- the strategic management field posits the construct of market frictions as its fundamental premise (Mahoney 2005). Market frictions are manifested in a variety of factors such as market power (Caves & Porter 1977); indivisibilities (Scarf 1994) leading to economies of scale (Chandler 1990); economies of scope (Teece 1980); sunk costs (Baumol, Panzar & Willig, 1982); asset specificity (Williamson 1985); imperfect information (Philps 1988; Rothschild & Stiglitz 1976); incomplete markets (Arrow 1974); asymmetric information (Holmstrom 1979); (negative and positive) externalities (Greenwald & Stiglitz 1986; Katz & Shapiro 1985), and positive transaction costs (Coase 1937; Williamson 1975).

Firms that proactively leverage these market frictions are posited to be more likely to obtain economic rents. For example, the development of reputation (Dierickx & Cool 1989;
Itami & Roehl 1987) as well as product differentiation through advertising expenditures (Sutton 1992) can be deployed to reduce asymmetric information problems on the part of consumers, and to capture consumers’ attention. Without such market frictions, there is no economic need for firms to exist, and there is no entrepreneurial opportunity for firms to generate and sustain positive economic profitability (Coase 1937; Mahoney 2001). Relaxation of these assumptions suggests the possibility of economic rents, which is a central topic within strategic management research. Therefore, an emphasis on market frictions provides a foundation for theory development, and the current paper considers various organizational economics theories (listed at the beginning of this paper) derived from this foundation.

We next provide a parsimonious, but fairly comprehensive, typology of market frictions, which underpins resource-based, transaction costs, as well as other organizational economics theories, with the purpose of identifying opportunities for developing a more systematic framework concerning the firm’s existence, boundary choice, heterogeneity, and economic rents.

**Market Frictions and the Boundary of the Firm**

**Figure 1** illustrates a market frictions logic concerning governance, which builds on Williamson’s (1975) market and organizational failures framework. Though different theories are concerned with different canonical problems (e.g., transaction costs theory focuses on governance choice, while the resource-based approach focuses on economic rents), the current paper develops the proposition that market frictions are fundamental building blocks for these organizational economics theories. Analysis of market frictions associated with a particular organizational economics theory enables identification of
convergences and divergences of these theories. To explain both governance and economic rents, we expand Williamson’s (1975) seminal framework to include: (1) human factors; (2) decision-making uncertainty; and (3) organizational factors. We consider each of these three factors in turn.

First, the human factors include bounded rationality, in which decision makers are intendedly rational, but limitedly so (Simon 1947), and potential opportunism, which is characterized as self-interest seeking with guile (Williamson 1996). The combination of these two human factors are important since **bounded rationality** leads logically to contracts typically being incomplete, and **opportunism** potentially leads to substantial contractual hazards. Such hazards are especially problematic when there is **asymmetric information** that can result in adverse selection (Akerlof 1970) and moral hazard (Holmstrom 1979) problems. In addition, contractual hazards due to incomplete contracts and opportunism may be especially severe with dedicated assets and **unilateral commitments** in physical, human, and site **asset specificity** (Williamson 1985), which can lead to substantial economic hold-up problems (Klein, Crawford & Alchian 1978; Sherry & Teece 2004).

Second, in terms of decision-making uncertainty, the presence of **externalities** (Coase 1960) can result in price signals no longer being sufficient statistics for the effective allocation of resources (Libecap 1989). Somewhat analogous to the externalities problems that may be found in inter-firm exchange, the presence of intra-firm **spillovers** can lead to internal transfer pricing that provides a poor signal for the effective allocation of capital (Dixit & Pindyck 1994). Further, **uncertainty** is a central market friction. Indeed, Arrow maintains that: “There is one particular failure of the price system … one that is absolutely central to the understanding of organizations. I refer to the presence of uncertainty” (1974:
Under Knightian (1921) uncertainty, it may not be feasible to assign probabilities to important contingencies or certain contingencies may not even be considered at all, \textit{ex ante}. Therefore, markets will, by definition, be incomplete (Arrow & Hahn 1970).

Third, organizational factors include \textit{production economies and capabilities} (Argyres 1996), which can be an important component of total costs (Walker & Weber 1984). \textit{Governance inseparability} is a construct that enables us to focus on path-dependent processes (Arthur 1994). In particular, previous governance choice commitments both enable and constrain the viability of future governance decisions (Argyres & Liebeskind 1999). Finally, \textit{undefined and poorly-defined property rights} (Barzel 1989; Demsetz 1967) can substantially increase transaction costs, making exchange difficult. It can lead to high measurement costs (Barzel 1982; Ouchi 1979) and resources being non-tradeable (Dierickx & Cool 1989).

Now that we have considered the constructs in Figure 1, we next analyze the linkages and inter-relationships of these constructs for the purpose of better explaining and predicting governance choices. Therefore, we consider, in turn, linkages from A to I in Figure 1 and their theoretical foundations.

**Linkage A: Uncertainty, Asymmetric Information and Opportunism.** A combination of asymmetric information and \textit{ex post} opportunism results in market frictions, and a representative example is the insurance problem, where insurance influences the behavior of the insured and can lead to moral hazard problems (Holmström 1979). A classic example that focuses on the consequences of asymmetric information and \textit{ex ante} opportunism is Akerlof’s (1970) “market for lemons,” where adverse selection recursively drives out suppliers of higher-than-average-quality used cars, and, in equilibrium -- without the aid of governance mechanisms, such as warranties for quality assurance -- the market completely
collapses. The concept of incomplete markets can be applied to a wide range of phenomena where asymmetric information exists (Arrow 1974).

Uncertainty, and asymmetric information combined with opportunistic behavior between borrowers and lenders can lead to incomplete and imperfect capital markets (Williamson 1975). Asymmetric information and opportunistic behavior have clearly played important roles in recent Wall Street developments. For example, in September 2008, Lehman Brothers, the fourth largest investment bank in the United States filed for Chapter 11 bankruptcy protection, ending its 158-year history. One explanation for Lehman Brother’s bankruptcy and later liquidation was its heavy involvement in financial derivatives, especially involving home mortgages and their subsequent collateralized debt obligations. In this process, a number of financial institutions participated in various activities such as borrowing and lending, packaging and re-packaging, securitizing, credit rating, and selling. Such a complex process entailing high degrees of specialization exacerbates the problem of information asymmetry along the chain of financial derivatives innovation, and imposes substantial difficulty in valuing and pricing the financial products being transacted. The fact that some lending is, indeed, sub-prime can be concealed as more layers of financial products are derived from it. Such severe information asymmetry problems deteriorate the ability of those financial institutions to respond effectively. In addition to dramatic market failures, government failure is also responsible since it did not meet its obligations for due diligence in its role of monitoring and auditing. Such government shortcomings contain their own elements of opportunism and asymmetric information problems. In short, the real-world economy provides us with profound examples of the consequences of organizational
and market failures due to a combination of uncertainty, asymmetric information, and opportunism (e.g., moral hazard problems).

**Linkage B: Bounded Rationality and Uncertainty.** The joint effects of bounded rationality and uncertainty result in incomplete contracting, as certain contingencies cannot be fully specified *ex ante* (Arrow 1974; Grossman & Hart 1986). Due to bounded rationality, contractual parties are not even remotely capable of foreseeing all relevant contingencies that may be involved in implementing a complex contract (Simon 1947; Williamson 1975), which results in market imperfections. Further, even with the aid of advanced computing technologies, substantial environmental uncertainty still results in contract incompleteness (Lajili & Mahoney 2006; Powell & Dent-Micallef 1997).

**Linkage C: Bounded Rationality and Opportunism.** Incomplete contracting and subsequent contractual hazards is a derivative condition of bounded rationality and opportunism. One governance solution to mitigate the potential contractual hazards associated with incomplete contracting is internalization. For example, a firm’s multidivisional internal capital market can attenuate efficiency losses due to asymmetric information (Williamson 1975). The multidivisional form is able to economize on bounded rationality through the loose coupling of tactical and strategic decisions (Chandler 1962), and is able to mitigate opportunistic behavior since it has better incentive alignment features because: (a) *property rights ownership and incentives change*, since each division has no preemptive claims on cash-flow streams; (b) *monitoring changes* as internal auditing is typically more effective than external auditing; and (c) *rewards change* since the multidivisional form can provide more fine-grained monetary and non-monetary rewards (Mahoney 1992; Williamson 1985).
**Linkage D: Opportunism and Asset Specificity.** Opportunism and asset specificity are two key constructs in transaction costs theory (Williamson 2003). The potential for opportunistic behavior by one of the exchange partners poses possible threats to the other exchange partner, especially when the latter has incurred substantial asset-specific investments. With a concern for the economic hold-up problem from the other exchange partner, internalization within the firm may be a better governance choice (Grossman & Hart 1986; Klein, Crawford & Alchian 1978). The extant empirical literature largely corroborates the prediction that potentially high market frictions due to opportunism and asset specificity result in economic safeguards such as internalization to mitigate these contractual hazards (De Figuerido & Teece 1996; Lajili, Madunic & Mahoney 2007; Sherry & Teece 2004).

**Linkage E: Asset Specificity and Externalities.** In general, externalities and positive transaction costs have been a concern in the property rights research literature (Coase 1960; North 1990). For our purpose here, we focus on a business context in which market frictions are likely to be critical, namely when there is high asset specificity (Williamson 1985). As an example, Kim and Mahoney (2002) examine oil field unitization to explicate the effects of investments characterized by high levels of asset specificity, ill-defined property rights, (negative) externalities and positive transaction costs. In particular, the negative externalities of oil migration towards those oil firms that drill early led to over-drilling at poor location choices. Consequently, a substantially sub-optimal aggregate amount of oil extraction was experienced for decades in Oklahoma and Texas (Libecap 1989).

**Linkage F: Asset Specificity and Spillovers:** The very same economic problem that may occur due to market frictions in inter-firm exchange involving asset specificity and externalities (Libecap 1989) can also occur with intra-firm exchange involving asset
specificity and spillovers (Kang, Mahoney & Tan 2009). For example, while a “strategic” outsourcing of workers may maximize the economic returns for a particular division, the benefits these workers provided in consultation with people in other divisions of the same company may be lost, in which case the economic returns to the enterprise can be negative. Another example of a substantial positive intra-firm spillover is R&D, whose knowledge creation may have positive spillover applications to other divisions within the company. However, due to potential multidivisional negotiating problems involving investments with high-levels of asset specificity, these potential positive spillovers may not be fully realized. In other words, the marginal benefits of R&D investment for a particular division may be smaller than the marginal costs for that particular division, in which case the investment might not occur despite the fact that the total net impact of the R&D investment at the firm-level would be positive when the full costs and benefits are accounted for in the decision calculus. The real options logic explicitly attempts to account for these positive spillovers within the construct of expansion and growth options (Dixit & Pindyck 1994).

Linkage G: Spillovers and Production Economies & Capabilities. Spillovers can occur with production capabilities that involve economies of scale and/or scope (Teece 1982). We emphasize here -- along with Williamson (1975: 83-84) and Teece (1980) -- that economies of scope, per se, only explain the co-location of resources. It does not, by itself, explain or predict governance choice, which is determined by the extent of market frictions. Thus, more refined theory building makes clear that when economies of scope are combined with spillovers and market frictions, the governance choice of internalization may be needed to achieve more fully the benefits of production economies vis-à-vis markets.
**Linkage II:** *Spillovers* and *Governance Inseparability:* Argyres and Liebeskind (1999) advanced theory by incorporating a time dimension into transaction costs economics to explain governance choice from a dynamic perspective. In particular, this line of reasoning maintains that governance choices, such as previous contractual commitments entailing sunk costs may enable or constrain a firm’s subsequent governance choices. This logic resonates with the emerging research literature on dynamic transaction costs theory (Langlois 1992; Spiller & Zelner 1997). An example for positive spillovers occurs when learning by doing is involved with contracting experience, and thus the decision to contract today may positively influence subsequent contracting capabilities (Mayer & Argyres 2004). An example of negative spillovers would be if internal procurement, internal expansion, and program persistence biases exist in the vertically-integrated firm (Williamson 1975: 118-122), and thus these internal organizational distortions that would constrain future governance choices should be included at the outset in the decision calculus.

**Linkage I:** *Capabilities* and *Un-defined and Poorly Defined Property Rights.* First, as Coase (1960) makes clear, an initial assignment of property rights facilitates subsequent trade. If property rights are undefined, markets may completely break down (North 1990). Further, if the initial property rights are poorly defined, the market friction consequences can still be substantial. Consider a multidivisional firm with highly-related business units where residual income to managers is based only on divisional ROA. Such property rights within this firm are poorly defined since the managers of these highly-related businesses are unlikely to share information under this particular compensation system. Hence, with this lack of knowledge transfer, the realized capabilities of such a firm will be far from the potential capabilities of this firm. A change in the property rights regime of the firm, where
divisional managers are rewarded based on corporate ROE, would almost surely improve information sharing within the firm, thereby facilitating greater capability-development. This example illustrates the more general principle that property rights and dynamic capabilities are tightly linked (Mahoney 2005; Wang & Barney 2006). Indeed, the next generation of resource-based and dynamic capabilities research should join the governance (and property rights) approach with the dynamic capabilities approach (Makadok 2001).

In summary, Figure 1 illustrates that the current paper’s developed market frictions logic, which joins agency theory, property rights theory, dynamic capabilities & real options theory, and transaction costs theory more fully explains and predicts governance structure choice. A question that naturally follows is: can this market frictions logic explain and predict economic rents to the same extent as it does for the governance structure question? We turn to this important question within our Strategy field in the following section.

**Market Frictions and Economic Rents of Firms**

In Figure 2, we outline several fundamental questions in the strategic management field, and embed these questions on a foundation of market frictions. The question of why firms exist corresponds to the economic analysis concerning the boundary choice of firms, and the questions of firm heterogeneity and performance differentials correspond to inquiry concerning the economic rents of firms. We submit that:

- The questions of (a) firm boundary choice; (b) firm resource heterogeneity; and (c) firm-level economic rents cannot be fully addressed by just one organizational economics theory;

- These three fundamental questions within the strategic management field share a common foundational antecedent: *market frictions*; and

- The market frictions that explain economic rents are sufficient market frictions to explain why firms exist.
Indeed, Mahoney (2001) maintains that market frictions are the critical concepts to both resource-based and transaction costs theory, and further posits that the “set of market frictions that explain sustainable firm-level rents would be sufficient market frictions to explain the existence of the firm” (2001: 655). The current paper advances this proposition by considering: What are these market frictions that can explain sustainable firm-level rents? To address this research question, we build upon several important works from the extant resource-based literature (Amit & Schoemaker 1993; Barney 1991; Grant 1991; Peteraf 1993); identify key constructs of this approach; connect and build upon these constructs with our developed market frictions logic; and suggest areas where the resource-based approach can be advanced further.

Barney (1991) suggests that the search for sources of sustainable competitive advantage focus on resources that are in demand (i.e., valuable), rare, inimitable and non-substitutable, the resource-based VRIN criteria, which constitute necessary conditions for economic rents. Grant (1991) traces the profit-earning potential of a resource or capability to: the extent and sustainability of the competitive advantage established through resources and capabilities, and the ability of the firm to appropriate the economic rents generated from its resources and capabilities. Amit and Schoemaker (1993) describe desired characteristics of a firm’s resources and capabilities that are posited to generate economic rents, such as scarcity, complementarity, and appropriability. And, Peteraf (1993) establishes the four cornerstones of competitive advantage: (a) superior resources; (b) ex post limits to competition; (c) imperfect resource mobility; and (d) ex ante limits to competition.

Figure 3 provides a synthesis of the various conditions derived from these seminal resource-based articles in the strategic management field in terms of (a) the key conditions
that help explain and predict the extent of economic rents that can be established through resources and capabilities; (b) the primary conditions that facilitate the sustainability of these competitive advantages; and (c) the fundamental conditions that enable appropriation of these economic rents.

Here, we join the current paper’s developed market frictions logic concerning governance choice in Figure 1 with the resource-based logic outlined in Figure 3, by further explicating the connectedness of our market frictions logic across all constructs contained in Figure 3. In particular, in Figure 3, we connect each resource-based construct with one or more market friction constructs.

(a) Extent of competitive advantages of resources

For a resource to create firm-level competitive advantage that resource in itself, or in combination with other resources in the firm, must satisfy several conditions (e.g., the resource is valuable and rare) and firms can differ in either getting access to, or figuring out ways of using these resources. We maintain that one fundamental explanation for such a difference is due to market frictions.

Scarcity/rare: Resources are rare if supply is limited, and expansion in supply cannot be easily achieved. Furthermore, resources may be scarcer for some firms than for others if the latter have better information to access those scarce resources. Peteraf (1993) maintains that the possession of superior resources by some firms within an industry can result from asymmetric information about the economic value of those resources when the firms started to develop them. A stronger form of resource scarcity is uniqueness of resources, e.g., resources that are created by a firm and henceforth are specific to that firm. Imperfect
information concerning availability of resources or the way to create resources reinforces their scarcity.

Relevance/value: A resource is relevant/valuable if there is demand for that resource or for products derived from it. Barney (1991: 106) regards a resource to have value when it enables a firm to conceive of or implement strategies that improve its efficiency or effectiveness. Peteraf and Barney (2003) define the “value” of a resource as the perceived benefits from a resource above its economic costs. Resource-based logic suggests that obtaining such a valuable resource can occur through luck and/or asymmetric information (Barney 1986).

Heterogeneity: Firms consist of resources that are not commonly employed by all firms in an industry. Heterogeneity is typically assumed in the resource-based approach (Peteraf & Barney 2003). But where does such heterogeneity come from? The key market friction is asymmetric information about resource attributes and availability. Further, when considering flow-and-stock dynamics (Dierickx & Cool 1989), information asymmetry combined with heterogeneity in a resource stock formed at time \( t \) influences firms’ future strategies in resource procurement, deployment and divestiture, and therefore, over time, the heterogeneity between firms may increase. Thus, the presence of firm-level heterogeneity implies the potential presence of various market frictions, such as asymmetric information, uniqueness/asset specificity and bounded rationality/causal ambiguity (Lippman & Rumelt 1982; Mahoney & Pandian 1992).

Complementarity: Amit and Schoemaker (1993) explain complementarity as a relationship between resources where the relative magnitude of strategic value of one resource may increase with an increase in the relative magnitude of other resources (Milgrom & Roberts 1990). Such complementarity can occur through positive inter-project
spillovers, such as economies of scope (Teece 1980) and through positive inter-temporal spillovers, such as asset accumulation or interconnectedness over time (Dierickx & Cool 1989). This resource complementarity can increase the difficulty in transferring, imitation or substitution of any single resource.

(b) Sustainability of competitive advantages

By constraining competition, a firm can increase sustainability of its economic rents. Therefore, impediments to economic activities (Yao 1988) or market frictions are especially relevant for sustainability of competitive advantage.

*Low transferability:* Low transferability of resources is an explicit manifestation of market frictions. Imperfect resource mobility can result from non-tradable or co-specialized assets. A resource may be non-tradable because of: (i) difficulties in pricing the resource due to information asymmetry between seller and buyer (e.g., brand image); (ii) impediments in transferring due to tacitness and intangibility (e.g., R&D capabilities); (iii) difficulties in partitioning (e.g., team-embodied skills); and (iv) isolating mechanisms such as patent and trademark enforcement (Rumelt 1984). Moreover, even if a resource is tradable, the potential economic rents from this resource will be attenuated if it is separated from its co-specialized assets (Teece 1986).

*Inimitability:* Inimitability is intrinsically related to low transferability, to the extent that some causes of low transferability such as non-tradable or co-specialized resources can lead to inimitability. Barney (1991) attributes inimitability to path-dependence, social complexity, and causal ambiguity (Lippman & Rumelt 1982), which can be derived from market frictions of externalities, spillovers, bounded rationality, and asymmetric information.
**Non-substitutability:** Substitutability or non-substitutability is a relative term, since for any resource there can be another resource that at least partially substitutes for the focal one, either cross-sectionally or longitudinally (Barney 1991). Therefore, there is often imperfect substitutability between firm resources. The resource-based construct of non-substitutability between resources is strategically important when the resources are valuable, rare, and inimitable. Moreover, non-substitutability and inimitability are interrelated. Indeed, Barney (1991) suggests that substitutability can take at least two forms, one form is that a firm can substitute a similar resource that enables it to conceive of and implement the same strategy, and the other form is that different firm resources are strategically equivalent if they achieve the same functional outcomes. These two forms of substitutability are similar to the resource-based construct of imitability of a resource, where substitutability can also be the imitation of a process, and thus can be a dynamic form of imitability. By connecting the conditions of inimitability and non-substitutability, market friction explanations for inimitability of firm resources are also applicable here.

**Durability:** If a strategic resource is not durable (e.g., erosion of physical assets) then competitive advantage based upon this resource cannot be sustainable (Dierickx & Cool 1989; Grant 1991). A focus on the criterion of resource durability leads to the issue of how a firm might be able to implement strategies to make its economic rents more durable. Concentrating on this strategic issue leads to more detailed consideration of market friction impediments to appropriation of economic rents, which we turn to next.

(c) **Appropriability of economic rents**

Appropriation of economic rents is a necessary condition for economic value capture and sustainability. Moreover, low appropriability of economic rents lowers the incentives to
invest in resources, activities, and dynamic capabilities in the value-creation process. The following conditions influence appropriability of economic rents created by a firm.

**Complementarity:** The possession of a complementary resource by a firm can lead to it appropriating economic rents. If the firm possesses a co-specialized resource (Teece 1986) then there is idiosyncratic bilateral synergy (Mahoney & Pandian 1992: 368) in which each firm possessing a co-specialized resource is expected to appropriate *some* of the gains from these joined resources. Even more dramatically, if one of the resources in the combination is held by a firm in a competitive factor market while the other firm possesses a valuable, rare (unique), inimitable, and non-substitutable resource, then resource-based logic makes clear that the firm possessing the unique resource will appropriate the *entire* economic value created in this resource combination (Barney 1986, 1991). For the purpose of the current paper, we emphasize that the market frictions, which result from *asset specificity*, not only help explain why these co-specialized resources are owned by one firm (Teece 1986) but also why co-specialized resources lead to economic rents (Mahoney 2001).

**Property rights:** Well-defined property rights regimes (e.g., patent and trademark protection, and equity share arrangements) provide direct mechanisms for appropriability. In contrast, undefined and poorly-defined property rights are fundamental sources of market frictions that create misleading price signals and therefore reduce efficiency in resource allocations (Coase 1960). The implications of property rights regimes for economic rents will be considered in the next section of the current paper.

**Organizational governance and incentives:** The organizational design of governance structures and incentive systems are closely related to property rights considerations. Assignment of residual claimancy (Alchian & Demsetz 1972) and residual rights of control
(Grossman & Hart 1986) are important components of organization design. For example, Grossman and Hart (1986) show that imperfections in residual control assignments (e.g., the contractual party having the greatest potential for increasing aggregate gains is not assigned residual control rights) can lead to poor resource allocation within the firm. In particular, if the contractual party contemplating firm-specific investment *ex ante* is not afforded residual control rights *ex post*, then a farsighted decision-maker will foresee the contractual hazards of having at least part of their investment appropriated, and consequently there will likely be an underinvestment in firm-specific capabilities (Wang & Barney 2006).

*Embeddedness:* Embeddedness entails both cross-sectional and longitudinal interconnectedness of resources (Dierickx & Cool, 1989). Consider the example of a university professor of chemistry who is tied to a university-owned research laboratory vis-à-vis a business school professor who has little ties to the physical resources. The embeddedness argument maintains that ownership of the physical asset leads to control over human resources (Grant 1991; Hart 1995). In particular, the owner of the physical resource (i.e., the university) is in the position to appropriate more of the economic value created by the professor of chemistry than the relatively more mobile professor of business.

In summary, the evolving resource-based approach within the discipline of strategic management has developed important constructs to explain the extent and sustainability of competitive advantage as well as the appropriability of economic rents. Key contributions of the current paper are to join several important constructs from various sources within the extant research literature, and to provide the foundational market friction building blocks that underpin this resource-based approach.
Look Ahead: A Theory of Economic Rents

Now that we have shown that market frictions provide the theoretical foundations for explaining and predicting both the firm boundary and economic rent questions, we consider next steps for the resource-based approach. Currently, the market frictions logic is better developed for the firm boundary question than for explaining firm economic rents. The current paper suggests, however, that more systematic theory development that utilizes the market frictions logic appears promising.

Table 1 provides a brief comparison of extant research literature on firm boundary and economic rent questions derived from transaction costs and resource-based approaches. We organize this comparison along with key market friction constructs, and select important research from these two approaches. The evidence provided in this Table indicates that market frictions logic is being used in both approaches but to different degrees in at least two senses of the term. First, transaction costs theory has progressed by more self-consciously and explicitly using the extant market failures literature (Williamson 1975, 1996), while the pattern of market frictions logic within the resource-based approach has been somewhat more emergent and implicit. We submit here, however, that we will now progress more rapidly through a systematic study of market frictions within our Strategy field. Second, several market friction constructs have been developed with a depth of theoretical and empirical research in transaction costs theory (Mahoney 2005; Williamson 1996) while the corresponding resource-based approach is currently less developed. For example, a potentially important market friction for future research in the resource-based approach is externalities, as well as the counterpart construct of firm-level spillovers. Such
an inquiry may require an examination of real options, dynamic capabilities, and dynamic transaction cost approaches (Ghemawat 1993; Mahoney 2005).

We propose here some paths in which the resource-based approach can be developed into a richer theory of economic rents and sustainable competitive advantage. First, a theory of rents must address not only potential economic value creation, but also realized value creation (Kim & Mahoney 2002). The resource-based logic in Barney (1991) and Peteraf (1993) offer key criteria for potential economic rents that can be derived from (a bundle of) resources. However, without well-defined property rights, including residual control rights and residual claimancy, the economic value creation potential of a certain resource or resource bundle may not be realized in full as predicted by the resource-based approach. A more systematic examination of market frictions in general and property rights (governance) in particular is required to ascertain the economic value of resources and capabilities (Makadok 2001; Mahoney 2005).

Second, a theory of economic rents must incorporate (inter-divisional and inter-temporal) spillover effects. Focusing on the transaction as the unit of analysis neglects spillover effects that can occur across multiple transactions (Argyres & Liebeskind 1999). Analogously, utilizing the individual resource as the unit of analysis will have similar limitations since it may neglect important spillover costs and benefits to the bundle of resources that the decision-maker either currently possesses or may possess over time. A richer resource-based analysis of economic value creation will therefore be achieved with greater attention to inclusion of the spillover effects of resources in the decision calculus.

Third, related to the concept of spillover effects, the unit of analysis in a theory of economic rents should not be confined to a single resource or even to a bundle of resources.
For example, Porter and Siggelkow (2008) review research studies that focus on interactions among a system of activities of a firm, and suggest effects of the interactions on a firm’s sustainable competitive advantage. This line of research literature on activities (Kauffman 1993; Milgrom & Roberts 1990) highlights a view of an activity system, which, we maintain, can be the unit of analysis for developing a theory of economic rents. An activity system view is complementary to the concept of resource (re)bundling processes in the resource-based approach (Penrose 1959), in that the activity system approach emphasizes the complementarities among activities, while the resource-based approach emphasizes the complementarities among resources. The activity system approach can be applied not only to the firm’s boundary choice, but also to the search for economic rents.

Fourth, in terms of econometric analysis, the challenging issues for testing transaction costs theory apply equally to testing the resource-based approach. Indeed, the fundamental premise within the strategic management field is that managerial decision-making in terms of governance decisions and the strategic choice of resource bundles are endogenous to their expected performance outcomes (Hamilton & Nickerson 2003). Therefore, more sophisticated econometric techniques (e.g., instrumental variable and two- and three-staged methods) are required to account for self-selection problems and unobserved heterogeneity (Heckman 1979; Shaver 1998). Failure to account for endogeneity can have important consequences. For example, in transaction costs theory, Masten (1993) shows that within the context of make-or-buy decisions in the shipbuilding industry, the obtained OLS estimates are approximately 30% lower for the average make component and almost 50% lower for the typical buy component. Thus, the uncorrected model provides biased estimates of the costs of governance choice mistakes. In applying these econometric issues to the resource-
based approach, empirical analysis that relates R&D intensity directly to economic performance may provide biased coefficients if important variables such as firm-level resources and capabilities, which are correlated with R&D intensity, are omitted from the model specification.

Finally, developing the market frictions logic within the resource-based approach is promising for providing practical implications for strategic managers. Indeed, such decision makers -- in parallel with researchers -- can proactively consider market frictions that enable cost minimization and the generation of firm-level economic rents. For example, managers can provide economic value in business situations where market frictions exist and therefore price signals are not perfect. Thus, the correct managerial calculus will include the spillover benefits and costs that are not typically priced in standard discount cash flow formulations. Such strategic connections show that the market frictions framework developed here enables us to better join Strategy theory and strategic management practice.

These are challenging issues for further inquiry within organizational economics in general and for the resource-based approach in particular. It is anticipated here that a more systematic examination of market frictions for the study of firm boundary, firm heterogeneity and firm-level economic rents will enable the next generation of research within our evolving field of Strategy to do even better.
References


Figure 1: A Market Frictions Framework of Governance

Human Factors:

1. Bounded Rationality (Causal Ambiguity)

Decision-Making Uncertainty:

2. Opportunism /Self-interest-seeking

Organizational Factors:

3. Asymmetric / Imperfect Information
   - Moral Hazard
   - Adverse Selection

4. Asset Specificity (Uniqueness)
   - Hold-up problem

5. Externalities

6. Spillovers

7. Uncertainty

8. Production Economies/ Capabilities

9. Governance Inseparability

10. Un-defined / Poorly Defined property Rights

Figure 2: Market Frictions and Theoretical Questions

(a) Why do firms exist?

(b) Why do firms differ?

(c) Why do firms perform differently?

As severity of the market friction increases

Market frictions foundation
Figure 3: Economic Rents and Sustainable Competitive Advantages

<table>
<thead>
<tr>
<th>(a) The extent of competitive advantage</th>
<th>(b) The sustainability of competitive advantage</th>
<th>(c) The appropriability of economic rents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Scarcity / Rare</td>
<td>5. Low Transferability</td>
<td>9. Complementarity</td>
</tr>
</tbody>
</table>

**Market Frictions:**

- Imperfect information
- Asymmetric information
- Asset Specificity/Uniqueness
- Externalities combined with positive transaction costs
  - Spillovers
  - Uncertainty
- Indivisibilities (production economies of scale)
- Undefined and Poorly-Defined Property Rights
### Table 1: Market Frictions in Transaction Costs and Resource-based Approach

<table>
<thead>
<tr>
<th>1st Order Constructs</th>
<th>Derived Constructs</th>
<th>Transaction Costs Theory</th>
<th>Resource-Based Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounded Rationality</td>
<td>Causal Ambiguity</td>
<td>• Simon (1947)</td>
<td>• Lippman &amp; Rumelt (1982)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Williamson (1975)</td>
<td>• Mahoney &amp; Pandian (1992)</td>
</tr>
<tr>
<td>Opportunism</td>
<td>Opportunistic behavior</td>
<td>• Parkhe (1993)</td>
<td>• Foss (1996)</td>
</tr>
<tr>
<td></td>
<td>Trust/ Relational Experience</td>
<td>• Chiles &amp; McMackin (1996)</td>
<td>• Conner &amp; Prahalad (1996)</td>
</tr>
<tr>
<td></td>
<td>Ex Ante Adverse Selection</td>
<td>• Chi (1994)</td>
<td>• Kim &amp; Mahoney (2006)</td>
</tr>
<tr>
<td></td>
<td>Ex Post Moral Hazard</td>
<td>• Chi (1994)</td>
<td>• Foss &amp; Foss (2005)</td>
</tr>
<tr>
<td>Production Economies</td>
<td>Economies of Scale (Indivisibilities)</td>
<td>• Arrow &amp; Hahn (1970)</td>
<td>• Lieberman &amp; Montgomery (1988)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
<td>------------------</td>
</tr>
<tr>
<td><strong>Uncertainty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Complexity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Asymmetric Information</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Externalities</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Spillovers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>