Dispersed knowledge and an entrepreneurial theory of the firm

Nicholas Dew¹,*, S. Ramakrishna Velamuri, Sankaran Venkataraman

¹Graduate School of Business and Public Policy, Naval Postgraduate School, Monterey, CA 93943, USA
²IESE Business School, Avenida Pearson 21, 08034 Barcelona, Spain
³Darden Graduate School of Business, University of Virginia, Charlottesville, VA 22901, USA

Abstract

In this article, we propose an entrepreneurial theory of the firm that is based on dispersed knowledge. We argue that the dispersion of knowledge over people and places and over time leads to uncertainty. This uncertainty, combined with heterogeneous expectations and the nexus of an individual and opportunity, explains the emergence of new firms. We then suggest that the theory of the firm proposed by us answers questions that have been overlooked by alternative theories. The specific question we discuss in this article is when and why an entrepreneurial opportunity will be taken to market through an existing firm, and when and why a new firm will be chosen as a vehicle for taking a new idea to market, i.e., whether the residual will be concentrated in an existing or in a new firm.

© 2003 Published by Elsevier Inc.

Keywords: Dispersed knowledge; Entrepreneurial theory; Uncertainty

1. Executive summary

Ever since Ronald Coase posed the question in 1937—“Why do firms exist?”—several explanations have been proffered for why, given markets, there is a need for the institutional modality of the firm. All the explanations that have hitherto been given have one common feature: they explain the rationale for the existence of firms in terms of the need to economize on costs, such as transaction costs, agency costs, and supervision and monitoring costs. In our
view, these explanations erroneously assume that a given transaction can be carried out either in a market or in a firm, and that it is only the cost saving that determines where it will in effect be carried out.

We offer an alternative explanation by arguing that the dispersion of economic knowledge over people, places, and over time leads to genuine uncertainty. A firm is created when an entrepreneur recognizes an opportunity in an environment of genuine uncertainty and is unable to sell this opportunity in the market. She then creates a firm by ensuring fixed payments to other parties such as employees and suppliers and by keeping for herself the residual rights. Our theory has three unique features: first, it traces the genesis of firms to the recognition of an entrepreneurial opportunity; second, unlike other theories that emphasize transactions and agency relationships, our theory places the individual at the center of the explanation of why firms exist; finally, it addresses the critical question of why particular firms exist.

2. Introduction

The peculiar character of the problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form but solely as the dispersed bits of incomplete and frequently contradictory knowledge, which all the separate individuals possess. (Hayek, 1945, p. 519).

Since the publication in 1945 of Hayek’s seminal essay, interest in the relationship between the knowledge possessed by economic agents and the process of recognition, pursuit, and exploitation of entrepreneurial opportunities has grown (Venkataraman, 1997; Shane, 2000). In this article, we explore the implications of dispersed knowledge for an entrepreneurial theory of the firm. Specifically, we seek answers to the question of why particular new firms are created, that is—why are some new ideas brought to market through existing firms, and others through new firms? We suggest that the answer to this question is to be found in the dispersed nature of knowledge.

The core of the theory presented here is that three phenomena arise from the dispersion of knowledge that are critical to an entrepreneurial theory of the firm. The first is that the dispersion of knowledge gives rise to genuine uncertainty, which necessitates the contractual structure that we recognize as a firm. The second is that dispersion of knowledge and genuine uncertainty contribute to the heterogeneity of expectations that must exist in order for one or more individuals to exploit the potential of the contractual structure of the firm. The

---

1 When we refer to a firm, we take it to possess the six characteristics stipulated by Alchian and Demsetz (1972, p. 794): (1) joint input production, (2) several input owners, (3) one party who is common to all the contracts of the joint inputs, (4) who has rights to renegotiate any input’s contracts independently of contracts with other input owners, (5) who holds the residual claim, and (6) who has the right to sell his central contractual residual status.
third is that dispersion of knowledge, genuine uncertainty, and heterogeneous expectations give rise to the nexus of the enterprising individual and the opportunity to discover, create, and exploit new markets. These three elements together explain why particular new firms are created. This theory is summarized in Fig. 1. As depicted in the figure, dispersion of knowledge (Hayek, 1945) serves as our starting point. We suggest that a dispersion of

![Diagram](image)

**Dispersion of Knowledge (Hayek, 1945)**

- **People**
- **Places**
- **Time**

**Cross-sectional Uncertainty (Akerlof, 1970)**

- A new nexus of contracted relationships is required (Knight, 1921)
- Need for a firm arises (Coase 1932; Knight 1921) because of the need for specific + residual contracts (Knight 1921; Grossman & Hart 1986)

**Longitudinal Uncertainty (Knight, 1921; Arrow, 1974)**

- Gives rise to heterogeneous expectations and nexus of an opportunity for a future product-market & an individual: i.e., a match is created in

**Create nexus through existing firms in the economy**

**Create nexus using a new firm in the economy**

New firm comes into existence when transaction costs are moot and incentives are unalignable.

Fig. 1. Dispersion of knowledge (Hayek, 1945).
knowledge across people and places leads to information asymmetry of the Akerlofian variety (Akerlof, 1970). Akerlofian information asymmetry, in combination with dispersion of knowledge over time, leads to Knightian uncertainty (Knight, 1921), and forms the core element of the entrepreneurial theory of the firm that we develop in this essay. We argue that dispersed knowledge, along with Knightian uncertainty, contributes to economic agents having heterogeneous expectations. These heterogeneous expectations, in turn, lead to the nexus of an individual and an opportunity. Exploiting this opportunity requires the unique contractual structure of the firm and leads to a choice between the two alternative modes of an existing or a new firm. The choice of the mode of exploiting the opportunity depends on whether the opportunity is discovered or created inside an existing organization or outside of it, whether any exclusive and defensible property rights (such as patents) exist over the opportunity, and the heterogeneous expectations of the parties relevant to the exploitation of the opportunity.2–4

This article proceeds as follows. We start by raising a new question relevant to entrepreneurship that traditional theories of the firm leave unaddressed. Next, we analyze the different ways in which knowledge is dispersed. We then argue that the dispersion of knowledge has an important role in three phenomena: genuine uncertainty, heterogeneous expectations, and the nexus of the individual and the opportunity. We then describe why these phenomena result in the distinctive contractual structure of the firm and why entrepreneurs create new products and services sometimes through existing firms and sometimes by creating new firms. We end by contrasting the theory offered here with some existing theories of the firm and offer some conclusions about what the theory developed here may have to offer researchers.

The theory developed here combines several perspectives on the theory of the firm to create a theory of the entrepreneurial firm that is relevant, meaningful, and fitting in the context of entrepreneurship. The theory can answer, with no loss of parsimony, the three central questions addressed by existing theories: Why do firms exist? What are the determinants of their scale and scope? Why do some firms persist over time? Second, it contributes a distinctly entrepreneurial perspective on the theory of the firm by addressing the issue of when and why new firms are created by entrepreneurs, i.e., why some entrepreneurial opportunities are pursued through new firms instead of through existing firms.

3. Background—new questions

Entrepreneurship—the process by which opportunities for the creation of markets for new products and services are recognized, pursued, and exploited (Venkataraman, 1997)—is a

2 We suggest that it is meaningful to talk of individuals, as opposed to firms, as the recognizers of opportunities.

3 Our discussion assumes that the property rights to the opportunity are held by the agent. If they are held by the principal, then the opportunity will be pursued through the existing firm.

4 A new firm that is created by the principals of the existing firm is considered an existing firm by us. The difference between a new and an existing firm is based on who the principals are.
phenomenon that is apparent to any casual observer of economic activity. Indeed, some of the most critically acclaimed economic theorists have suggested that the creation of new product markets is at the very core of economic activity (for instance, Schumpeter, 1976). It seems reasonable to make at least two conjectures about the genesis of new product markets (Venkataraman and Sarasvathy, 2001). First, of all the product possibilities that could come into existence from a new technology, only a subset will actually be created and brought into existence within a finite period. The history of invention suggests that long lead times exist between invention and the production of novel economic artifacts, that a slew of economic failures often precedes the eventual successful application of technologies (Wiener, 1996), and that many product possibilities languish on research benches and never result in new product markets. The second conjecture is that only a subset of the products that are brought into existence from a given invention will be launched through existing firms, whereas the rest will be launched through new firms. This conjecture is supported by casual observation that suggests that many new products are brought to market by new firms even in industries dominated by existing firms. For specific examples, one only needs to think of Amazon.com in book retailing and Wal-Mart in discount retailing. In this article, our focus will be on exploring the second conjecture.

Rather curiously, traditional theories of the firm do not take the phenomenon of the genesis of particular firms or markets as the subject matter of their theorizing (for an exception, see Rumelt, 1987), let alone recognize the importance of individuals in this process. Therefore, they do not attempt to directly address the key question of the role of existing and new firms in bringing new products and services into the economy. Instead, traditional theories focus on explaining the existence of the firm as a mode of economic organization. Holmstrom and Tirole (1989) state that a theory of the firm must answer two questions: first, why firms exist; and second, what determines their scale and scope, or, stated differently, why are all nonmarket activities not organized through one single giant firm (Coase, 1937)? Other theorists have suggested that additional questions are important; such as why do some firms persist over time (e.g., Hannan and Freeman, 1984). The best-known theory—the transaction cost theory of Coase (1937) and its derivatives—has developed into a rich research stream focused on these questions.

Nevertheless, this body of research leaves unaddressed important questions about the nature of firms that are crucial to entrepreneurship, namely, why do entrepreneurs create particular new firms to pursue opportunities instead of utilizing existing firms? Our suggestion is that there exists a distinctive entrepreneurial perspective on the theory of the firm that answers this question.

---

5 Sarasvathy and Venkataraman (2001) define technology broadly as knowledge that can be embedded in a product or service.
6 It is possible for some new products and services to be brought into existence through institutions that are not firms, e.g., Linux computer software.
7 Under the general description of “traditional theories” we include the remarkable original insights of Coase (1937) and all its subsequent derivatives (such as Williamson (1985) as well as contributions of Alchian and Demsetz (1972), Penrose (1995), Jensen and Meckling (1976), etc.
4. Uncertainty and the dispersion of knowledge

The concept of the dispersion of knowledge was formulated by Hayek (1945) in his famous essay on the role of knowledge in economic theory. Importantly, Hayek distinguished between two types of knowledge: first, the body of scientific knowledge, which is stable and can be best known by suitably chosen experts in their respective fields; second, the dispersed knowledge of particular time and place, whose importance only the individual possessing it can judge. Hayek pinpointed the harnessing of this latter type of knowledge as a key and underestimated element in the economic development of society. Fundamentally, because of the dispersion of knowledge, different people know different things and in this article we focus especially on the characteristics and nature of these differences as a fundamental facet in the explanation for the existence of firms.\(^8,9\)

Two aspects of the important relationship between the dispersion of knowledge and uncertainty have received extensive theoretical treatment. First, in a famous article, Akerlof (1970), pointed to the ways in which the dispersion of knowledge across people and places leads to asymmetries of knowledge that lead to uncertainty. In his original exposition of this issue, Akerlof pointed to the phenomenon of “lemons” (used cars for sale that are defective): the set of economic transactions that were closed because of the uncertainty caused by the “lemons” phenomenon. Akerlof argued that if the percentage of “lemons” in the total population of used cars for sale were to be high, then potential buyers, who have limited ways of telling a good car from a “lemon” due to their information asymmetry vis-à-vis sellers, would demand discounts on the prices of all used cars (including the reliable ones). The sellers of reliable used cars would then be discouraged from putting their cars up for sale. This would lead to an adverse selection, such that a larger and larger percentage of “lemons” would make up the population of used cars for sale. The information asymmetry between buyers and sellers could reach a point at which the market for used cars could come to a

---

\(^8\) For convenience, we use the term knowledge to include information in this article. We recognize that knowledge is normally described as “justified true belief” whereas information can be true or false or unverifiable and that the distinction between knowledge and information is sometimes important in an argument. However, our thesis rests on the issue of dispersion, not on the differences between knowledge and information.

\(^9\) We make a distinction between three types of knowledge: contingent, idiosyncratic, and specialized. Contingent knowledge is the knowledge that an agent receives by virtue of being in a particular place at a particular time and by interacting with particular individuals. The agent does not search for this knowledge but instead comes to know certain things serendipitously, by “blind luck” (Nelson and Winter, 1982). The size and characteristics of the agent’s social networks influence the type of contingent knowledge received (Granovetter, 1973). Idiosynratic knowledge is the knowledge that a particular person possesses or receives simply by virtue of being that person. The degree of heterogeneity of an individual with respect to other individuals in a group will determine the likelihood of her receiving or possessing differential knowledge. Specialized knowledge is knowledge that comes from particular educational training and professional experience. Specialized knowledge can be explicit or tacit. Tacit knowledge can be difficult to codify and transmit. A particular aspect of tacitness that Polanyi (1966) pointed out is ignorance of your own knowledge, i.e., until a situation arises in which it is required, the agent is not aware she has the requisite knowledge to make the decision. Since knowledge is contingent, idiosynratic, specialized, and partly tacit, different people know different things. In other words, knowledge is dispersed.
standstill. We describe Akerlofian information asymmetry as cross-sectional uncertainty because it is the uncertainty that arises when all the knowledge pertaining to a transaction is available in the system at one moment in time, but its unequal distribution creates uncertainty.

Knightian uncertainty (Knight, 1921) is different from cross-sectional uncertainty in that no agent in the system can possess accurate knowledge of the future owing to the fact that much knowledge relevant to the prediction of the future has not yet been created by any economic agent. Another way of looking at Knightian uncertainty is to see it as resulting from the lack of futures markets for most goods and services (Arrow, 1974; LeRoy and Singell, 1987). Uncertainty is therefore present because the knowledge relevant to future predictions is not simply unknown, but unknowable (Shackle, 1979). We describe Knightian uncertainty as longitudinal uncertainty owing to its time-dependent characteristics. This aspect of uncertainty has received less theoretical treatment in the literature and yet is critical to the theory of the firm.

At the heart of Knight’s exposition of longitudinal uncertainty is the insight that the creation of genuine novelty is a regular occurrence. Novel knowledge defies prediction in three of its aspects. First, the content of novelty cannot be predicted—to predict an invention is to go a long way toward inventing it. A second point that follows from the first is that the timing of the arrival of the invention cannot be predicted. Third, it logically follows from the previous two points that the effects of novel inventions cannot be predicted.10 This genuine lack of predictability owing to the production of novelty is the core aspect of Knightian uncertainty.

Perhaps the simplest and most elegant observation on the relationship between dispersed knowledge and Knightian uncertainty can be gleaned from Schumpeter’s (1934) metaphor of the role of combinations in the production of novelty. Schumpeter (pp. 65–66) points out that:

To produce means to combine materials and forces within our reach. To produce other things, or the same things by a different method, means to combine these materials and forces differently. . . . In so far as. . . . the new combinations appear discontinuously, then the phenomenon characterizing development emerges. . . . Development in our sense is then defined by the carrying out of new combinations.

Novel economic artifacts, such as innovations, appear here then as a combination of at least two prior artifacts, artifacts that were dispersed prior to their combination. Koestler’s (1964) famous metaphor of the “biosociative” process of creation may be seen as an elegant restatement (and broadening across the spectrum of creative acts) of Schumpeter’s basic metaphor of combinations of dispersed knowledge brought together to yield novel new economic artifacts. Similarly, in his discussion of invention, Wiener (1996) emphasizes in a

---

10 Edison’s invention of the phonograph is a wonderful example not only of the creation of genuine novelty but also of the unpredictability of the eventual effects of an invention: The invention Edison anticipated revolutionizing office work (as a record-keeping device) instead revolutionized (after a 100-year gestation period) the face of popular entertainment by making the record industry possible (by providing a cheap way of listening to music).
very similar way the dispersion of relevant knowledge across people and places (as well as the relationship between physical dispersion and temporal dispersion). The dispersion of knowledge matters in the production of new knowledge because when different people know different things the combination of “dispersed bits of incomplete and frequently contradictory knowledge” (Hayek, 1945) by an individual, or group of individuals, sometimes results in the creation of new knowledge that no individual had previously foreseen. As a result, new knowledge enters into the world.

One of Knight’s theoretical contributions was to point out the significance to economic affairs of such new knowledge by pressing home the distinction between uncertainty (outcomes whose probabilities are unknowable) and risk (outcomes whose probabilities can be subjectively assigned). Langlois (1984) argues that the distinction between these two concepts might be understood as the difference between structural and parametric knowledge. According to Langlois (p. 29), neoclassical economic theory assumes that the agent always has perfect structural knowledge and only leaves room for imperfect parametric knowledge:

What is often overlooked in the neoclassical literature of information and uncertainty is that its formalism implies certain-knowledge as much as it allows for uncertainty. The agent is implicitly presumed to have an exhaustive list of possible actions and states of the world and, equally important, a means–end framework relating the actions and the states of the world to his utility. We might say that the agent has certain-knowledge of the structure of the problem he faces or, put another way, that he has perfect structural knowledge. Imperfections in the agent’s knowledge extend only to specific parameters of the problem—the $x_i$—that are obscured from his vision. He may have imperfect parametric knowledge but never imperfect structural knowledge; he may acquire parametric information but never structural information.

An example of an extreme case of structural uncertainty is the range of outcomes from a basic research program. Assume that an entrepreneur wanted to explore the possibility of profiting from a discovery made in the current year by a theoretical physicist. How would she even go about structuring the problem? The discovery may have a commercial application in 50–100 years (a time frame well beyond the investment horizons of most individuals) and this possibility is itself a function of a multiplicity of other discoveries and events in the intervening period. A complete structure cannot be assigned to the problem because to create the structure would require the creation or discovery of the relevant future inventions. It is because of high structural uncertainty that there is no market for most outcomes from basic research programs.

As a point of contrast, risk is a situation that is insurable, meaning simply that the outcomes can be classified, subjective probabilities can be assigned, and a premium can be charged for possible unfavorable outcomes. For example, let us assume that an executive who has worked for 20 years for an established Fortune 100 company is contemplating joining a newly established firm as an employee. Clearly, there is risk associated with a change of employment and several things could go wrong in his future job—he may not get along with his colleagues, he may be unable to perform his job to the satisfaction of his superiors, or his
future employer could close down, to name just a few. These possible outcomes can be classified in a meaningful way and probabilities assigned to each outcome. The classification will be subjective, as will the probabilities, which means that different individuals will classify the outcomes differently and will assign different probabilities to the outcomes. Therefore, the problem is structurally manageable but parametrically uncertain.

Clearly, the practical differences between Akerlofian and Knightian uncertainty are differences in degree rather than differences in kind. However, these differences have important consequences for the theory of the firm. The important point about risk is that there is a mutual acceptance of the concept of a premium, and therefore of a technically insurable situation. The mutual acceptance of a premium opens the possibility for the parties to conclude the transaction using the price mechanism, which in effect means that the transaction has been concluded in the market. Uncertain situations are ones that exhibit failure of intersubjective agreement between the relevant parties, which means that the price mechanism cannot be used to conclude the transaction: In Knight’s (1921, p. 225) words, “[T]here is no valid basis of any kind for classifying instances.” This failure of the price mechanism therefore leads to one of the parties resorting to the contractual structure that we define as a firm.

5. Heterogeneous expectations and the individual opportunity nexus

In the previous section, we have argued that the dispersion of knowledge over people and places and over time leads to Knightian uncertainty. The dispersion of knowledge means that different people know different things, and in combination with Knightian uncertainty, it also contributes to different people having different expectations. Uncertainty both creates the opportunity and the necessity for expectations, and the dispersion of knowledge (among other things) makes it inevitable that those expectations will be heterogeneous.

At least four relevant observations can be made about the role of dispersed knowledge in the construction of heterogeneous expectations. The first is that expectations are an imaginative faculty, not immune from intersubjective agreement, but always subjectively held since another person cannot directly observe the imagination of another individual. Shackle (1979) reasonably speculated that the set of rival possibilities for the future that are imaginatively exercised by an individual is much larger than the actual single future that eventually prevails. Therefore, individuals seek to make the future more predictable for themselves by imagining alternatives but at the same time enlarge the possible set of future outcomes by introducing imaginative possibilities that other actors may not have thought of. They therefore increase aggregate uncertainty by increasing the number of decision possibilities that may occur in future scenarios. In addition, these imaginative possibilities lead individuals to have different expectations about the alternative values of resources.

11 In Langlois and Cosgel’s (1993, p. 460) words, “When the categories of knowledge themselves are unknown, they cannot form the basis of interpersonal agreement and market exchange. In such situations, firms supercede markets.”

12 Other factors that may contribute to heterogeneous expectations include stable psychological traits, upbringing, values, social reference points, etc.
Second, it seems reasonable to suggest that even if knowledge was not dispersed in the way Hayek described it, it is likely that individuals would have different expectations. Individuals possess different aspiration levels, psychological traits, cognitive processing capabilities, and prior knowledge. Much of the difference between individuals in these respects is tacit, and therefore, inaccessible not only to outsiders but often to the individuals themselves. These differences alone would lead individuals to have heterogeneous expectations even if they shared exactly the same knowledge of time and place. This explains why two individuals may know the same things but one may not see any potential for profit in it, while the other may see an entrepreneurial opportunity.

Third, costs are subjectively perceived by individuals (Buchanan, 1969). Cost is what the decision-maker sacrifices or gives up when he selects one alternative rather than another. Cost is therefore a forward-looking ex ante concept that depends on the imaginative faculties of the decision-maker. Since knowledge is dispersed, resulting in different people knowing different things, different people will have different costs; costs that depend on the array of alternative scenarios they imagine. This fact explains why individuals may have different opportunity costs, which cause them to have differing expectations about the value of an opportunity.

Fourth, expectations are notoriously volatile (Shackle, 1979). Expectations are constantly being modified as the results of past actions, both of the focal agent and of the other agents, start to become known. There is thus a two-way feedback—expectations lead to agents creating a future (which only partly matches their ex ante expectations) and as the future starts to unfold in unexpected ways the agents modify their expectations for further action. Network theoretic analysis of expectations, including the phenomena of fads and fashions, confirms the criticality of the dispersion of knowledge to expectation formation by pointing out how information channeling through social networks effects expectation formation by economic agents (Bikhchandani et al., 1992).

The implication of heterogeneous expectations is that we cannot talk about opportunities without talking about specific individuals. It is hard to imagine how certain firms could have come to be aside from the particularity of certain individuals: Wedgewood without Josiah Wedgewood, Ford without Henry Ford, General Electric without Thomas Edison, Wal-Mart without Sam Walton. The reason for this is that recognition of an opportunity usually (but not exclusively) depends on the prior knowledge and particular expectations of the individual (Venkataraman, 1997; Shane, 2000). The intersection of an individual and an entrepreneurial opportunity can occur either through a search process or fortuitously. However, the acquisition of knowledge does not guarantee that the individual will recognize and actively pursue an entrepreneurial opportunity.

Particular individuals are therefore likely to discover, create, and exploit particular opportunities. One implication of our emphasis on the individual and the opportunity is to highlight the incompleteness of focusing on the individual alone. Our discussion points to a reason for the general lack of success of the traits-based research on entrepreneurship. As we

\[13\] In contradiction to the theory presented here, some economists have theorized that general risk-taking propensity can be used to distinguish entrepreneurs from nonentrepreneurs, for instance, Kihlstrom and Laffont (1979).
have acknowledged, stable psychological characteristics do contribute to the process of opportunity recognition. However, taken in isolation and without due consideration to the quality of the opportunity, they are poor predictors of the likelihood of recognition of an entrepreneurial opportunity.

In summary, dispersion of knowledge and Knightian uncertainty lead economic agents to form heterogeneous expectations about an uncertain future. This heterogeneity is vital to the decision some agents make to create a structure of contracts we recognize as a firm.

6. The genesis of particular new firms

In the previous sections, we have established links among three phenomena: dispersed knowledge, Knightian uncertainty, and the heterogeneous expectations of economic agents. We have argued that these three phenomena lead particular individuals to pursue particular opportunities; we refer to this as the individual-opportunity nexus. In this section, we use the arguments developed so far to explain why some opportunities are pursued through existing firms and why some result in particular new firms coming into existence.

6.1. Opportunities identified by individuals outside existing firms

Knightian uncertainty is of special relevance to entrepreneurship literature since the central pursuit of the field is how, why, and when entrepreneurs manage, in the absence of markets, to create novel economic artifacts, such as new product and service markets (Venkataraman, 1997). The absence of markets is clearly a situation in which the entrepreneur faces Knightian uncertainty as the entrepreneur chooses to exploit an opportunity to bring new products and services to the market or to use new production processes where the value of final outputs are unknowable at the time the resources are committed. Conceptually, we distinguish between three types of markets—markets for the entrepreneur’s services, product markets, and factor markets. In the remainder of this section, we discuss how the presence or absence of these markets influences the manner in which the opportunity is taken to market.

6.1.1. Market for the entrepreneur’s services

Knight (1921) was clear that when an individual exercises judgment under uncertainty and decides to pursue an opportunity, there is no market for the entrepreneur’s services. This is because there is no intersubjective agreement, which is necessary for a market transaction to take place. If the entrepreneur could sell the opportunity she has recognized at a price that satisfies her, then we would say that a market exists for the opportunity. It is because no such market exists that she has to hire herself by starting a firm. In the recent literature on theories of the firm, several scholars have sought to clarify the distinctive characteristics of an original

---

14 In Boudreaux and Holcombe’s (1989, p. 147) words, “[T]he primary function of the Knightian firm is entrepreneurial; it is a necessary component in the creation of markets.” And, “[T]he decisions of entrepreneurs are crucial in determining what particular goods and services will be available on the market” (p. 150).
Knightian theory of the firm (Langlois and Cosgel, 1993; Boudreaux and Holcombe, 1989, Foss, 1993). In summary form, the argument runs as follows. Contractual rights can be of two types: specific and residual. The essence of the residual contract arises from the fact that specific contracts cannot be written to cover situations of Knightian uncertainty, since specific contracts cannot be written for situations that are at the present time not only unknown, but unknowable, given the emergence of new categories of novel possibilities. Where contracts are incomplete there will be residual rights, and therefore a residual contract. The residual contract enables the completion of the nexus of contracts that would otherwise remain incomplete. The residual contract holder will be the entrepreneur, whose specialized service of exercising judgment under situations of true uncertainty is ultimately noncontractible in the market.

For opportunities that are discovered or created outside existing firms, the creation of a particular new firm explains how uncertainty, heterogeneous expectations, and the individual-opportunity nexus are “organized” by the entrepreneur into a contractual structure that makes action possible. In essence, the entrepreneur expects to capture value from the difference between the ex post value of the firm and the aggregate ex ante price of the necessary factors of production. To access these factors of production, the entrepreneur has to establish a basis for exchange, given that markets for entrepreneurial services are closed in cases of Knightian uncertainty. Ex post, it is the residual contract that is the isolating mechanism for capturing profits from the ex ante conjectures, insights, judgments, and imagination of the entrepreneur, insights and conjectures that were not possible to trade in markets. The creation of a firm is thus an entrepreneurial response to the absence of a market for the entrepreneur’s services.

6.1.2. Product markets

As far as product markets are concerned, it is clear that the entrepreneur proposes to introduce some novelty into the market; the question is how much novelty? Clearly, the novelty involved in introducing an updated version of an existing automobile is not the same as that involved in introducing a completely new product, such as Mosaic, precursor of Netscape. The updated automobile may involve more comfortable interiors, better fuel consumption, and other similar features, but the firm that launches the updated product usually has a large amount of information that it can use to plan the launch. A launch such as that of Mosaic is a totally new product concept made available for consumers to do things they had never done before. The degree of uncertainty involved is positively related to the degree of novelty associated with the opportunity. The greater the uncertainty, the greater the likelihood that the opportunity cannot be transacted in the market.

---

15 Two alternative Knightian theories of the firm have also been prominent in the scholarship: first, Kihlstrom and Laffont’s (1979) model, suggesting that individuals with a taste for risk will become entrepreneurs; second, Barzel’s (1987) and LeRoy and Singell’s (1987) articles, suggesting moral hazard and adverse selection issues lead to the failure of markets for entrepreneurial services, necessitating entrepreneurs to start up their own firms.

16 Although the theory proposed by Grossman and Hart (1986) is remarkably similar to these Knightian theories, they do not use a Knightian rationale for their argument. We point out a difference between their position and a Knightian interpretation in the concluding part of this essay.

17 Clearly, this is not the same as saying a firm is a result of “market failure” (Williamson, 1985).
6.1.3. Factor markets

Finally, Knight (1921) was also unequivocal that the entrepreneurs “guarantee to those who furnish productive services a fixed remuneration,” (p. 271) meaning that the entrepreneur was the sole bearer of uncertainty, and purchased factors at fixed prices in existing factor markets. This statement can be refuted by offering counterexamples of the many independent entrepreneurs who often cannot resort to factor markets for their inputs, and others who choose not to do so. We therefore make a distinction between the existence (or absence) of factor markets and the ability (or inability) of entrepreneurs to resort to them. Specifically, we seek to explain those situations in which entrepreneurs with very low resource endowments manage to create long lasting and successful firms. Faced with a lack of resources, and factor markets that are inefficient to varying degrees, these entrepreneurs have no choice but to resort to nonmarket sources, such as family and friends, to finance the acquisition of their factor inputs. Using strategies that have been referred to as “bootstrapping,” these entrepreneurs manage to find solutions to factors markets that for all practical purposes are closed to them. The important point about the entrepreneurial use of nonmarket solutions is that through creativity and imagination the entrepreneur keeps the price of the important factors low enough to generate a positive outcome for the venture. By leveraging her human, social, and limited financial capital, the entrepreneur acquires factors at prices well below what they would otherwise be in the competitive factor markets under conditions of high uncertainty.

Of course, an entrepreneurial opportunity discovered or created independently may also be developed through other processes. In cases where a reasonable intersubjective agreement exists over the value of an opportunity, a second option exists, the sale of the idea to another firm. For this option to exist, less extreme perceptions of uncertainty and less heterogeneous expectations must attend the opportunity, i.e., the entrepreneurs sell ideas when less uncertainty pertains to them. Opportunities based on new knowledge are, in some exceptional cases, developed this way, as when an independent inventor licenses a patent to an existing firm in exchange (normally) for a fixed royalty payment and stream of licensing fees, contingent upon some measure of the utilization of the patent. Of course, in the case of the absence of prior formal property rights over the opportunity, the main factor mitigating against the sale of an idea is the indivisibility of information (Arrow, 1962), i.e., the

---

18 One of the most important factor markets that is notoriously inefficient is the credit market—often credit is not available at any cost to entrepreneurs who have neither a track record nor an asset base.

19 We say “exceptional” because most of the literature on patents shows that the development of new chemical compounds is the only class of opportunities for the application of new knowledge where patent protection is consistently meaningful. Lack of patentability and invent-arounds are the general rule.

20 Consistent with Arrow’s (1962) thesis, information dispersion through use is the general rule: even patents are granted for a limited period only. Especially worth noting is the relevance of Arrow’s arguments for the notion that differences in governance result from efforts to protect tangible versus intangible knowledge (also commonly referred to as explicit versus tacit knowledge). We believe these differences are based on an incorrect analysis of the nature of intangible/tacit knowledge and an incorrect attribution to Polanyi (1966) of the nature of the distinction (Tsoukas, 1996). The only difference we emphasize here is the difference between opportunities for the application of new knowledge that have prior legal property right protection (i.e., a defensible patent) and those that do not. The vast majority of entrepreneurial opportunities fall in the latter category.
entrepreneur must reveal the nature of the opportunity to sell it. However, given the legal protection sometimes afforded by a patent, the valuation of the opportunity depends upon the heterogeneous expectations of the parties. For patents with a wide range of potential applications, the patent holder commonly licenses the patent for specific uses with different firms, thus ensuring the maximum dispersion of the patent. The option value of the patent is thereby maximized by dispersing it across several market opportunities.

6.2. Opportunities identified by individuals within existing firms:

Much empirical evidence suggests that enterprising individuals often exit existing firms and form particular new firms as mechanisms to exploit entrepreneurial opportunities (Christensen, 1997). This scenario occurs when an opportunity is discovered or created by an individual or a team of individuals in an existing firm,21 but then exploited by the creation of a new firm. We make use of the concept of intersubjective agreement to explain whether an opportunity that is discovered or created in an existing firm will be pursued in the existing firm or result in the creation of a new firm.

One of Penrose’s (1995) insights was that existing firms have a natural perception “corridor” based on the knowledge of their employees, corporate goals, and incentive structures. This means that existing firms often become the entrepreneurial party in the discovery and creation of new markets by invisibly extending their existing residual contract structure to new entrepreneurial opportunities. In other words, individuals in existing firms often have a set of expectations that grant a higher valuation to particular opportunities than other stakeholders have, owing to the nature of their knowledge corridor.22 Therefore, firms seek endogenous growth, as Penrose suggested, when they have higher expectations of the value of an opportunity. These higher expectations are linked to current knowledge base, their current product markets, and their current incentive structures.

We have suggested earlier that the difference between risk and uncertainty can be viewed in terms of the degree of intersubjective agreement on the value of the opportunity. At one extreme, when there is no uncertainty attached to an opportunity, then a very large number of economic agents agree on its value; we could say that the opportunity exists in some objective sense, or that it is not individual-centric.23,24 A typical opportunity for arbitrage meets this criterion, in that one or more of a large number of individuals could exploit the opportunity. At the other extreme, when a very high uncertainty is attached to an opportunity, then only one individual sees value in it; very often, there is some attribute the individual possesses—prior knowledge, a particular network of connections, or a very specific asset—that leads her

21 As Christensen’s (1997) and other researchers have amply testified, individuals in existing firms are a major source—perhaps the major source—of new entrepreneurial opportunities. Our suggestion is that the notion that existing firms incubate most new opportunities has been underresearched to date.

22 We take this to be the reality of what Barney (1986) means when he suggests firms have “superior” expectations, i.e., the existing firm is often in the best position to judge the value of an opportunity.

23 This is not meant to be an ontological statement, but simply a convenient way of looking at things.

24 This is the case when a short-term profit possibility arises in the perfectly competitive market—a large number of economic agents swoops in to compete the profit away.
to exercise judgment. In this case, we could say that the opportunity resides totally in the
individual’s mind\textsuperscript{25} or that it is totally individual-centric.

We use intersubjective agreement as our organizing principle to describe four scenarios
(see Fig. 2). On the vertical dimension, we have the degree of intersubjective agreement
within the firm. At the heart of our thesis is the assertion that even when an opportunity is
identified within a firm, it is an individual or a team of individuals who recognize opportunities. These individuals must obtain agreement with other individuals in the firm on the value of the opportunity, which is not always easy. We can think of innumerable examples of opportunities that were identified within existing firms, but could not be pursued because the backing of key individuals in the organization could not be obtained. The case of an existing firm pursuing a radically uncertain opportunity raises important questions about the development of new opportunities in such organizations. The vertical dimension thus measures the degree of intersubjective agreement that exists between members of a firm on the value of a given opportunity. The greater the novelty of the opportunity identified, with respect to the firm’s current activities, the more difficult it will be to achieve a high degree of intersubjective agreement within the firm.

\textsuperscript{25} We could think of the radically subjectivist view of Shackle (1979) to be close to this position, although Shackle may claim that the opportunity often does not even reside in the individual’s mind—it emerges with action.
On the horizontal axis, we have the degree of intersubjective agreement outside the firm. This is degree to which individuals outside the firm would agree (if they had the information) on the value of a given opportunity recognized within the firm. The more objective the opportunity (i.e., the less individual-centric), the greater the intersubjective agreement outside the firm.

Quadrant 1 describes the scenario wherein intersubjective agreement is high both within the firm and outside it. This is a case of the opportunity being very obvious and existing in some objective sense. If this is the case, then the opportunity can be exploited within the existing firm or in the market (without the need for a new firm to come into existence). That is, a residual contract is not a necessary condition for this opportunity to be exploited. If it is exploited within an existing firm, it could be for two reasons: first, although it is individuals who recognize opportunities, the firm may have the property rights to the opportunity; second, even if the firm does not own the property rights, the opportunity may not be sustainable enough into the future to warrant the departure from the firm of the individual who identified it. On the other hand, if the opportunity is valuable enough, it is possible that the existing firm and the individual may both evaluate the opportunity positively and both elect to pursue it.

Quadrant 2 represents the scenario wherein the intersubjective agreement within the firm is high, but that outside the firm is low. This usually happens when individuals within firms identify opportunities that rely heavily on complementary assets, both tangible and knowledge-based, that are very specific to the firm (Rumelt, 1987). For example, molecules discovered by the R&D department of a pharmaceutical firm within its area of therapeutic specialization would come under this category. Both Quadrant 1 and Quadrant 2 scenarios lend themselves to explanation by the transaction cost literature. Specifically, the issues involved are the relative costs of organizing the opportunity within the firm vis-à-vis the market.

Quadrant 3 represents the situation in which the intersubjective agreement on the value of the opportunity outside the firm is high, but is low within the firm. Examples of this scenario are when the opportunity cost for the firm of pursuing the opportunity is high. From the viewpoint of the firm, the opportunity cost may be influenced by alternatives with greater materiality to the firm, or more within the existing strategic focus of the firm’s overall operations. Therefore, different expectations of what constitutes a “good return” in the context of the opportunity and its alternatives may lead to the pursuit of the opportunity by the entrepreneur through the creation of a particular new firm rather than the existing institutional alternative.

Quadrant 4 describes situations in which the intersubjective agreement is low both within firms and outside firms. This is a situation of true Knightian uncertainty. Typically, one individual or a very small group of them identifies an opportunity while working in a firm. The other members of the existing firm do not see the opportunity and are therefore uninterested in the firm pursuing it. The lone individual or small group of individuals then

---

26 We suggest that opportunities that fall within this Quadrant are rarely valuable or sustainable, precisely because they are so obvious.
leaves the employment of the existing firm to independently pursue the opportunity. However, unless the individuals who have spotted the opportunity have very high resource endowments, it is likely that they will be unable to access existing factor markets. This is because the high uncertainty associated with the opportunity (i.e., the product market uncertainty) is likely to have spilled over to the factor markets; factors will demand too high a premium for the individuals to be interested. In this scenario, there is a very high likelihood of a “bootstrapping” approach—the use of family and friends for resources in place of factor markets—until such time that the uncertainty over the opportunity has cleared sufficiently for the individuals to access factor markets.

7. Existing theories

We recognize that existing theories such as incomplete contracts theory (Grossman and Hart, 1986), transaction cost economics (Williamson, 1975, 1985), and agency theory (Jensen and Meckling, 1976) do possess the resources within their conceptual frameworks to explain why an opportunity will be pursued through a new firm rather then through an existing one; however, we aver that these theories have focused their attention on the explanation of other phenomena.

First, following Grossman and Hart (1986), the language of incomplete contracts has become relatively commonplace in institutional economics. Grossman and Hart (1986, p.691) posit that “when it is costly to list all specific rights over assets in the contract, it may be optimal let one party purchase all residual rights. Ownership is the purchase of these residual rights.” The point we emphasize here is that the difference between the contractual theory of the firm offered by Grossman and Hart and the one offered here is that under Knightian uncertainty the set of bilateral contracts that constitutes the firm is not just incomplete but incompletable without the stockholder contract. The use of the residual contract results from the fact that a complete set of contracts ultimately cannot be written27 owing to Knightian uncertainty. The residual contract is therefore better thought of as an institutional innovation—as a way of establishing property rights over uncertainty (De Soto, 2000). The contractual organization of the firm reflects how Knight described the state of our knowledge of the future: that it is structurally uncertain because of the dispersion of knowledge. The stockholder contract essentially does one simple thing: it is a contract that is written to cover the things that cannot be written in contracts.

Second, the emphasis of transaction cost economics is on the relative merits of the firm as an institutional modality vis-à-vis markets, and the structural features of the firm that result from the potential for ex post opportunistic behaviors by the contracting parties (Williamson,

---

27 We believe that this is what Grossman and Hart (1986) mean when they refer to “costly contracts.” These authors refer to the fact that contracts are written incomplete rather than the fact that they are incompletable. We simply make explicit the notion that under structural uncertainty, contractual incompleteness is not discretionary, and the cost of specifying all contingencies is therefore infinite. As Hodgson (1999) points out, in such situations there is no way of saying firms exist because they economize on transaction costs. Instead, we suggest the firm is better thought of as an institutional innovation to cope with contractual incompleteness.
According to transaction cost economics, firms exist to economize on transaction costs. The three building blocks that Williamson (1975, 1985) uses to develop transaction cost theory are information impactedness (different people have different information), uncertainty, and bounded rationality. These phenomena give rise to different costs associated with different modes of governance—firms, markets, etc.—and with different structural features—horizontal and vertical integration, etc. So for instance, transaction cost theory would explain the pursuit of an opportunity through a new firm rather than through an existing one by focusing on the transaction costs of the two alternatives. The alternative that minimizes these costs of transactions will prevail. However, as we have explained vis-à-vis incomplete contracts theory, the key issue in entrepreneurial scenarios subject to Knightian uncertainty is not that contracts are incomplete but rather that they are incompletable. Cost comparisons between market and firm contract structures are therefore moot. We would add that, as Conner (1991) has pointed out, it is the value-creating potential of an opportunity rather than simply transaction costs that determines the choice between markets and firms and between firms.

Third, agency theory emphasizes the characteristics of the firm that ensue from the degree of congruence in the incentives of the principals and agents (Jensen and Meckling, 1976; Holstrom and Tirole, 1989). Thus, agency theory relies on “incentive structures” instead of costs for explaining firms. The building blocks are similar to those of transaction cost economics. Given information asymmetry, uncertainty, and conflicting objectives (somewhat similar to our concept of heterogeneous expectations), those modes with the appropriate incentive structures that solve these problems will prevail. For example, due to information asymmetry, principal agent conflicts, and misaligned incentives, the discoverer of the opportunity may have an incentive to incorporate a new firm to appropriate the rents from the discovery or creation of a new opportunity. The existing firm may not have the right incentive structure to prevent such a creation or may prefer not to incur the costs of compromising or altering an existing incentive structure to accommodate the new opportunity.

While this theoretical viewpoint has much to commend it, the point we emphasize here is that in entrepreneurial scenarios subject to Knightian uncertainty it is not very meaningful to talk about aligning the incentives of principals and agents for all the possible future instances in which such opportunities may be detected. A perfect alignment will mean one of two things: symmetric utility functions (i.e., that the principal and agent are, in effect, the same person) or that the principal is willing to incur any cost to specify ex ante that all uncertain opportunities will be organized within the existing firm. Even if this were possible—and Knightian uncertainty suggests such an alignment for all time is ephemeral—the costs of achieving such an alignment will in many instances be so high as to neutralize the benefits of organizing the opportunities within the existing firm. Some principals will therefore resign themselves to the fact that certain opportunities detected by their agents will be pursued through new firms since the totality of scenarios requiring principal agent alignment cannot be foreseen. In essence, the alignment of contracts is also ultimately moot.

---

28 We posit that firms in which the detection of such opportunities is very likely will use stock options as a way of bringing the incentives of agents closer to those of principals.
8. Conclusion

In conclusion, an entrepreneurial theory of the firm based on dispersed knowledge generates several novel perspectives on the firm.

First, previous theories of the firm only provide a justification for the existence of firms as an institutional modality (usually vis-à-vis markets). They do not explain how firms come into existence and when a new firm, as opposed to an existing one, will be chosen as a vehicle for taking a new idea to market, i.e., whether the residual will be concentrated in existing or in new firms. The theory presented here provides answers to these questions, although our discussion has focused on the latter question.

Second, previous theories overlook the important role of particular individuals in originating novel economic artifacts, such as markets and firms. For example, neoclassical economic theory defines the firm as a production function that processes information freely available from the external environment, i.e., any risk-seeking individual can grasp any entrepreneurial opportunity. In contrast, our view emphasizes the individual-knowledge nexus in the creation of opportunity, and therefore the significant role that the idiosyncrasies of individuals play in the formation of markets and firms. Another example is previous theories such as the evolutionary view (Nelson and Winter, 1982) that emphasizes routines and activities as the carriers of knowledge within firms. Thus, again the role of the individual is a secondary consideration. Yet nothing is more apparent in economic history than the importance of both routines and individuals. Moreover, we would argue that the role of individuals is the primary issue that deserves explanation, since the establishment of routines depends on the prior realization of the nexus of the individual and opportunity and the subsequent establishment of a firm. In other words, the origin of routines must be explained in order for a theory of the firm based on routines to be robust. Therefore, our exposition posits that knowledge is dispersed and specific to individuals and explains how this contributes to the creation of routines. Rather than taking costs, incentives, or routines as the basic units of analysis, we take the individual-opportunity nexus as the basic unit of analysis.

Third, previous theories emphasize the contractual structure of the firm as an efficient solution to incomplete contracts, transaction cost minimization, or incentive alignment issues. The theory developed here explains why each of these issues is moot given the absence of markets that characterizes Knightian uncertainty: that contracts are not just incomplete but incompletable; that as a result of the incompleteness of contracts transaction cost considerations are moot; and that incentive structures are ultimately unalignable over all future scenarios. Instead, the contractual structure of the firm ought to be thought of as an institutional innovation that is a robust coping mechanism for the problems and opportunities arising from the prevalence of Knightian uncertainty in entrepreneurial situations.

In closing, we conjecture that at least two areas of potential further investigation arise from this article. The first is that the concept of the dispersion of knowledge (Hayek, 1945) might

\[29\]

For example, in the work of Chandler (1990) and McCraw (1997).
be seen as having the potential to bring coherence to several different aspects of entrepreneurship scholarship, such as the concept of arbitrage (Kirzner, 1973; Burt, 1992), creative entrepreneurship (Schumpeter, 1976), the costs of coordinating economic activity (Coase, 1937; Alchian and Demsetz, 1972), the role of firm growth (Penrose, 1995), and decision-making under uncertainty (Knight, 1921). The second is that a better linking of the creative processes of entrepreneurship and the rent-creating and rent-sustaining processes of business strategy is possible by situating these processes in the context of the dispersion of knowledge. Our suggestion is that a better understanding of how opportunities, which we take to be the source of all rents, arise and are converted into economic artifacts can improve the literature on business strategy.

Acknowledgements

We gratefully acknowledge the financial support of the Batten Institute of the Darden School of Business Administration, University of Virginia, for carrying out this research. We are also grateful to Linda Cyr, Laura Dunham, Ed Freeman, Mary Hamilton, Joo-Heon Lee, John McVea, participants in a seminar at CMIT, National University of Singapore, and participants in the 2002 Darden-Lally Entrepreneurship Retreat for comments on earlier versions of this article.

References

J. Polit. Econ. 94 (4), 691–719.