Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Organizational Structure and Product Choice in Knowledge-Intensive Firms

Yanhui Wu

MANAGEMENT SCIENCE
Vol. 61, No. 8, August 2015, pp. 1830–1848
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.2080

Organizational Structure and Product Choice in Knowledge-Intensive Firms

Yanhui Wu
Marshall School of Business, University of Southern California, Los Angeles, California 90089,
yanhuiwu@marshall.usc.edu

This paper formulates a model in which a firm simultaneously chooses its organizational structure and product position. The firm's production is knowledge intensive, requiring employees to solve problems. A vertical hierarchy, in which workers refer unsolved problems to managers facilitates the acquisition and leveraging of managers' superior knowledge. I show that a larger span of control is complementary to the provision of high-value products. Moreover, this complementarity is sustained when employees acquire sufficient knowledge and is further strengthened when the firm enhances its capability of communicating knowledge. The model yields testable implications concerning (1) the fit between a firm's product position and span of control, (2) the effect of information technology on product innovations and skill-biased organizational changes, and (3) the heterogeneity in hierarchical structure and human resource management in professional service firms.

Keywords: strategy and structure; strategic fit; product selection; organizational design; problem solving; knowledge-intensive firms
History: Received July 25, 2013; accepted September 16, 2014, by Bruno Cassiman, business strategy. Published online in *Articles in Advance* March 13, 2015.

1. Introduction

During the recent recession, four Wall Street law firms—Davis Polk & Wardwell, Cahill Gordon & Reindel, Sullivan & Cromwell, and Cravath, Swaine & Moore—won high praise from their peers, because they were able to "keep partner ranks lean" (Coe 2014). The rationale for such praise is that a lean organizational structure, measured by a large associates-to-partner ratio, enables a law firm to leverage the vast manpower of its associates and thus boost partner profitability. Contrary to this praise of big, lean law firms, some legal scholars argue that excessive partner-to-associates leverage may damage the quality of legal services and lead to the death of big law firms (e.g., Ribstein 2010).

Whether a lean organizational structure is a good thing is a question common to knowledge-intensive firms, where the main input in production is workers' knowledge. A notable phenomenon among professional service firms (PSFs)—one of the most prominent types of knowledge-intensive firms—is the substantial variation in their organizational structure. For example, the associates-to-partner ratio in accounting and engineering service firms can be more than 30, whereas the ratio in consulting and law firms is usually no more than five (Broderick 2011). In the legal industry, the associate-to-partner ratio varies from four to five for firms that provide regular corporate legal services to one to two for firms that focus on legal services related to sophisticated financial transactions such as merger and acquisition. A PSF's product choice and its organizational structure display a "strategic fit," as recognized by researchers (e.g., Maister 1993, Lowendahl 2005, Delong et al. 2007).

The strategic fit between product choice and organizational structure is a traditional topic in the study of business strategy. Since Chandler (1962), strategy scholars have primarily theorized on the fit between the strategy of product scopes and business portfolio and the structure of multidivisions and multilocations (e.g., Lawrence and Lorsch 1967, Rumelt 1974, Miles et al. 1977, Teece 1982, Hill and Hoskisson 1987). Although these studies offer important insights into the management of capital-intensive manufacturing and retailing firms, they are limited in addressing managerial issues in knowledge-intensive firms. For knowledge-intensive firms, products are largely specific to particular customer needs, and the primary organizational goal is to integrate specialized knowledge. Thus, the considerations of product scopes and coordination across functional or geographical units are secondary to concerns about product specialization and knowledge utilization.

In this paper, I formulate a model for firms' joint decisions of product choice and organizational structure in the context of knowledge-intensive production. Central in the model is a problem-solving approach that relates a firm's ability to acquire knowledge to its value creation through solving problems that occur in production. A hierarchy that features a vertical division of labor emerges to integrate specialized knowledge, to leverage superior knowledge, and to mediate between knowledge acquisition and product choice. From the model, I derive testable hypotheses about how the fit between product positioning and the span of control affects firm performance and how this fit responds to changes in a firm's ability to acquire and integrate knowledge.

I start from a problem-solving approach in which individual workers acquire knowledge, such as expertise and know-how, to solve problems that deliver value to the firm. Superior knowledge is more valuable, but is more costly to acquire. After it has been acquired, knowledge can be used concurrently by many individuals, without diminishing its availability to any of the users. As Garicano (2000) shows, to minimize the cost of knowledge acquisition and increase the intensive use of superior knowledge, a knowledge-based hierarchy emerges: production workers acquire knowledge about the most common or easiest problems confronted and pass the more exceptional or harder problems to problem solvers who acquire more-advanced knowledge and specialize in helping their subordinates. Such a model articulates Demsetz's (1988) observation that efficient knowledge acquisition entails specialization and that the efficient application of knowledge requires the integration of specialized knowledge.

Based on this problem-solving approach, I model the decisions of a firm, which is endowed with the technologies of acquiring and integrating knowledge, with respect to three choices: (1) the vertical position of a product that correlates its market value positively with the complexity of problems in production; (2) the hierarchical structure that specifies the extent of knowledge leverage measured by the span of control, i.e., the subordinates-to-supervisor ratio; and (3) the levels of knowledge acquired by all workers. The essential insight of the model is that the alignment between product position and hierarchical structure resolves the conflict between value creation through selecting superior products, on the one hand, and production efficiency through leveraging knowledge, on the other hand. For instance, a high-end product creates more unit-value for a firm, but workers have a high probability to encounter unknown problems and must frequently interact with managers. For a manager with time constraints, frequent interactions with one worker imply fewer interactions with other workers. Thus, the scope of knowledge integration is reduced and the extent of leveraging superior knowledge is limited. Conversely, a low-end product generates less market value, but allows a firm to easily sustain a high level of knowledge leverage and to achieve production efficiency. To simultaneously attain a high-end product position and a high level of knowledge leverage, a firm must acquire substantial knowledge for its employees and maintain a balanced distribution of knowledge across hierarchical layers.

The model establishes a mechanism that ties together a firm's knowledge input, organizational structure, and product selection to achieve superior performance. A firm's ability to acquire knowledge and its efficiency in integrating knowledge are two key drivers. I view a firm's ability to acquire knowledge as a capability that enables workers to overcome their cognitive limitations during the process of problem solving. Thus, better training and mentoring of workers, the adoption of advanced information-extraction technology, and the use of new learning methods all improve a firm's ability to acquire knowledge. I show that such improvements increase a firm's profitability by simultaneously expanding its span of control and upgrading its product as the consequence of an even increase in the knowledge level of all workers. A firm's ability to integrate knowledge depends on the codifiability of knowledge, which distinguishes whether knowledge is explicit or tacit and governs the cost to communicate knowledge. The model predicts that firms that primarily use coded knowledge in production, such as engineering and accounting firms, tend to employ a hierarchy with a large span of control, matched to an unbalanced distribution of knowledge across layers, whereas firms whose production relies more on tacit knowledge, such as consulting and law firms, tend to employ a hierarchy with a narrow span of control, matched to a balanced distribution of knowledge across layers. Furthermore, I show that improvements in communicating knowledge—either coded or tacit—lead a firm to expand its span of control, but may induce it to upgrade or downgrade its product, depending on whether knowledge in the production of a superior product becomes more or less codifiable. I apply these results to study the effects of advances in information and communication technology on skill-biased organizational change and product innovation.

"The fundamentals of the professional service business are brutally simple; it's about talent, it's about clients, and it's about teaming to bring it all together to create and deliver value," claims Jim Quigley, the former CEO of Deloitte Touche Tohmatsu Limited (Broderick 2011, p. 9). Our theoretical model captures the three key elements of Quigley's claim: (1) talent—the knowledge that is embodied in human minds; (2) product positioning, as determined by clients' valuation; and (3) organizational structure that brings the first two elements together. To illustrate the intuition and business relevance of the model, I draw extensively on observations from PSFs, which themselves comprise an increasingly important economic sector.[1] However, the basic insights of this paper apply to a wide range of other activities, as will be discussed in the conclusion.

1.1. Related Literature

This paper stands at the intersection of organizational economics and strategy research. Despite the burgeoning interest in the economic study of organizational design, organizational economists have not yet paid much explicit attention to strategy, as noted by Argyres et al. (2012) and Roberts and Saloner (2013). Several papers (e.g., Milgrom and Roberts 1990, 1995) model the complementarities among elements of strategy and organization. However, they focus on the fit of a variety of managerial instruments inside firms, taking product choice as given. This paper builds on a strand of modern organizational economics literature in which organizational structure is modeled as a coordination system to acquire, gather, and process dispersed information (e.g., Radner 1993, Bolton and Dewatripont 1994, Van Zandt 1999, Garicano 2000), but it departs from this literature by explicitly introducing product choice. The introduction of this new element shifts the focus from the resource-allocation role of organizational structure to the strategic role of organizational structure as a mediator between available resources and product choice. In this respect, this paper bridges a gap between organizational economics and strategy research.

Conceptually, this paper is closely related to the knowledge-based theory of the firm. The economics foundation of the hierarchy model can be traced back to Hayek's (1945) fundamental idea that organizations exist, to a large extent, to integrate knowledge that is dispersedly distributed among different individuals. Elaborating on this idea, numerous studies in strategy research have developed various theories of the firm as an institution for integrating knowledge.[2] Among these studies, Grant (1996a, b) are particularly similar to our paper in the emphasis on the coordination mechanisms through which knowledge is integrated to create capabilities. Complementary to his verbal arguments, I employ a formal model to articulate the premises (assumptions), the drivers (exogenous variables), and the effects (endogenous variables). Such formalization not only clarifies some implications in the existing literature and brings them closer to empirical research, but also generates several new insights regarding the interplay between knowledge acquisition and knowledge integration and regarding the formation of a firm's knowledge-based capabilities.

The behavior foundation of this paper is rooted in Simon's (1991, 1997) theory of "bounded rationality," in which agents are subject to cognitive limitations in learning, decision making, and communication. This bounded rationality view provides a common cornerstone for various studies of knowledge-based production.[3] Our problem-solving approach is conceptually close to Nickerson and Zenger (2004), which theorizes about the differentiation of organizational alternatives based on their efficiency in resolving conflicts during the search for solutions to problems. Their research emphasizes the cognitive process of acquiring knowledge and the effects of this process on a firm's boundaries, whereas our research emphasizes the economic efficiency of acquiring and integrating knowledge and its effects on firms' internal structure and the fit between strategy and structure. Although our economics approach differs substantially from those used in organization science, our analysis of the knowledge-integration mechanism bears some similarity to some recent studies of epistemic interdependence between agents and the resulting information processing within organizations (e.g., Puranam et al. 2012).

The rest of this paper is organized as follows. In §2, I present a simple example to sketch the basic idea of the paper. In §3, I present the formal model. In §4, I perform comparative statics analysis of the model to derive a number of testable empirical hypotheses. In §5, I discuss several extensions. In §6, I conclude. Proofs of lemmas and propositions are presented in the appendix.

2. An Example

In this section, I construct a simple example to outline the basic idea of this paper. A firm is modeled as a team of $1+s$ agents: one manager and s workers. This two-layer team aims to solve problems that lead to the production of a good. For example, in a firm that provides engineering services, a senior engineer supervises junior engineers to solve customers' problems; in a consulting or law firm, a partner supervises associates to provide the services demanded by customers.

[1] Herrendorf et al. (2013) show that of the 20% increase in the share of services in GDP that has taken place over the last 50 years, almost half was driven by the surge of professional and business services such as computer services, consulting and legal services.

[2] See Foss (2006) for an extensive survey of the literature on the knowledge-based view of organizations.

[3] See Garicano and Prat (2013) and Garicano and Wu (2012) for extensive discussions of this body of work and its implications for strategy research.

Production. Problem solving is the basic activity of knowledge-intensive production. Solving problems requires knowledge specific to problems, such as expertise and know-how. To capture the uncertainty of applying knowledge to problem solving, I assume that an agent, with a set of acquired knowledge A, can successfully solve a problem with probability $F(A)$. For simplicity, I take an agent's knowledge set as exogenously given.[4] Consider two agents endowed with knowledge sets A_1 and A_2, respectively. As long as A_1 and A_2 are not identical, the knowledge set of a single agent is a subset of the knowledge sets of the two agents. This specification includes two important situations in which: (1) one agent's knowledge encompasses the other's knowledge, and (2) one agent's knowledge supplements the other's knowledge. The first situation describes a master-apprentice relationship, such as the partner-associate relationship in law firms and consulting firms; the second situation describes a relationship between two specialists, in which one agent specializes in advanced knowledge and the other specializes in elementary knowledge, such as engineers in engineering service firms and doctors in healthcare firms. In either situation, I label the agent with superior knowledge—the one with either broader or more-advanced knowledge—as a manager (she) and the other as a worker (he). I denote the knowledge set acquired by the manager A_m and the one acquired by a worker A_w. Reasonably, $F(A_m \cup A_w) > F(A_w)$.

Organization. In the aforementioned production process, one way of organizing production is that the manager specializes in helping workers who fail to solve a problem. As a result, the manager's knowledge, once acquired, can be used multiple times in combination with workers' knowledge. This is precisely the notion of knowledge leverage in the management of PSFs. However, the extent of knowledge leverage is constrained by the cost of communication in referring unsolved problems because effective communication takes time—an extremely scarce resource. Without loss of generality, I assume that the communication cost, denoted as $h \in (0, 1)$, is all borne by the manager in terms of her time. The manager, who supervises s workers, faces the following constraint: $[1 - F(A_w)]sh \leq 1$, where $1 - F(A_w)$ is the probability that a problem is beyond a worker's knowledge set and thus referred to the manager. The right-hand side of the inequality is the manager's available work time normalized to one, which is the upper limit that a manager can use to help her s workers. In this two-layer hierarchy, the organizational structure of the firm is fully characterized by s, which I refer to as the span of control or the extent of knowledge leverage.

Product. In addition to the choice of the span of control s, the firm can choose the complexity of problems, indexed with a scalar k. A problem is more complex when it is less decomposable and requires more knowledge to solve it (e.g., Simon 1962, Kauffman 1993). Consequently, for an agent with a constant knowledge level, the probability of solving a more-complex problem is lower: $F(A, k') < F(A, k)$ if $k' > k$. However, solving problems that are more complex can create more value for customers and permit charging a higher price: $\phi(k') > \phi(k)$ if $k' > k$. I thus call a product indexed with a larger k as a superior product.

In sum, endowed with the knowledge sets (A_m, A_w) and a communication cost h, a firm aims to maximize its output:

$$\max_{\alpha, s, k} \; y = \phi(k)\big[\alpha F(A_m, k) + (1-\alpha)F(A_m \cup A_w, k)s\big] \quad (1)$$
$$\text{s.t.} \;\; [1 - F(A_w, k)]sh \leq 1.$$

In this objective function, the manager can spend α fraction of her time solving problems by herself and $1 - \alpha$ fraction of her time helping workers. Because $F(A_m, k) < F(A_m \cup A_w, k)$ and $s > 1$, it is optimal for the manager to fully specialize in helping workers, that is, $\alpha = 0$. Then, the constraint is binding, as the profit-maximizing firm will use up all the manager's available time. Rewrite the binding constraint as $s = (1/h)(1/(1 - F(A_w, k)))$. This means that the span of control is constrained by the workers' ability to solve problems and the efficiency of communication between the manager and workers. Substituting $\alpha = 0$ and the transformed constraint into the objective function, the optimization problem boils down to a problem of product choice:

$$\max_k y = \phi(k)\frac{F(A_m \cup A_w, k)}{1 - F(A_w, k)}\frac{1}{h}. \quad (2)$$

Suppose that a firm currently offering product k considers moving to a superior product position k' so as to reap a higher profit margin: $\phi(k') > \phi(k)$. This adjustment, however, would incur two costs. First, because a superior product requires solving problems that are more complex, the firm's production is less effective: $F(A_m \cup A_w, k') < F(A_m \cup A_w, k)$. This is the production cost. Second, workers encountering problems that are more complex must ask the manager for help more frequently, which reduces the span of control and consequently restricts the intensive use of the manager's knowledge. This is the organizational cost. Only when the superior product's market value

[4] In the formal model, workers' knowledge will be endogenous, and whether they acquire identical knowledge or not will be a result of optimization.

exceeds the production and organizational costs can a firm profit from choosing it. Even when production cost is low because of a highly talented manager, a high organizational cost can limit the leverage of the manager's knowledge and thus constrain the firm's product choice.

An important result arises from the previous analysis: The match between a firm's product choice and organizational structure depends on the distribution of knowledge across hierarchical layers. This can be seen from the term $F(A_m \cup A_w, k)/(1 - F(A_w, k))$ in (2) that indicates the probability of a problem being solved by the manager conditional on it being unsolved by a worker. For a team composed of a competent manager but less-able workers, $F(A_m \cup A_w, k)$ is large, implying that the probability for the firm to solve complex problems is high; however, $1 - F(A_w, k)$ is large as well, implying that the less-able workers will frequently request help. Thus, it is difficult to leverage the manager's knowledge, and a firm optimally chooses a superior product and a narrow span of control. Such a strategy focuses on a high profit margin, forgoing production efficiency per unit. Conversely, with a team composed of an incompetent manager but able workers, a firm can easily leverage the manager's knowledge, although the probability for the firm to solve complex problems is low. Consequently, the firm prefers the combination of a low-end product and a wide span of control; its strategy focuses on production efficiency—the low average cost because of more intensive use of knowledge, forgoing a high profit margin. Only when both the manager and the workers are competent, can a firm match a superior product with a wide span of control and realize the dual benefits of a high profit margin and the intensive use of the manager's knowledge.

The previous example illustrates the essential idea of this paper: Organizational structure mediates between available resources and product choice. In the example, however, available resources—the agents' knowledge—are exogenously provided, and the division of labor among agents is prescribed. An analysis of knowledge-intensive organizations is incomplete without understanding how knowledge is acquired and how labor is divided among workers. A more sophisticated model is required to track these variables together with a firm's product choice and organizational structure.

3. The Model

In this section, I introduce a general approach to represent the problem-solving process and elaborate on the basic concepts mentioned in the previous section. The model preserves the basic features of the example in §2: knowledge is the main input in production and problem solving is the primary productive activity. However, the model differs from the example in that agents must acquire knowledge to solve problems. This new feature gives rise to a fundamental trade-off between the cost of acquiring knowledge and the benefit of using knowledge intensively, which in turn determines the division of labor among agents and thus the efficient allocation of resources inside firms.

3.1. Setup

3.1.1. Production. As in §2, production requires a worker to use knowledge to solve problems. I restrict attention to knowledge that is (1) used to solve specific problems and (2) embodied in the human mind. This restrictive definition of knowledge differentiates our study from the vast economic studies of information and technology. In our study, knowledge is above all a cognitive capability. Its acquisition requires the mobilization of scarce cognitive resources, and its integration is subject to cognitive limitations. I see this bounded rationality view of knowledge-based production as applicable to general production processes.

Following Garicano (2000), I use a stochastic method to formulate the previous problem-solving process. I define a random variable Z to represent the problems that a producer will potentially encounter during production. Let $\Omega \subset \mathbb{R}^+$ be the set of all possible problems and $A \subset \Omega$ be the set of problems that the producer is able to solve, referred to as the knowledge set. The distribution of problems associated with a particular task is represented by a continuous and differentiable probability distribution $F(Z)$ with a density function $f(Z)$ over Ω. Output is produced when a problem randomly drawn from the distribution $F(Z)$ lies within a worker's knowledge set A.

The density function $f(Z)$ measures the frequency of problems. I normalize $f(Z)$ to be nonincreasing so that problems are ordered from most to least common. Such normalization captures the idea that when performing tasks, workers identify and attempt more-common problems before less-common ones. With this normalization, I can compare two knowledge sets $A_1 = [Z_1, Z_2]$ and $A_2 = [Z_3, Z_4]$ such that $Z_4 > Z_3 > Z_2 > Z_1$. The problems encompassed in set A_2 are less common than those in A_1. I thus call the knowledge in A_2 more advanced than that in A_1. Moreover, the density function $f(Z)$ provides a convenient characterization of a task. For instance, routine tasks can be characterized by a density distribution with high frequency of common problems (i.e., a small value of Z), whereas innovative tasks are associated with a high frequency of uncommon problems (i.e., a large value of Z). For expositional simplicity, I use the term "complexity" to describe a task that is characterized by the distribution of common and uncommon problems encompassed in it. For instance, a task is more complex if $F(Z)$ has a thicker tail, implying a higher probability of encountering uncommon problems. Conceptually, our notion of complexity is consistent with the one that is used in the studies of complex system (e.g., Simon 1962, Kauffman 1993).[5]

Provided that the knowledge used in production is specific to the problems to be solved, a worker must expend effort and time to acquire the knowledge needed to identify problems, search for solutions, and eventually match solutions to problems. This process of knowledge acquisition, even in the form of reproducing existing knowledge, is costly. For simplicity, I assume that the cost of acquiring a knowledge set A (learning all the problems in A) is proportional to its size $\mu(A)$, i.e., the Lebesgue measure of the set A. For example, if $A = [0, Z_w]$, then $\mu(A) = Z_w$, and the cost of acquiring the knowledge is $c \cdot \mu(A)$. Here, the parameter c measures the marginal cost of acquiring knowledge that I refer to as a producer's learning cost.

Generally, a worker can acquire a knowledge interval $[Z_3, Z_4]$, for some arbitrary Z_3 and Z_4 such that $Z_4 > Z_3 > 0$. Compared to a knowledge interval with the same size but smaller boundary values, say, $[Z_1, Z_2]$ such that $Z_4 > Z_2 > Z_1 \geq 0$, the cost of knowledge acquisition is the same. However, because of the assumption of a nondecreasing $f(Z)$, the problems encompassed in $[Z_3, Z_4]$ are less common and more difficult to solve than those in $[Z_1, Z_2]$. Holding the probability of solving problems the same, the cost of acquiring more-advanced knowledge that is used to solve less-common problems is effectively higher. Thus, it is unnecessary to assume that the unit cost of knowledge acquisition increases with the complexity of problems.[6]

In sum, facing a task with a distribution of problems, $F(Z)$, a single worker acquires knowledge $A_w = [0, Z_w]$ to maximize the expected net output y:

$$y = \Pr(Z \leq Z_w) - cZ_w = \int_0^{Z_w} f(\varphi)\, d\varphi - cZ_w. \tag{3}$$

Note that in this single-worker production problem, a worker acquires a knowledge interval starting from zero, which indicates the most common problems because of the assumption of the nonincreasing density function for the random variable Z.

[5] In these studies, the complexity (of a system) is defined by the number of components in the system and the frequency of interactions among these components. In my model, a problem can be decomposed into a sequence of interdependent subproblems. A problem is solved only when all subproblems are solved. The larger the number of subproblems, the less likely that a problem is solved. Thus, a task is more complex when the problems encountered during performing the task consist of a larger number of subproblems.

[6] The current modeling framework is equivalent to modeling a constant marginal benefit of solving problems with increasing marginal costs of acquiring more-advanced knowledge.

3.1.2. Organization. Similar to the example in §2, a firm can employ a team to improve the efficiency of production. Differently, I now allow for an arbitrary formation of a team, and a hierarchy will emerge as an equilibrium outcome. As in Garicano (2000), a team is defined as a partition of its members into L classes of size β_i (measured in terms of a fraction of the organization), where $i \in [0, L]$ indexes a particular class of the organization. For notational convenience, I treat i as if it were an integer. Let $A_i = [Z_{i-1}, Z_i]$ denote the knowledge interval to be acquired by an agent of class i, and $z_i = Z_i - Z_{i-1}$ represent the size of the knowledge set. Then, $F(z_i)$ is the probability of the problem being solved by an agent in class i, and cz_i measures the cost that this agent expends to acquire the knowledge. An agent in class i can refer an unsolved problem to an agent in class $i+1$. The probability that a problem is solved when it is passed through workers from the bottom ($i = 0$) to the top ($i = L$) is $F(\sum_{i=0}^{L}(Z_i - Z_{i-1}) = F(\sum_{i=0}^{L} z_i)$. As described in §2, referring a problem to another worker, regardless of whether the problem is solved, incurs a communication cost, $h < 1$, in terms of the receiver's time. With this specification, the firm maximizes its expected output by allocating the size of knowledge intervals, z_i, to members at each layer, the fraction of members, β_i, at each layer, and the number of layers, L, of the hierarchy, as follows:

$$y = \max_{z_i, \beta_i, L} \; F\left(\sum_{i=0}^{L} z_i\right)\beta_0 - c\sum_{i=0}^{L} z_i \beta_i, \tag{4}$$

$$\text{s.t.} \quad \left[1 - F\left(\sum_{j=0}^{i-1} z_j\right)\right]\beta_0 h \leq \beta_i \quad \text{for all } i > 0; \tag{5}$$

$$\text{and} \quad \sum_{i=0}^{L} \beta_i = 1. \tag{6}$$

The term $[1 - F(\sum_{j=0}^{i-1} z_j)]\beta_0$ in constraint (5) measures the number of unsolved problems that pass through layers from 0 to $i-1$ and must be dealt with by agents at layer i. Thus, the constraint simply means that the time that an agent in class i spends helping agents in class $i-1$ when they refer unsolved problems cannot exceed her available time, which is normalized to one. Constraint (6) is merely an identity to guarantee that all the fractions of agents add up to one.

3.1.3. Product. I now introduce the key strategic variable—product choice—on top of the previous production and organization problems. A firm can choose a product, indexed with a scalar k, from a continuum of product position. As in §2, the index k indicates two sides of a product. On the supply side, it indicates the complexity of tasks in the production process. A larger k implies greater complexity of tasks and a larger probability of encountering uncommon problems. On the demand side, k indicates the market position of a product. A larger k implies a higher product position, at which consumers are willingness to pay a higher price. This specification of k is analogous to a model of vertical product differentiation in the industrial organization literature. In the rest of the paper, I call a product indexed with a greater k "a product at a higher position," or simply, "a superior product."

The positive relationship between a product's market value and the complexity of production tasks is prominent in the professional service industries. For example, Maister (1993) classifies professional services into three types: (1) the "commodity" type that only requires solving routine and standardized problems, (2) the "process" type that requires solving a wide range of normal problems with occasional occurrences of exceptional ones, and (3) the "customization" type that requires solving complex problems with frequent encounters of exceptional ones. Commodity services, including taxation and financial services for individual customers and basic software diagnosis and engineering tests, have low market value. Process services, including corporate accounting and legal services and the design of modestly-scaled information or engineering systems, have intermediate market value. Customization services, including strategic management consulting, legal services for corporate financial transactions, and the design of innovative information or engineering systems, have high market value.

As in §2, I denote the price of a unit of product k as $\phi(k)$, with $\phi'(k) > 0$. A firm's profit-maximization problem is

$$\pi = \max_{k,\, z_i,\, \beta_i,\, L} \quad \phi(k) F\left(\sum_{i=0}^{L} z_i ; k\right)\beta_0 - c\sum_{i=0}^{L} z_i \beta_i - g(k), \qquad \text{s.t. (5) and (6).} \tag{7}$$

Compared to (4), the objective function (7) accommodates the parameter k in three places. First, $\phi(k)$ is the market price of product k. Second, as a measure of task complexity, k directly affects the functional form of the knowledge distribution. Third, mostly for technical purposes, I add the term $g(k)$ with $g'(k) > 0$. Economically, $g(k)$ can represent two costs: the entry cost and the cost of capturing customers. $g(k)$ increases with k, as it would be more costly to enter a high-end market segment and to maintain the customer relationship for a superior product. Note that I take the price function $\phi(k)$ as given, as this paper focuses on a firm's decisions about the alignment between product selection and internal structure. One can endogenize this pricing function in a market equilibrium model, as will be discussed later in §5.4.

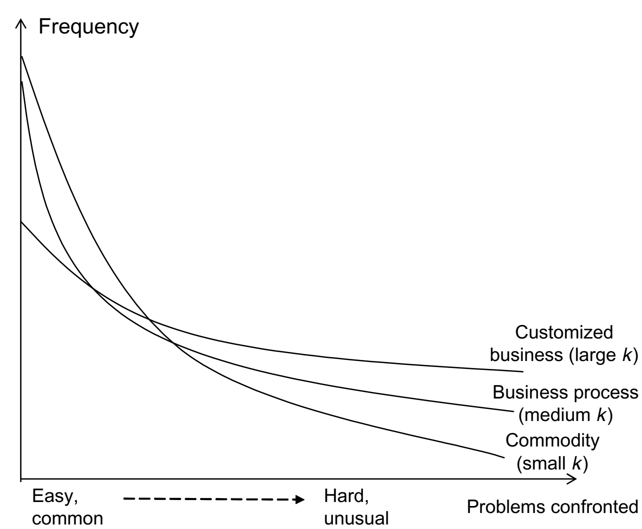
3.2. Assumptions

From the single-worker production problem (3) to the organizational problem (4) to the strategy problem (7), a firm must make three sets of decisions regarding (1) the levels of knowledge to acquire for all workers and the distribution of knowledge among them, (2) the organizational structure that specifies the height (the number of layers) and width (the span of control) of the hierarchy, and (3) the product position that determines the market value and the complexity of problems that occur during the production process. These decisions are interdependent, and solving the entire optimization problem is nontrivial. To simplify matters, I make two additional assumptions.

Assumption 1. *The complexity of tasks is characterized by an exponential distribution of problems encountered in production*: $F(x) = 1 - e^{-(1/k)x}$ *with* $x \geq 0$ *and* $k \geq 0$.

The qualitative results of this paper hold for a general class of probability distributions with nondecreasing density. The exponential distribution brings about substantial technical convenience because of its memoryless property. Moreover, it has a natural economic interpretation of complexity. The single parameter $1/k$, which governs the shape of the distribution, determines the frequency of encountering uncommon problems, the average difficulty (the mean of the distribution), and the predictability (the variance). Thus, a larger k indicates a more-complex task in the sense that uncommon problems are encountered more often, and that problems are on average more difficult to solve and are less predictable. Maister's classifications of the three types of professional services correspond, respectively, to exponential distributions with a small, intermediate, and large k. Figure 1 depicts such a correspondence.

Figure 1 **Characterization of Tasks with Exponential Distributions of Problems**



Assumption 2. *The cost function $g(k)$ is sufficiently convex relative to $\phi(k)$: $g''(k) - \phi''(k)$ is positive and sufficiently large.*

This technical assumption is imposed to rule out the extreme situation in which the marginal value of a higher product position increases so quickly that a firm will always choose a corner solution. In the rest of this paper, Assumptions 1 and 2 are taken as given unless otherwise specified.

3.3. Solutions

I break down the profit-maximization problem (7) into two stages. First, taking product position k as given, the firm optimally chooses its organizational structure and acquires knowledge for all of its employees. This can be thought of as a short-run scenario in which a structure is designed to follow a selected product strategy. Second, the firm chooses its product position. This reflects a long-run scenario in which a firm can adjust both its product strategy and its organizational structure.

3.3.1. Organizational Structure. For a given k, I divide all terms on both sides of (7) by $\phi(k)$ to obtain a normalized objective function:

$$\frac{\pi}{\phi(k)} = \max_{z_i,\,\beta_i,\,L} f\left(\sum_{i=0}^{L} z_i, k\right)\beta_0 - \frac{c}{\phi(k)}\sum_{i=0}^{L} z_i\beta_i - \frac{g(k)}{\phi(k)}. \quad (8)$$

This objective function, together with the constraints (5) and (6), is a variant of (4). I apply the method in Garicano (2000) to solve such an optimization problem. First, with the normalization that more-common problems encountered first (recall that the density distribution of problems is nonincreasing) and the assumption that the cost of acquiring knowledge is proportional to the size of the knowledge, it is optimal for an agent in class $i > 0$ to completely specialize in helping agents in class $i-1$.[7] Therefore, the constraint in (5) is always binding. Then, I can use the recursive structure of this binding constraint to eliminate β_i from the objective function (8). From $\sum_{i=0}^{L}\beta_i = 1$, I can also eliminate β_0 and write (8) as a function exclusively of the knowledge acquired by agents at each layer. Absent integer constraints, the optimum number of layers (L) is unlimited, because the exponential distribution of $F(\cdot)$ has the memoryless property. To keep the model tractable, I do not impose an integer constraint on the number of layers.[8] Implicitly, the employment size of the organization is measured in terms of labor that is infinitely divisible, instead of in terms of people, so it is reasonable not to impose integer constraints.

Intuitively, after a piece of knowledge has been acquired with a fixed cost, the firm wishes to use this knowledge as many times as possible. To minimize the cost of knowledge acquisition, division of labor is needed: each agent acquires a different set of knowledge;[9] to intensify the use of knowledge, the agents whose knowledge is more costly to acquire specialize in helping other agents with their unsolved problems. Thus, a hierarchical structure based on agents' specialization in knowledge is used to trade-off the intensive use of knowledge against the cost of acquiring knowledge. Analytically, the agents at the bottom ($i = 0$) specialize in production without helping any others, whereas agents at all the upper layers specialize in helping subordinates solve problems. Because of this pattern of vertical specialization, I will refer to agents at the bottom layer as workers and agents above the bottom layer as managers, among whom I distinguish between senior and junior managers when necessary. In the context of PSFs, workers are associates who directly handle customers' problems, and managers are principals or partners who focus on problems that are referred by associates.

With an exponential $F(\cdot)$, all managers, regardless of which layer, acquire the same size of knowledge at optimum, i.e., for $i > 0$, $Z_i - Z_{i-1} = Z_{i+1} - Z_i$ or $z_i = z_{i+1}$. This gives a closed-form solution to the optimal size of knowledge intervals acquired by workers and managers, respectively,

$$z_w^* = k\ln\left(\frac{\phi(k)}{ck} - \ln h\right) + k\ln h, \quad (9)$$

$$z_m^* = k\ln\left(\frac{\phi(k)}{ck} - \ln h\right). \quad (10)$$

The span of control, defined by the ratio of agents across two adjacent layers, is constant:

$$s = \frac{\beta_i}{\beta_{i+1}} = \frac{\phi(k)}{ck} - \ln h.^{10} \quad (11)$$

The intuitions of the previous solutions can be illustrated by the two figures next: Figure 2(a) illustrates the optimal acquisition of knowledge for each agent, and Figure 2(b) illustrates the optimal hierarchical structure. In Figure 2(a), agents in class $i = 0$ acquire a knowledge set $[0, Z_0]$, agents in class $i = 1$ acquire

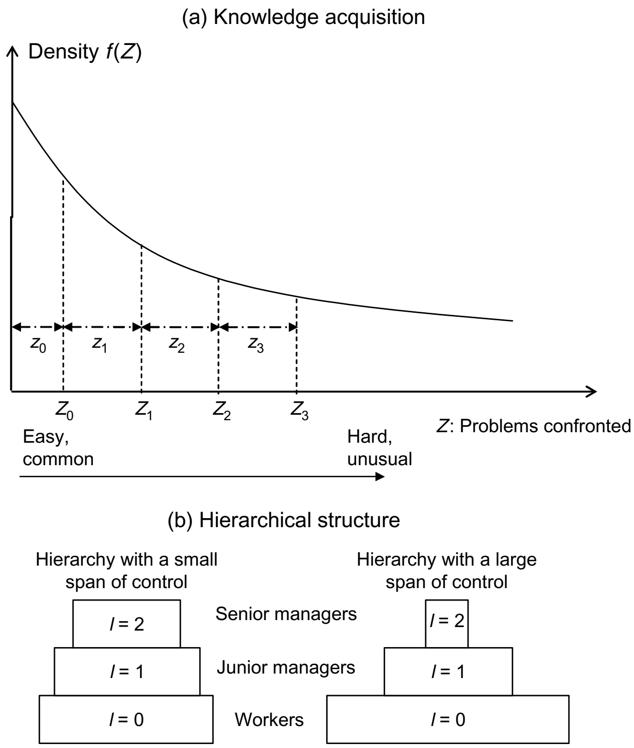
[7] Proof of this result in a more general setting can be found in Garicano (2000).

[8] If the number of layers is restricted, the basic results obtained in this paper remain unchanged at the cost of rather complex mathematics. A version in which the hierarchy is restricted to a finite number of layers is available upon request.

[9] The results will be qualitatively unchanged if the knowledge sets of agents between layers ought to overlap, i.e., a helper's knowledge should encompass her subordinate's knowledge.

[10] To guarantee the existence of a positive z_m^*, we assume $\phi(k)/ck - \ln h > 1$, which also implies $s > 1$.

Figure 2 Optimal Knowledge Acquisition and Organizational Structure



$[Z_0, Z_1]$, agents in class $i = 2$ acquire $[Z_1, Z_2]$, and so on. Because of this division of labor, agents who ex ante have the same ability to acquire knowledge are ex post different: agents in class $i > 0$ acquire knowledge that is more advanced—the knowledge that is used to solve less-common problems—than those in class $i-1$. A problem does not reach an agent at layer i until it goes through the previous $i-1$ layers. This "management by exception" is a way to protect agents with advanced knowledge from dealing with common problems. In the words of Alfred Sloan, "I work fairly hard, but on exceptions" (Sloan 1924, p. 137).

Figure 2(b) depicts two hierarchies that have a similar pyramid structure, with the number of agents at an upper layer decreasing exponentially. In this model, the constant span of control in (11), owing to the exponential distribution of problems, entails such a pyramid hierarchy. This result remains true in more-general models. Intuitively, the knowledge of upper-layer agents is used to solve less-common problems, and the cost to acquire such knowledge (spread over solved problems) is higher. To economize the costs of knowledge acquisition, it is optimal to allocate fewer agents at a higher layer of the hierarchy. This is the core idea of knowledge leverage in the management of PSFs, in which a partner mentors several principals, each of whom in turn mentors a number of associates to achieves scale economies in the use of the partners' advanced knowledge.

The two hierarchies in Figure 2(b) have the same size in terms of units of labor, but differ in their spans of control. The hierarchy shown in the left graph features a narrow span of control, with the size of the bottom layer (workers) being small and the size of the managerial layers slowly convergent toward zero. In contrast, the hierarchy shown in the right graph features a broad span of control, with the size of the bottom layer (workers) being large and the size of the managerial layer quickly shrinking. A greater span of control means that the knowledge of agents at a higher layer can be used more intensively. In the figure, the knowledge of the senior manager (at the second layer) can be used s^2 times, which is a convex function with respect to s.

Several features with regard to solutions (9)–(11) are worth noting. First, the solution in (10) can be written as $z_m = z_w - k \ln h$, implying that the manager's knowledge increases with the worker's knowledge. This is because a more able worker can release his manager's time to help other workers, resulting in an increase in the returns to the manager's knowledge. Second, given k and h, the span of control s pins down the knowledge distribution across layers in the hierarchy.[11] A larger span of control encourages a manager to acquire more-advanced knowledge because her knowledge can be used more intensively. It also encourages a worker to further acquire knowledge because of the endogenous complementarity between managers' knowledge and workers' knowledge. Third, an organization's production efficiency can be fully characterized by the span of control. Thus, despite that the optimal number of layers in this model is unlimited, I can regard the hierarchy as consisting of two layers: a layer of workers and a layer of managers.

3.3.2. Product Choice. I now turn to the decision of product choice. Substituting the optimal solutions (9)–(11) into (7), the optimal output can be written as

$$\frac{\pi^*}{\phi(k)} = 1 - \frac{ck}{\phi(k)}\left[1 + \ln\left(\frac{h\phi(k)}{kc} - h \ln h\right)\right] - \frac{g(k)}{\phi(k)}.$$

The firm's objective function boils down to

$$\max_k \pi^* = \phi(k) - g(k) - ck\left[1 + \ln h + \ln\left(\frac{\phi(k)}{kc} - \ln h\right)\right]. \tag{12}$$

The first term in (12) is the market value of product k, the second term is the entry cost of attaining and capturing such a product position, and the last term indicates the cost of acquiring knowledge for all agents to solve the problems associated with product k.

[11] These results are obvious when they are rewritten as $z_m^* = k \ln s$ and $z_w^* = k(\ln s + \ln h)$.

Lemma 1. *Under Assumptions* 1 *and* 2, *there always exists a unique interior solution* k^* *to the firm's profit-optimization problem* (12).

Proof. See the appendix. □

With Lemma 1, the optimal product choice is determined by the following first-order condition:

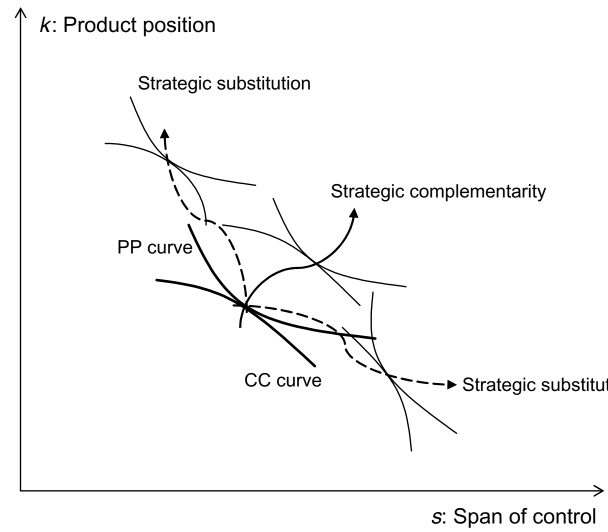
$$g'(k^*) + c\left[1 + \ln\left(\frac{h\phi(k^*)}{k^*c} - h\ln h\right)\right] = \phi'(k^*) + c\left[\frac{\phi(k^*) - k^*\phi'(k^*)}{-ck^*\ln h + \phi(k^*)}\right]. \quad (13)$$

This condition reflects that a firm optimally chooses its product position to balance costs and benefits. The left-hand side captures two incremental costs that arise when a firm moves its product toward a higher position. First, the term $g'(k^*)$ indicates that the firm must pay a larger entry cost. Second, producing a superior product requires solving more-complex problems. If the agents' knowledge levels remain unchanged, the probability of successfully solving problems would decrease. Thus, to maintain the same probability of success, a firm facing more-complex problems must pay an additional learning cost for the entire team. The second term captures this cost.

On the right-hand side of (13), the first term is the incremental market value of a product at a higher position. The second term captures the indirect effect of adjusting the product position on a firm's organizational structure. For instance, when $\phi(k)$ is a convex function (i.e., $\phi(k) < k\phi'(k)$), $c[(\phi(k^*) - k^*\phi'(k^*))/(-ck^*\ln h + \phi(k^*))] < 0$. This negative term induces the firm to choose a product at a higher position. Intuitively, the convexity of the price function entails a greater marginal value of moving toward a higher position; thus, a firm has a strong incentive to acquire more knowledge for its agents that in turn allows the firm to enlarge its span of control. This organizational adjustment reinforces the choice of a higher market position. Conversely, when $\phi(k)$ is a concave function (i.e., $\phi(k) > k\phi'(k)$), the marginal value of moving toward a higher position decreases. If the firm chooses a higher product position, the benefit is not enough to justify the cost of acquiring additional knowledge to cope with more-complex problems. Employees then become less competent and will refer unsolved problems more frequently; the extent of knowledge leverage decreases. Anticipating such a consequence, the firm may prefer a product at a lower position to avoid the costs of acquiring and integrating knowledge.

The previous analysis has important implications for the strategic fit between a firm's product positioning k and span of control s. The position of product indicates how much unit value a firm can create by solving customers' problems. The span of control indicates how efficiently a firm can extract economies of scale in the utilization of knowledge. The former choice emphasizes value creation, and the latter choice emphasizes production efficiency. The optimal fit between the two choices precisely reflects how a profit-maximizing firm resolves the conflict between value creation and production efficiency. Figure 3 illustrates this basic idea. In the (s, k) space, the PP curve depicts the iso-profit function such that a firm can attain the same level of profitability through different combinations of (s, k). It slopes downward because increasing one variable while holding the other constant would generate more profit, as seen from (12). The CC curve depicts the cost function because of the communication constraint faced by a manager who specializes in helping workers with unsolved problems. It slopes upward because workers producing a superior product face more-complex problems and ask for help more frequently; thus, the span of control is reduced.[12] The tangent point between the two curves—the pair (s^*, k^*)—pins down the optimal fit between a firm's product position and span of control.

Figure 3 **Optimal Matching Between Product Choice and Organizational Structure**



Ideally, a firm wants to simultaneously choose a superior product and a large span of control, as the firm can thus create greater value and appropriate this value by improving the efficiency of production. I refer to such a combination of product choice and organizational structure as strategic complements (the solid arrow line in Figure 3). However, to deploy

[12] I do not explicitly derive the expression for the CC curve in Figure 3, because the derivation is mathematically involved. However, the shape of the CC curve can be seen from the constraint in Equation (1) in §2.

such a strategy, a firm must be capable to acquire substantial knowledge for its employees, maintain a balanced distribution of knowledge across hierarchical layers, and effectively communicate knowledge across layers. If a firm lacks these capabilities, it must compromise between value creation and production efficiency. When a firm pairs a high-end product with a narrow span of control or pairs a low-end product with a wide span of control (the dotted arrow lines in Figure 3), I refer to either of these combinations as strategic substitutes.

4. Applications: Building Knowledge-Based Capabilities

In the previous model, a firm optimally chooses its product position and hierarchical structure in response to two parameters: c, the efficiency of learning knowledge; and h, the efficiency of communicating knowledge. These two parameters reflect a firm's intangible ability to acquire and integrate knowledge—a core ability in the development of a firm's competitive advantage, as argued by Nonaka and Takeuchi (1995) and Grant (1996a), among many others. In this section, I articulate the effects of changes in the two parameters on the alignment between product choice and organizational structure. To facilitate the analysis, I make the following assumption to replace Assumption 2.

Assumption 3. (a) *The price of a product with position k is $\phi(k) = \alpha k - t$, with $\alpha > 0$ and $t > 0$*; (b) *The cost function $g(k)$ is sufficiently convex*: *$g''(k) > 0$ and is sufficiently large relative to α*.

Part (a) of this assumption specifies a linear pricing rule. The two parameters α and t capture the market condition under which a firm creates value by offering a unit of product to its customers. Specifically, α measures the marginal market value of increasing product position, and t is an industry-wide transfer from firms to consumers. Because of linear pricing, the assumption on the cost function in part (b), as a special case of Assumption 2, is necessary to guarantee the existence of an interior solution. Assumption 3 also allows us to gauge other pricing forms with the linear pricing rule.

4.1. Learning Ability and Knowledge Acquisition

I take a problem-solving approach in which an individual worker acquires knowledge to solve problems. Knowledge is learned to overcome the cognitive limitation in problem solving, and the cost of knowledge acquisition is primarily the time that a problem solver expends to identify customers' demand and match solutions to problems. A firm can obtain better learning ability through two channels. First, at the individual level, "all the learning occurs in the human mind" (Simon 1991). A firm can improve individual workers' efficiency to acquire knowledge by superior hiring processes that allow for the selection of more-competent employees, better training and mentoring of employees, and the adoption of new technology and methods that enhance workers' ability to extract information, identify problems, reproduce knowledge, and learn new solutions. Second, at the organizational level, a firm can structure organizational routines to facilitate the matching of solutions to problems and develop organizational culture to motivate employees to acquire knowledge. This second aspect is particularly important when knowledge required to solve problems is tacit and difficult to measure, as it is a source of nonimitable organizational capability (e.g., Kogut and Zander 1992). In this paper, I regard all sorts of improvements in a firm's learning ability as a reduction of parameter c in the model.

In the model, for a given product position, an improvement in a firm's learning ability encourages both managers and workers to acquire more knowledge, which generates a higher probability of successfully solving problems. Moreover, with more knowledge, lower-layer agents (workers) ask for help from upper-layer agents (managers) less frequently. Thus, a manager can use her knowledge more intensively by helping more subordinate workers; this improved knowledge leveraging induces the manager to acquire further knowledge. Having acquired more knowledge, the firm may consider choosing a superior product. Such a new product choice will generate two opposing effects on the firm's incentive to acquire knowledge. On the one hand, the increased unit value of the new product encourages knowledge acquisition; on the other hand, the increased complexity of problems encountered in production discourages knowledge acquisition. This trade-off determines the optimal level of knowledge acquisition and the distribution of knowledge across hierarchical layers, which in turn determine the firm's product choice in conjunction with its hierarchical structure. In the following proposition, I demonstrate a case in which greater learning ability enables a firm to exploit strategic complementarity between product positioning and hierarchical structure.

Proposition 1. *Under Assumptions* 1 *and* 3, *an improvement in learning ability* (*i.e., a decrease in* c) *induces a firm to enlarge its span of control and choose a product at a higher position*, *associated with more knowledge acquired for both managers and workers*.

Proof. See the appendix. □

Proposition 1 shows that the ability to acquire knowledge is crucial for a firm to sustain a high-end market position. Moving to a higher product position without the ability to acquire more knowledge not only decreases the quality of the product, as measured by the probability of solving problems, but also reduces the intensity of using existent knowledge. This provides an explanation for the fact that high-end law, consulting, and financial service firms expend great efforts to hire experts with in-depth industrial knowledge, to develop research-related manpower or knowledge centers, and to cooperate with universities.

The linear pricing rule in Assumption 3 is an important, though not necessary, condition for achieving complementarities among a firm's knowledge acquisition, product positioning, and span of control. During the offering of a higher-end product, workers encounter more uncommon problems, and the probability of solving these problems is reduced. A firm's incentive to acquire knowledge would decrease if the market value of moving to a higher position is not great enough. Linear pricing ensures that the marginal value of a product does not decline with its market position. This offsets the firm's reduced incentive to acquire knowledge because of the need to handle more-complex problems. In the real business world, I certainly do not expect the linear pricing rule to apply to every circumstance. For example, when a superior product does not proportionally increase consumers' willingness to pay, a firm's pricing rule may be a concave function with respect to its product position. In such a case, an improvement in a firm's learning ability may induce it to choose a lower product position, because it may be more valuable for the firm to focus on production efficiency by expanding the span of control, forgoing the high unit value.

4.2. Communication Ability and Knowledge Integration

In the field of epistemology, one key property that distinguishes different types of knowledge is codifiability (e.g., Polanyi 1958, 1966). Knowledge is more codifiable when it is easier to express, store, divide, and transfer. The distinction of knowledge based on its codifiability is an important element in the knowledge-based theory of organizational capabilities (e.g., Nonaka 1994, Grant 1996b), because the codifiability of knowledge determines the cost of integrating knowledge within the firm. This distinction is particularly relevant to the PSFs. As well documented (e.g., Halliday 1987, Morris and Empson 1998, Malhotra and Morris 2009), the codifiability of the knowledge involved in solving customers' problems distinguishes among different professional services. For example, the knowledge required for engineering services is usually based on scientific formulae and data and thus is easy to code and transfer. In contrast, the knowledge required for the provision of legal services relies heavily on a personal understanding of institutions and personal judgment and is difficult to code and transfer.

The current model captures the notion of the codifiability of knowledge by parameter h, which measures the efficiency of referring problems and communicating knowledge among agents. In this subsection, I analyze how a change in this parameter, referred to as a firm's communication ability, affects a firm's strategic match between product choice and organizational structure.

Suppose that a firm improves its communication ability, whereas its learning ability remains constant. How should the firm adjust its strategic match between product position and the span of control? Intuitively, better communication improves the efficiency of referring problems within the organization, and a manager can expand her span of control by helping more workers. This increased span of control encourages the manager to acquire more knowledge, because that knowledge will be used more intensively. With more able managers, a firm then has an incentive to choose a superior product. The result is summarized in the following proposition.

Proposition 2. *Under Assumptions* 1 *and* 3, *an improvement in the ability to communicate knowledge* (*i.e., a decrease in* h) *induces a firm to enlarge its span of control and choose a product at a higher position, associated with more knowledge acquired for managers.*

Proof. See the appendix. □

Apparently, the effect of an improvement in communication ability on a firm's product position and span of control is similar to the effect of an improvement in learning ability. However, the mechanisms underlying these effects are different. Improved communication directly affects a manager's span of control, which then improves agents' incentive to acquire knowledge, whereas improved learning ability directly increases the level of knowledge acquired by agents, which in turn allows a manager to maintain a larger span of control. Because of this difference, an improvement in communication does not necessarily increase workers' knowledge, whereas an improvement in learning ability unambiguously does. In other words, an improvement in communication ability may enhance the asymmetry of knowledge distribution across hierarchical layers. This implies that a firm's expansion is less constrained by the knowledge of individuals at low layers of the hierarchy.

The implication of Proposition 2 corresponds to a codification strategy that has been widely practiced by PSFs. For example, Hansen et al. (1999) noted that some top consulting firms, such as Anderson Consulting and Ernst and Young, used a "people-to-documents" approach to "develop an electronic document system that codifies, stores, disseminates, and allows the reuse of knowledge" (p. 108). This approach enables a firm to use large teams with a high associates-to-partners ratio and intensify the use of its top talent's knowledge. Many law firms also adopt a similar codification strategy that uses process and technology to identify recursive patterns in legal forms and routinize legal practices. Legal scholars (e.g., Susskind 2008) argue that such a codification strategy, owing to technological change, has transformed legal work from customized (e.g., court room practice), to systematized (e.g., a document assembly system), and to commoditized (e.g., IT-based legal product). This transformation provides an explanation for the continuous increase in the associates-to-partner ratio among top U.S. law firms over the last two decades (Henderson 2014).

The cross-sectional implications of Proposition 2 are particularly relevant to the management of different types of PSFs. Following the slogans used in PSFs (Delong et al. 2007), I distinguish the following three types of professional services: (1) "procedure" (low h), (2) "gray hair" (intermediate h), and (3) "rocket science" (high h). In the procedure business, knowledge input in production is primarily data based and highly codifiable. The cost of communicating this type of knowledge is low. It is more efficient to use a flat organizational structure with a manager supervising a large number of qualified associates. Substantial investments in human capital are unnecessary. This is the situation in engineering service and accounting firms, which have a large associate-to-manager ratio and employees who tend to be newly minted college graduates. In the gray hair business, knowledge input is largely based on experience but often is reusable. The cost of communication is higher than in the procedure business. To sustain high-quality services, a firm involved in the gray hair business requires a smaller associate-to-manager ratio. Organizational capabilities hinge on the complementarity between managers and associates. Therefore, investments in human capital are important. Finally, in the rocket science business, knowledge input is often based on experimentation and highly innovative. Because this type of knowledge is highly costly to codify and communicate, it is difficult to leverage the talent of stellar managers. This partially explains why business-strategy consulting firms and Wall Street law firms have a much lower associate-to-partner ratio than do other PSFs.

4.3. Advances in Information and Communication Technology

Numerous studies have shown that the advances in information and communication technology (ICT) have had a profound impact on firms' workplace organization and managerial practices.[13] The literature stresses that ICT enters the production function as a substitute for unskilled labor but a complement for skilled labor, which induces changes in organizational structure that permit more efficient management of skilled labor (e.g., Brynjolfsson and Hitt 1996, Caroli and Van Reenen 2001, Bresnahan et al. 2002, Bertschek and Kaiser 2004). Recently, several empirical studies have drawn attention to the impact of ICT on the production process (e.g., Bartel et al. 2007) and changes in managerial function (e.g., Guadalupe et al. 2014). These mechanisms are of great importance for manufacturing and retailing firms, in which production involves sizeable physical assets, diverse activities, and complex managerial functions.

Our model provides a new mechanism to examine the impact of changes in ICT on organizational structure and product choice in knowledge-intensive firms. As documented by Ofek and Sarvary (2001) and Banker et al. (2002), almost all leading accounting firms have adopted audit software and knowledge-sharing applications. Many consulting firms use computer-mediated knowledge systems and digitized data bases. Engineering services have heavily relied on computer-aided design, data documentation, and professional software. One major consequence of adopting these technologies is that knowledge becomes much easier to extract, code (digitize), store, cross-reference, and transfer. Corresponding to the theoretical model, the improvement of ICT allows firms to reduce both the costs of acquiring and communicating knowledge. Propositions 1 and 2 immediately lead to the following result.

Corollary 1. *Under Assumptions* 1 *and* 3, *advances in information and communication technology lead to the following effects on a firm's product choice and organizational structure*:

Product Upgrading. *Firms tend to choose superior products that create greater per-unit value*.

Skill-Biased Organizational Change. *Firms tend to enlarge the span of control and invest more in the human capital of managers*.

This corollary yields two readily testable hypotheses regarding the impact of advances in ICT on product innovations and organizational changes that has been the focus of a large number of empirical studies on manufacturing firms. Systematic evidence using data from knowledge-intensive firms such as PSFs is relatively limited and would be particularly welcome to test these two hypotheses. In Corollary 1, I do not distinguish the effects of different ICTs—those that facilitate knowledge acquisition and those that improve knowledge communication—on the knowledge distribution across layers in hierarchies. Empirically, it would be interesting to draw such a distinction, which has important implications for the managerial practices of manufacturing firms, as recognized by Bloom et al. (2014).

[13] See Brynjolfsson and Milgrom (2013) for a recent survey and the references therein.

5. Extensions

The results in §4 are pertinent to some of the model's specific assumptions. Nevertheless, the insight of the model is more general than these results. In this section, I discuss several extensions of the model to address more-sophisticated managerial issues.

5.1. Task Complexity and Codifiability of Knowledge

In the model, a firm's communication ability, indicated by parameter h, does not vary with the market position of a product, k. This assumption, treating communication ability as a firm's innate property, allows us to articulate the effect of changing communication ability on the alignment between product positioning and organizational structure. However, the independence between a firm's communication ability and product position may not be generally true when one thinks of knowledge codifiability as a key determinant of a firm's communication ability (recall the discussion in §4.2). For instance, in the professional service industry, customers in the high-end segment tend to demand complex problems that require innovative and customized solutions. The knowledge encompassed in these solutions is often tacit in nature and is difficult to communicate. In contrast, in the low-end market segment, solutions to consumers' problems require only standardized knowledge that can be easily coded and communicated.

The introduction of a positive correlation between a firm's communication ability and product position does not change the basic insights of the model, but will affect some comparative statics results. Essentially, the presence of this correlation exacerbates the conflict between selecting a superior product and employing a broader span of control, because moving toward a higher product position not only requires superior knowledge input but also reduces the efficiency of communicating knowledge. To resolve this conflict, a firm must acquire sufficient knowledge for workers so that managers can be protected from frequent interactions with a small number of workers. Thus, for firms with low learning ability (i.e., high c), an improvement in their learning ability may induce them to choose a superior product but a narrower span of control, whereas an improvement in their communication ability may induce them to choose a wider span of control but a lower-end product position. The results in Propositions 1 and 2 that improvements in firms' learning and communication ability enhance the strategic complementarity between product positioning and the span of control hold only for those firms with sufficiently high learning ability (i.e., low c).

The extension in this subsection has important implications for the management of PSFs whose services simultaneously fall into two classifications that I have discussed: one based on the complexity of tasks and the other based on the codifiability of knowledge. For PSFs involving commodity tasks and procedure knowledge (small k and small h), solving customers' problems does not require advanced knowledge, and managers' knowledge can be easily leveraged. Thus, it is optimal for PSFs of this kind to match a low-end product position to a large span of control without substantial investments in workers' human capital. For PSFs involving process tasks and gray hair knowledge (medium k and medium h), the prices of their services are modest, and firms' profits come from high-quality and repetitive services. A knowledge-based hierarchy with high-quality employees and a sizeable span of control to effectively leverage top managers' knowledge is crucial to achieve efficiency. For PSFs involving customization tasks and rocket science knowledge (large k and large h), solving customers' problems requires exceptional knowledge. Although the high prices of these services provide incentives for firms to acquire substantial knowledge, the considerable communication cost limits the extent of leveraging knowledge. Thus, the span of control is narrow, and business success heavily relies on a few superstars who have acquired remarkable knowledge.

Given the correlation between task complexity and knowledge codifiability, the advances in ICT are likely to have different impacts on the organizational structure of PSFs in different segments. Numerous evidence suggests that the impact of ICT on knowledge transfer and information communication is most effective when knowledge is codifiable. In contrast, when knowledge is tacit, the effect of ICT on knowledge transfer within organizations is limited. Therefore, the advances in ICT are likely to substantially enlarge the span of control of PSFs in the low-end segment that perform commodity tasks, but barely affect the span of control of PSFs in the high-end segment that perform customization tasks. This differentiated effect can amplify the difference in the hierarchical structure of PSFs across market segments.

5.2. Endogenous Communication Ability

In this paper, I treat a firm's communication ability as a parameter that is subject to exogenous shocks. For example, in §4.3, I show that economy-wide advances in ICT improve firms' communication ability that leads firms to change their organizational structure and product choice. For an individual firm, however, its communication ability, or more generally its ability to integrate knowledge, is likely to be endogenous. For instance, aiming to select a superior product position or to increase the extent of knowledge leverage, a firm may invest more heavily in ICT. Within the modeling framework in §4, it is straightforward to show that a firm's investment in ICT, product choice, span of control, and the acquisition of knowledge for top manager are complementary to one another. A more sophisticated approach to endogenize a firm's communication ability is to explicitly model the communication process within a firm. One such example is Cremer et al. (2007), which models how a firm designs its codes to facilitate communication.

Endogenous communication ability is particularly important when knowledge used to solve problems is tacit, for instance, the knowledge used in strategic business consulting and legal services for sophisticated financial transactions. In these situations, effective communication of knowledge within organizations requires a "people-to-people" approach, instead of a "people-to-documents" approach. How people interact with each other to create and communicate knowledge is largely an institutional problem and is endogenous to a firm's culture. An extension to endogenize a firm's efficiency in communicating tacit knowledge is naturally connected to the study of organizational culture, either treated by economists as the general principle and the means for coordination (Kreps 1990) and a common language of coding (Cremer 1993), or regarded by strategy scholars as an element of routines (Nelson and Winter 1982) and a system of shared experience and common perspectives (Nonaka 1994).

5.3. Unbalanced Knowledge Distribution

In the model, I implicitly regard a firm's learning ability as a unidimensional capability that applies equally to workers at all layers. Thus, improved learning ability induces all agents to acquire further knowledge in a balanced manner so that a firm can choose a high-end product without reducing its span of control. The general insight is that a balanced distribution of knowledge across hierarchical layers is crucial for a firm to exploit the complementarity between its product position and span of control. Without being protected by competent subordinates, even the most-talented managers, who can solve the most complex and valuable problems, may be outperformed by less-talented managers supported by competent subordinates.

In reality, not all firms can achieve a balanced distribution of knowledge across hierarchical layers because more-advanced knowledge is likely more costly to acquire. I can adapt the current model, in which individual workers acquire knowledge internally and firms bear the costs of knowledge acquisition, to address the unbalanced distribution of knowledge within firms. For instance, in industries where formal education constitutes a core part of workers' human capital, firms often acquire knowledge by hiring talented workers who have acquired their knowledge externally; in this case, the costs of acquiring knowledge are the wages paid to workers. Because labor market conditions differ across education levels and expertise areas, the wages for workers across hierarchical layers are likely to differ, resulting in an unbalanced distribution of knowledge within firms. CEOs of PSFs often complain about the difficulty of hiring competent mid-level managers because of a labor-force shortage of well-educated workers or the inefficiency of the managerial labor market. Without a strong middle layer, a firm can expand only along an unbalanced path (recall Figure 3), with a firm's product position and the span of control being strategic substitutes. This is part of the "getting stuck in the middle" phenomenon, in which a firm fails to move into the high-end market segment and to keep pace with market growth. Here, "getting stuck in the middle" happens because firms are stuck in building a strong middle layer.

Consistent with the model implications, several managerial practices are widely used in top PSFs to mitigate the unbalanced distribution of knowledge within firms. Top PSFs often locate their offices in areas where elite universities cluster. This approach enables a firm to quickly search and identify high-quality employees who have the potential to become mid-level managers within the first several years of their career. Almost all top PSFs make enormous investments in training their junior employees, even if they anticipate high turnover. The value of such a large investment in human capital is not only to improve new employees' ability to solve customers' problems, but also to facilitate the leveraging of top managers' talent.

5.4. Heterogeneous Firm Performance

Using detailed within-firm data from specific industries, a number of recent empirical studies (e.g., Pierce 2012, Bennett 2013, Natividad 2014) show that organizational structure, most notably the form of asset integration, has a significant impact on knowledge acquisition and knowledge transfer among workers and is thus an important factor that accounts for the observed heterogeneous firm performance. These findings are broadly consistent with the view that firms can design efficient organizational structure to optimize knowledge flow and integration for the achievement of superior performance (e.g., Nonaka 1994, Kogut and Zander 1996, Nickerson and Zenger 2004). Along this line of research, our model provides a formal framework to address the relationship among knowledge management, internal organizational structure, and firm performance. In particular, the current model can be extended to a market equilibrium setting, in which firms with heterogeneous capabilities of learning and communicating knowledge compete in the market. In such a model, the complementarities among knowledge acquisition, the span of control, and product positioning amplify a firm's initial advantage and enhance resource allocation across firms. Consequently, substantial differences in firms' performance appear even if their initial ability differs only slightly. Under reasonable assumptions on market structure, the equilibrium stratifies firms into two groups: one group consists of more-productive firms whose product position and span of control are complementary; the other group consists of less-productive firms whose product position and span of control are substitutes. These results are empirically testable, as large-scaled employer-employee data are now available in various countries. The comparative statics analysis concerning the impact of ICT (§4.3) also provides a potential source of identification for empirical studies.

Provided that the product choice element in our model explicitly relates a firm's internal organization to the product market condition, another potential extension of the model is to study firms' strategy and performance over the product life cycle. The basic idea is sketched as follows. At the introduction stage, firms enter the market with innovative products; competition is weak and the products' prices are high. Firms have great incentives to acquire further knowledge so as to enhance the market value of their products. However, at this nascent stage, the knowledge required for production is often new and difficult to communicate among individuals; thus, it is difficult to leverage the knowledge possessed by a limited number of individuals within the firm. According to the model, the optimal strategy for a firm at the introduction stage is to focus on the creation of high-value products and the acquisition of new knowledge, forgoing the intensive use of existing knowledge. At the development stage during which the products become more mature, competition increases, and the marginal value of product upgrades begins to decrease. Simultaneously, the knowledge required for production becomes more familiar to workers and easier to codify. A firm's optimal strategy at this stage is to exploit the complementarity between value creation and knowledge leverage. Finally, at the mature stage when products become standardized and imitable, competition is fierce and the profit margin is low. It is optimal for a firm to focus on improving the codifiability and integration of knowledge, forgoing value creation through product innovation. Such an extension offers an explanation for the transformation of strategy and organizational structure from "exploration" to "exploitation" over the product life cycle, as documented in previous studies (e.g., Hofer 1975, Anderson and Zeithaml 1984).

6. Conclusion

The fit of important elements of organization and of the organization as a whole with its strategy and environment is a central topic in strategy research. In this paper, I formulate a particular aspect of the fit between strategy and structure—the fit between the choice of vertically differentiated products and the structure of vertically integrated hierarchies. This specific focus captures the essential features of knowledge-based production and delivers far-reaching implications for the strategic management of knowledge-intensive firms, particularly PSFs. The major managerial implications that I have discussed are collected in Table 1. These implications are applicable to the management of other economic activities in which knowledge and expertise are major inputs, for instance, R&D, product design, and creative businesses.

Despite our focus on knowledge-intensive firms, the economic principles underlying the model are highly relevant to many other settings. First, organizational structure plays an important strategic role in mediating between available resources and product choice. The premise is that producing more-valuable products requires the acquisition and use of more-expensive resources—in our model, superior knowledge. When an organizational structure is designed to facilitate the utilization of those resources, a firm increases its incentive to acquire more-expensive resources and to choose more-valuable products. This principle is relevant to settings in which production is innovation-oriented and the resources critical to production are difficult to acquire. Second, to achieve superior performance, a firm should not only align its organizational structure with its product strategy, but also actively explore the complementarity between product choice and organizational structure. From a dynamic perspective, this complementarity is an important driver for continuous product innovations and firm growth. I stress the importance of building the knowledge-based capabilities—the ability to acquire and integrate knowledge, precisely because these capabilities are crucial for exploiting the complementarity between product choice and organizational structure. Third, the model establishes a theory of strategy and structure that is neither "structure follows strategy," as in Chandler's (1962) dictum, nor is it "strategy follows structure" (e.g., Bower 1970, Burgelman 1983). Instead, strategy and structure are jointly determined in equilibrium in response to environmental changes. This view of "strategic fit" is particularly important for firms' long-term strategic choices, as stressed by Roberts and Saloner (2013).

Modeling a fit of some elements of strategy and organization entails the simplification and omission of other elements. One limitation of this paper is that the

Table 1 The Fit Between Strategy and Structure in Professional Service Firms

Panel A. Classification by product position and task complexity (parameter k)

Product position	Commodity (small k)	Business process (medium k)	Customized business (large k)
Profit margin	Low	Intermediate	High
Task complexity	Low	Intermediate	High
Human capital	Low investment	Modest investment	High investment
Degree of knowledge leverage	Large	Intermediate	Depend on the quality of human capital

Panel B. Classification based on codifiability of knowledge (parameter h)

Attributes of knowledge input	Procedure (small h)	Gray hair (medium h)	Rocket science (large h)
Codifiability of knowledge	Data-based standard knowledge; highly codifiable	Experienced-based tacit knowledge; partially codifiable	Experimentation-based innovative knowledge; hardly codifiable
Human capital	Low investment	Modest investment	High investment
Degree of knowledge leverage	Large	Intermediate	Small

Panel C. Classification by task complexity and knowledge codifiability

Composite business type	Commodity and procedure	Business process and gray hair	Customized business and rocket science
Organizational structure	Large scale hierarchical structure	Knowledge hierarchy	Minimum hierarchy
Role of management	Manage by administration	Manage by knowledge transmission and direction	Manage by empowerment
Competitive advantage	Efficiency-based competency; low cost and fast delivery of services	Experience-based competency; high-quality and repetitive services	Expertise-based competency; creative and unique services

model of organizational structure is confined within the team-theory tradition (e.g., Marschak and Radner 1972), in which agents in an organization share the same objective, thus muting the incentive problems. Another limitation of this modeling framework is its inadequacy to incorporate some important epistemic properties of productive input, other than the codifiability of knowledge, into the analysis. With regard to the issue of product choice, this model focuses on the selection of vertically differentiated products, leaving the important issues of product scopes under explored. These limitations raise challenging but potentially fruitful opportunities for future research.

Acknowledgments

The author is very grateful to Luis Garicano for his guidance. The author thanks department editor Bruno Cassiman, the associate editor, and an anonymous referee for their comments and suggestions. The author also thanks Victor Bennett, Nan Jia, Tony Marino, John Matsusaka, Heikki Rantakari, Feng Zhu, and seminar participants at the Center for Research in Economics and Strategy Conference (Washington University in St. Louis), the London School of Economics, and the University of Southern California for their helpful comments.

Appendix

Proof of Lemma 1

Define the first derivative of the optimization problem (12):

$$\Omega \equiv \phi'(k) - c\left[1+\ln\left(\frac{h\phi(k)}{kc} - h\ln h\right)\right] - c\left[\frac{-\phi(k)+k\phi'(k)}{-ck\ln h+\phi(k)}\right] - g'(k).$$

Then,

$$\begin{aligned}\frac{d\Omega}{dk} &= \phi''(k) - g''(k) - c\frac{d\ln[\phi(k)/(kc) - \ln h]}{dk} \\ &\quad - c\frac{d[[-\phi(k)+k\phi'(k)]/(-ck\ln h+\phi(k))]}{dk} \\ &= \phi''(k) - g''(k) - c\frac{\phi'(k)-\phi(k)/k}{\phi(k)-ck\ln h} \\ &\quad - c((k\phi''(k)[\phi(k)-ck\ln h] \\ &\quad - [\phi'(k)-\phi(k)/k][k\phi'(k)-ck\ln h]) \\ &\quad \cdot([\phi(k)-ck\ln h]^2)^{-1}).\end{aligned}$$

For notational simplicity, let $x = ck\ln h$. Then,

$$\begin{aligned}\frac{d\Omega}{dk} &= \phi''(k) - g''(k) \\ &\quad - \frac{c}{[q(k)-x]^2}[(\phi'-\phi/k)(\phi-x) + k\phi''(\phi-x) \\ &\qquad - (\phi'-\phi/k)(k\phi'-x)] \\ &= -g''(k) + \phi''(k) \\ &\quad + c\frac{[k\phi'(k)-\phi(k)]^2 - \phi''(k)k^2[\phi(k)-x)]}{k[\phi(k)-x]^2} \\ &= -((k[\phi(k)-x][g''(k)(\phi(k)-x) - \phi''(k)(\phi(k)-x-ck)] \\ &\quad - c[k\phi'(k)-\phi(k)]^2)\cdot(k[\phi(k)-x]^2)^{-1}). \qquad (14)\end{aligned}$$

Given that $z_s^* = k\ln(\phi(k)/(ck) - \ln h) > 0$, $\phi(k) - x - ck = \phi(k) - ck(1+\ln h) > 0$. Since

$$\begin{aligned}&g''(k) - \phi''(k) > 0, \\ &g''(k)(\phi(k)-x) - \phi''(k)(\phi(k)-x-ck) \\ &\quad > [g''(k) - \phi''(k)][\phi(k)-x] > 0.\end{aligned}$$

When $g''(k) - \phi''(k)$ is sufficiently large, the first term in the numerator of the (14) dominates the second term. Then, $d\Omega/dk < 0$. The second-order condition of the maximization problem is satisfied. □

Proof of Proposition 1
Under Assumptions 1 and 3, the first-order condition of the optimization problem is

$$\begin{aligned}\Phi &\equiv \alpha - c\left[1+\ln\left(\frac{h(\alpha k-t)}{kc} - h\ln h\right)\right] - c\left[\frac{t}{-ck\ln h+\alpha k-t}\right]\\ &\quad - g'(k)\\ &= 0.\end{aligned}$$

Note that $\alpha - c$ must be positive to satisfy the first-order condition. Then,

$$\begin{aligned}\frac{\partial\Phi}{\partial c} &= -\left[1+\ln\left(\frac{h(\alpha k-t)}{kc} - h\ln h\right)\right] - c\frac{-((\alpha k-t)/(kc^2))}{(\alpha k-t)/(kc)-\ln h}\\ &\quad - t\frac{[-ck\ln h+\alpha k-t]+ck\ln h}{[-ck\ln h+\alpha k-t]^2}\\ &= -(\ln s+\ln h) - \big[((\alpha k-t-ck\ln h)^2-(\alpha k-t)\\ &\quad \cdot(\alpha k-t-ck\ln h)+t(\alpha k-t))\cdot((\alpha k-t-ck\ln h)^2)^{-1}\big]\\ &= -(\ln s+\ln h) - \frac{-ck\ln h(\alpha k-t-ck\ln h)+t(\alpha k-t)}{(\alpha k-t-ck\ln h)^2} < 0.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{dk}{dc} &= -\frac{\partial\Omega/\partial c}{\partial\Omega/\partial k} < 0;\\ \frac{ds}{dc} &= \frac{d[\phi(k)/(ck) - \ln h]}{dc} = -\frac{\alpha k-t}{c^2k} + \frac{t}{ck^2}\frac{dk}{dc} < 0;\\ \frac{dz_s^*}{dc} &= \frac{d[k\ln(\phi(k)/(ck)-\ln h)]}{dc} = \frac{dk}{dc}\ln s + \frac{k}{s}\frac{ds}{dc} < 0;\\ \frac{dz_w^*}{dc} &= \frac{d[k\ln(\phi(k)/(ck)-\ln h)+k\ln h]}{dc}\\ &= \frac{dk}{dc}(\ln s+\ln h) + \frac{k}{s}\frac{ds}{dc} < 0. \quad \square\end{aligned}$$

Proof of Proposition 2
Without invoking Assumption 3, the first-order condition of the optimization problem (12) is

$$\begin{aligned}\Omega &\equiv \phi'(k) - c\left[1+\ln\left(\frac{h\phi(k)}{kc} - h\ln h\right)\right] - c\left[\frac{-\phi(k)+k\phi'(k)}{-ck\ln h+\phi(k)}\right]\\ &\quad - g'(k) = 0.\end{aligned}$$

Then,

$$\begin{aligned}\frac{\partial\Omega}{\partial h} &= -\frac{c}{h} - c\frac{-1/h}{\phi(k)/(kc)-\ln h} - c\frac{-\phi(k)+k\phi'(k)}{-[-ck\ln h+\phi(k)]^2}\left(-\frac{ck}{h}\right)\\ &= -\frac{c}{h}\frac{1}{\phi(k)-ck\ln h}\\ &\quad \cdot\left[\phi(k)-ck\ln h-ck+\frac{ck[-\phi(k)+k\phi'(k)]}{-ck\ln h+\phi(k)}\right].\end{aligned}$$

A positive z_s^* requires $\phi(k) - ck\ln h - ck > 0$. By Assumption 3, $-\phi(k) + k\phi'(k) = t > 0$ and thus $\partial\Omega/\partial h < 0$. Then,

$$\begin{aligned}\frac{dk}{dh} &= -\frac{\partial\Omega/\partial h}{\partial\Omega/\partial k} < 0;\\ \frac{ds}{dh} &= \frac{d[\phi(k)/(ck)-\ln h]}{dh} = \frac{1}{c}\underbrace{\frac{\phi'(k)k-\phi(k)}{k^2}}_{>0}\underbrace{\frac{dk}{dh}}_{<0} - \frac{1}{h} < 0;\\ \frac{dz_s^*}{dh} &= \frac{d[k\ln(\phi(k)/(ck)-\ln h)]}{dh} = \frac{dk}{dh}\ln s + \frac{k}{s}\frac{ds}{dh} < 0. \quad \square\end{aligned}$$

References

Anderson C, Zeithaml C (1984) Stage of the product life cycle, business strategy, and business performance. *Acad. Management J.* 27(1):5–24.

Argyres N, Felin T, Nicolai F, Zenger T (2012) Organizational economics of capability and heterogeneity. *Organ. Sci.* 25(6): 1213–1226.

Banker R, Chang H, Kao Y (2002) Impact of information technology on public accounting firm productivity. *J. Inform. System* 16(2):209–222.

Bartel A, Ichniowski C, Shaw K (2007) How does information technology affect productivity? Plant-level comparisons of product innovation, process improvement and worker skills. *Quart. J. Econom.* 122(4):1721–1758.

Bennett V (2013) Organization and bargaining: Sales process choice at auto dealerships. *Management Sci.* 59(9):2003–2018.

Bertschek I, Kaiser U (2004) Productivity effects of organizational change: Microeconometric evidence. *Management Sci.* 50(3): 394–404.

Bloom N, Garicano L, Sadun R, Van Reenen J (2014) The distinct effects of information technology and communication technology on firm organization. *Management Sci.* 60(12): 2859–2885.

Bolton P, Dewatripont M (1994) The firm as a communication network. *Quart. J. Econom.* 109(4):809–839.

Bower JL (1970) *Managing the Resource Allocation Process* (Graduate School of Business Administration, Harvard University, Boston).

Bresnahan TF, Brynjolfsson E, Hitt LM (2002) Information technology, workplace organization, and the demand for skilled labour: Firm-level evidence. *Quart. J. Econom.* 112(1):339–376.

Broderick M (2011) *The Art of Managing Professional Services* (Prentice Hall, Upper Saddle River, NJ).

Brynjolfsson E, Hitt LM (1996) Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Sci.* 42(4):541–558.

Brynjolfsson E,Milgrom P (2013) Complementarity in organizations. Gibbons R, Roberts J, eds. *The Handbook of Organizational Economics* (Princeton University Press, Princeton, NJ), 11–55.

Burgelman RA (1983) A model of the interaction of strategic behavior, corporate context, and the concept of strategy. *Acad. Management Rev.* 8(1):61–70.

Caroli E, Van Reenen J (2001) Skill biased organizational change? Evidence from a panel of British and French establishments. *Quart. J. Econom.* 116(4):1449–1492.

Chandler A (1962) *Strategy and Structure* (MIT Press, Cambridge, MA).

Coe E (2014) 4 elite New York law firms keep partner ranks lean. *Law 360* (March 23), http://www.law360.com/articles/520626/4-elite-new-york-law-firms-keep-partner-ranks-lean.

Cremer J (1993) Corporate culture and shared knowledge. *Indust. Corporate Change* 2(1):351–386.

Cremer J, Garicano L, Prat A (2007) Language and the theory of the firm. *Quart. J. Econom.* 122(1):373–407.

Delong T, Gabarro J, Lees R (2007) *When Professionals Have to Lead* (Harvard Business School Press, Boston).

Demsetz H (1988) The theory of the firm revisited. *J. Law, Econom., Organ.* 4(1):141–162.

Foss NJ (2006) *Strategy, Economic Organization, and the Knowledge Economy: The Coordination of Firms and Resources* (Oxford University Press, New York).

Garicano L (2000) Hierarchies and the organization of knowledge in production. *J. Political Econom.* 108(5):874–904.

Garicano L, Prat A (2013) Organizational economics with cognitive costs. Acemoglu D, Arellano M, Dekel E, eds. *Advances in Economics and Econometrics: Theory and Applications, Proceedings of the Tenth World Congress of the Econometric Society* (Cambridge University Press, New York), 342–390.

Garicano L, Wu Y (2012) Knowledge, communication, and organizational capabilities. *Organ. Sci.* 23(5):1382–1397.

Grant RM (1996a) Toward a knowledge-based theory of the firm. *Strategic Management J.* 17(S2):109–122.

Grant RM (1996b) Prospering in dynamically-competitive environments: Organizational capability as knowledge integration. *Organ. Sci.* 7(4):375–387.

Guadalupe M, Li H, Wulf J (2014) Who lives in the C-suite? Organizational structure and the division of labor in top management. *Management Sci.* 60(4):824–844.

Hansen MT, Nohria N, Tierney T (1999) What's your strategy for managing knowledge. *Harvard Bus. Rev.* 77(2):106–116.

Hayek F (1945) The use of knowledge in society. *Amer. Econom. Rev.* 35(4):519–530.

Halliday T (1987) Knowledge mandates: Collective influence by scientific, normative and syncretic professions. *British J. Sociol.* 36(3):421–47.

Henderson W (2014) From big law to lean law. *Internat. Rev. Law Econom.* 38:5–16.

Herrendorf B, Rogerson R, Valentinyi A (2013) Two perspectives on preferences and structural transformation. *Amer. Econom. Rev.* 103(7):2752–2789.

Hill CWL, Hoskisson RE (1987) Strategy and structure in the multiproduct firm. *Acad. Management Rev.* 12(2):331–341.

Hofer CW (1975) Toward a contingency theory of business strategy. *Acad. Management J.* 18(4):784–810.

Kauffman S (1993) *The Origins of Order* (Oxford University Press, New York).

Kogut B, Zander U (1992) Knowledge of the firm, combinative capabilities, and the replication of technology. *Organ. Sci.* 3(3): 383–397.

Kogut B, Zander U (1996) What firms do? Coordination, identity, and learning. *Organ. Sci.* 7(5):502–518.

Kreps D (1990) Corporate culture and economic theory. Alt J, Shepsle K, eds. *Perspectives on Positive Political Economy* (Cambridge University Press, Cambridge, UK), 90–143.

Lawrence PR, Lorsch JW (1967) Differentiation and integration in complex organizations. *Admin. Sci. Quart.* 12(1):1–47.

Lowendahl B (2005) *Strategic Management of Professional Service Firms* (Copenhagen Business School Press, Copenhagen).

Maister DH (1993) *Managing the Professional Services Firm* (Free Press, New York).

Malhotra N, Morris T (2009) Heterogeneity in professional service firms. *J. Management Stud.* 46:6:895–922.

Marschak J, Radner R (1972) *Economic Theory of Teams* (Yale University Press, New Haven, CT).

Miles RE, Snow CC, Meyer AD, Coleman HJ Jr (1977) Organizational strategy, structure, and process. *Acad. Management Rev.* 3(3):546–562.

Milgrom P, Roberts J (1990) The economics of modern manufacturing: Technology, strategy and organization. *Amer. Econom. Rev.* 83(3):511–529.

Milgrom P, Roberts J (1995) Complementarities and fit strategy, structure, and organizational change in manufacturing. *J. Accounting Econom.* 19(2–3):179–208.

Morris T, Empson L (1998) Organization and expertise: An exploration of knowledge bases and the management in accounting and consulting firms. *Accounting, Organ. Soc.* 23(5–6): 609–624.

Natividad G (2014) Integration and productivity: Satellite-tracked evidence. *Management Sci.* 60(7):1698–1718.

Nelson RR, Winter SG (1982) *An Evolutionary Theory of Economic Change* (Belknap Press, Cambridge, UK).

Nickerson JA, Zenger TR (2004) A knowledge-based theory of the firm—The problem-solving perspective. *Organ. Sci.* 15(6): 617–632.

Nonaka I (1994) A Dynamic theory of organizational knowledge creation. *Organ. Sci.* 5(1):14–37.

Nonaka I, Takeuchi H (1995) *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation* (Oxford University Press, New York).

Ofek E, Sarvary M (2001) Leveraging the customer base: Creating competitive advantage through knowledge management. *Management Sci.* 47(11):1441–1456.

Pierce L (2012) Organizational structure and the limits of knowledge sharing: Incentive conflict and agency in car leasing. *Management Sci.* 58(6):1106–1121.

Polanyi M (1958). *Personal Knowledge* (University of Chicago Press, Chicago).

Polanyi M (1966) *The Tacit Dimension* (Routledge & Kegan Paul, London).

Puranam P, Raveendran M, Krudsen T (2012) Organization design: The epistemic interdependence perspective. *Acad. Management Rev.* 37(3):419–440.

Radner R (1993) The organization of decentralized information processing. *Econometrica* 61(5):1109–1146.

Ribstein L (2010) The death of big law. Working paper, University of Illinois, Champaign.

Roberts J, Saloner G (2013) Strategy and organization. Gibbons R, Roberts J, eds. *The Handbook of Organizational Economics* (Princeton University Press, Princeton, NJ), 799–852.

Rumelt RP (1974) *Strategy, Structure, and Economic Performance* (Division of Research, Harvard Business School, Boston).

Simon HA (1962) The architecture of complexity. *Proc. Amer. Philos. Soc.* 106(6):467–482.

Simon HA (1991) Bounded rationality and organizational learning. *Organ. Sci.* 2(1):125–134.

Simon HA (1997) *Administrative Behavior*, 4th ed. (Free Press, New York).

Sloan A (1924) The most important thing I ever learned about management. *System: Management Bus.* 46(August):137–141.

Susskind R (2008) *The End of Lawyers? Rethinking the Nature of Legal Services* (Oxford University Press, London).

Teece DJ (1982) Towards an economic theory of the multiproduct firm. *J. Econom. Behav. Organ.* 3(1):39–63.

Van Zandt T (1999) Real-time decentralized information processing as a model of organizations with boundedly rational agents. *Rev. Econom. Stud.* 66:633–658.