Investment in Organization Capital*

Bruce Ian Carlin  
*UCLA Anderson School of Management*

Bhagwan Chowdhry  
*UCLA Anderson School of Management*

Mark J. Garmaise  
*UCLA Anderson School of Management*

JEL classification: D23, G34

Keywords: Organization Capital, Corporate Governance, Managerial Turnover, Executive Compensation, Mergers and Acquisitions

* Correspondence to Bruce Carlin, Bhagwan Chowdhry, or Mark Garmaise, UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA 90095-1481, or e-mail: bruce.carlin@anderson.ucla.edu or bhagwan@anderson.ucla.edu or mark.garmaise@anderson.ucla.edu. We thank Tony Bernardo, Matthias Kahl, Sheridan Titman, the Finance seminar participants at Southern Methodist University, Stanford, UCLA, Indian School of Business and the participants at the ASSA 2004 meetings in San Diego for many useful comments on earlier drafts.
Investment in Organization Capital

Abstract

We study a firm’s investment in organization capital by analyzing a dynamic model of language development and intrafirm communication. We show that firms with richer internal languages (i.e., more organization capital) have lower employee turnover, higher diversity in skill, and greater wage dispersion. The model predicts that senior managers will more frequently be promoted from within in firms with a rich language. Our results also suggest that firms with lower asset betas and higher geographic concentration will invest more in organization capital by retaining their employees more often. Our model has implications for the management of human capital, executive compensation and mergers.
Introduction

Organization capital was first defined by Prescott and Visscher (1980) to be the accumulation and use of private information to enhance production efficiency within a firm. This capital can be a significant source of firm value. For example, Atkeson and Kehoe (2005) estimate that the payments that arise from organization capital are more than one-third the size of those generated by physical assets, and represent more than 40% of the cash flows generated by all intangible assets in the U.S. National Income and Product Accounts (NIPA). Despite its importance, however, studying the dynamics of investment in organization capital has not received much attention in the academic literature.

In this paper, we seek to fill this void by analyzing a theoretical model of organization capital and deriving both cross-sectional and time-series empirical implications that are of interest to corporate finance. We specifically address the following questions: How does investment in organization capital affect investment in alternative sources of value creation? How does such investment affect the dispersion of executive compensation within a firm? How do firm characteristics such as geographic location and beta risk (i.e., systemic risk) affect such investments? What is the relationship between a firm’s level of organization capital and its propensity to promote from within in its senior management ranks?

To address these questions, we develop a model of organization capital, viewing it as a form of intrafirm language. This captures the idea that the value of organization capital depends on its being shared across managers and that it must be transmitted to the next generation of employees to be preserved. A firm’s language summarizes informal work routines, convenient technical jargons and a vocabulary of patterns remembered from past experiences. It creates complementarities among managers because it facilitates communication and enhanced production (Crémer, 1993). Indeed, as Arrow (1974) points out, one of the advantages of an organization is its ability to economize in communication through a common code. The richness of a firm’s language, then, measures the breadth of the set of tasks covered by its communications channels, and is an important input to productivity within the firm.

We begin by analyzing a static model in which a firm exists for two periods and is then liquidated. The firm is endowed with a language that covers some of the types of
business opportunities that it may face. The firm has both junior and senior managers, and the key strategic decision it faces is how many incumbent managers to promote to senior management versus hiring from the outside. Ex ante, internal and external managers have the same expected productive quality, but the key difference between them is that incumbent managers may produce more efficiently by employing the firm’s language. In choosing whether to retain an incumbent or to hire an external manager the firm must trade off the incumbent’s valuable access to the firm’s language against the potentially higher personal productivity of an outside replacement.

In equilibrium, we show that the firm’s retention decision can be expressed as a threshold policy in which a quality is set above which incumbent employees are retained. Further, we show that greater richness of the firm’s language leads to a lower minimum quality requirement for retention. As a result, in a firm with a richer language, there is less turnover, more diversity in quality, and a greater difference between the highest and lowest paid incumbent senior managers. Empirically, then, the model predicts that firms with a richer language are more likely to exhibit decreased employee turnover, greater diversity in skill, higher incumbent wage dispersion, and more frequent promotion of senior managers from within the organization. These findings arise from the fact that the firm increases the support from which managers are drawn (by lowering the required lower bound). We also show that managerial compensation rises more quickly in firms with more organization capital, as managers learn to exploit the internal language.

These empirical predictions require a proxy for the richness of a firm’s language. The most straightforward candidate is the density of social networks that exists within an organization. Another potential proxy is the quality of relationships within those networks. Measuring these proxies has become increasingly feasible, for instance by exploiting new technologies for the content analysis of intrafirm e-mail communications. Our model’s predictions can tested by empirically studying the relationship, for example, between the density of a firm’s social networks and the variability of its managerial compensation.

Some firms may operate in industries with stable external environment whereas others may operate in industries with relatively volatile external environment. We interpret stability of the external environment in the model as the likelihood that a new business opportunity faced by the firm is covered by its language. Our model predicts that firms in stable industries
will retain more of their incumbents.

We then extend our analysis to consider a dynamic study of how language evolves endogenously through time. We develop an overlapping generations model with an infinitely-lived firm and managers who remain with the firm for a maximum of two periods (if they are promoted to senior management from within the organization). We consider that language is transmitted from senior management to juniors with some probability, and that this probability is increasing in the number of incumbent managers that are retained. This provides an additional incentive to promote an incumbent employee to senior management because he will help transmit the firm language to the new generation of junior managers and thereby assist the firm in preserving its organization capital.

As with any asset that affects production, accumulating organization capital requires investment and the allocation of resources. Optimal investment may require substituting away from alternative forms of productivity. A richer language induces a firm to retain incumbents of lower quality. In our model firms invest in their organization capital by retaining incumbent managers with relatively low personal capabilities. These managers generate relatively low cash flows for the firm in the current period, but help to maintain the firm’s organization capital and therefore to create greater productivity in the future.

The dynamic model, in addition to preserving the insights of the static model, shows that investment in language is higher in firms with a lower discount rate. Since firms that undertake lower systematic risk projects apply a lower discount rate to future cash flows, this has several clear empirical ramifications. Specifically, our analysis predicts that firms with lower asset betas are more likely to invest in language. This then implies that low asset beta firms will have higher wage dispersion, more skill diversity, and lower employee turnover. To our knowledge, these implications have not been tested before, but are the subject of future research.

We further show that investment in the dynamic model is greater when the firm has a more responsive transmission function (i.e., when incumbent retention is more effective in generating successful language transmission). There might be cross-sectional differences across firms in their ability to transmit their languages and preserve their organization capital from one generation of managers to the next. For example, firms that are geographically more concentrated may be able to transmit their languages to future generation of managers
with greater probability. Our model therefore predicts that such firms will have greater employee retention and more within-firm variability in executive compensation.

Finally, we discuss the implications of our model for mergers. An implication of our theory is that the most valuable mergers are between firms with very different levels of organization capital. This is because the constituent-firm language that is not adopted by the merged firm is lost. Mergers that minimize the destruction of organization capital create more value.

Our model contributes to the literature on intrafirm communication, and is distinctive in several respects. Many studies take the information in the firm as given, and analyze the optimal way to employ that capital. For example, Bolton and Dewatripoint (1994) analyze the costs and benefits of centralizing information networks within firms, given an exogenous flow of new information. Harris and Raviv (2002) analyze the formation of optimal organization design, given that information as a scarce is exogenously given. Crémér, García, and Pratt (2007) study the development of optimal codes within organizations to employ the capital from information in the most efficient way. In contrast, in our paper, we focus on the nature of investment in language, and study the effect that this has on the firm’s labor management policy. Intrafirm communication is explicitly promoted by the firm’s decisions to retain incumbent managers and these decisions also influence employee compensation. Our model therefore describes endogenous differences in quality diversity and wage dispersion in the firm.

The rest of the paper is organized as follows: Section I poses and solves the two-stage version of our model. There we characterize the firm’s retention policies, and their ramifications for diversity and wage dispersion. In Section II, we analyze the overlapping generations model and characterize the firm’s investment in language. Section III considers the issue of merger integration. Section IV concludes. The Appendix contains all of the proofs.

I. Production and Organization Capital

We begin by considering a firm that employs managers for two periods and is then liquidated. We characterize the firm’s hiring and firing decisions, employment compensation, investment in organization capital through its decision to retain its incumbent managers, and diversity
Figure 1: At the beginning of $t = 1$, $N$ junior managers are hired. At the beginning of $t = 2$, each incumbent manager’s quality $y_i$ is realized. Based on the firm’s scope $K$ and the firm’s organization capital (language) $L$, the firm chooses which incumbent managers to promote and which ones to replace with outsiders. After this, the project $k$ is realized and the senior management produces $Y$ in aggregate based on whether $k \in L$ or not. Finally, senior managers are compensated and the firm is liquidated.

of skill levels in the organization. In Section II, we will embed this two-stage interaction into an overlapping generations model in which an infinitely-lived firm employs managers who work for two periods and then retire.

A. Language and Production

The timing of the game is outlined in Figure 1. In the first period ($t = 1$), a group of new managers are hired and are considered to be “junior” to an existing set of “senior” managers. The firm consists of $2N$ workers, $N$ seniors and $N$ juniors. During the first period, junior managers assist the seniors in producing the output of the firm, but do not create value independently from senior management. Their type (quality) $\tilde{y}$ is unknown to both the firm and the manager, and does not affect value creation for the firm when they are junior. We assume that for each manager, $\tilde{y}$ is distributed according to a twice continuously differentiable, strictly increasing, log-concave distribution function $F$ over the support $[0, Y]$, where $E[\tilde{y}] = \bar{y}$. As such, our scope of analysis here is fairly wide, including common distributions such as the uniform, truncated normal and exponential distributions.

At the beginning of the second period ($t = 2$), the types of all incumbent junior managers are revealed to both managers and firms. Existing senior managers retire and some junior incumbents are promoted to form the new senior management, depending on the firm’s retention policy and the willingness of management to remain with the firm (to be specified shortly). Define $n$ as the number of incumbent junior managers who are promoted to senior
management, and \( N - n \) as the number of senior managers who are recruited from outside the firm. We assume in what follows that the firm’s physical capital (real assets) is fixed, in order to analyze the firm’s investment in organization capital through its retention decision.

In any given period, the firm is faced with carrying out a project \( k \in \mathcal{K} = \{1, 2, \ldots, K\} \) for some \( K > 1 \) that arises randomly according to a uniform random variable that is i.i.d. across time periods. These projects may be thought of as representing different strategic initiatives or market opportunities.

Each senior manager \( i \) produces an individual output \( y_i \) that directly depends on his own quality. Given the number of incumbent managers, all managers may increase their production through language. This type of production enhancing communication is only possible, though, when the language includes the particular project at hand. Let \( \mathcal{L} \subseteq \mathcal{K} \) denote the set of all projects that are part of the firm’s language and let \( L \) denote the number of such projects. As such, \( L \) provides a measure of the level of the firm’s organization capital, and by construction, the probability that any task \( k \in \mathcal{L} \) is \( \frac{L}{K} \). (We describe the evolution of \( \mathcal{L} \) in the dynamic game in the following section.)

If the given task is part of the language, then the incumbent managers will foster more communication among all senior managers. In total, language increases the total productivity of all managers and increases production by \( G(n) \), where \( G(\cdot) \) is an increasing function.\(^1\)

Production within the firm may then be calculated as

\[
Y = \sum_{i=1}^{N} y_i + \Psi_{k \in \mathcal{L}} G(n)
\]

where \( \Psi \) is an indicator function that is equal to one if \( k \in \mathcal{L} \) and zero otherwise. The expected productivity (from real assets) of incumbent managers will depend in equilibrium on \( L, K, \) and \( n \). We denote this dependence as \( E[\tilde{y}_L | L, K, n] \) and will calculate this quantity in the next section. Assuming that all outside managers are randomly chosen from the same distribution \( F \), we can compute the firm’s expected production as

\[
E[Y | L, K, n] = n E[\tilde{y}_L | L, K, n] + \frac{L}{K} G(n) + (N - n)\bar{y}.
\]

\(^1\)DeMarzo, Vayanos and Zwiebel (2001) and Garicano (2000) provide other models of communication in organizations. Crémér, Garicano and Pratt (2007) analyze the optimal design of a code within an organization and consider implications for integration across different groups of agents. A specific example of such a language based on internal jargon, shared values and common experiences is found in the workings of the consulting firm McKinsey described in The McKinsey Mind by Raisel and Friga (2002).
The sum of the first and second terms is the expected productivity that arises from incum-
bents seniors, whereas the third term reflects the expected contribution from new seniors
from the outside.

Compensation within the firm proceeds as follows. Every senior manager has a reservation
utility of $\bar{u}$, which does not depend on his particular realization of $\hat{y}$.\(^2\) While we do not model
the determinants of $\bar{u}$ explicitly, its value is common to all managers and reflects conditions
in the labor market such as competition, market power, and differentiation in skill. This
implies that $\bar{u}$ is expected to be higher when the skills that particular employees provide are
difficult to replace. As such, participation by any particular manager will occur if and only
if the firm meets their participation constraint.

Compensation for each manager is determined according to a Nash bargaining game in
which the firm pays the manager a fraction $\theta \in (0, 1)$ of the value of the production in
exchange for the remainder. Here, we follow Radner and Van Zandt (1992) and Garciano
(2000) in that we set incentives within the firm aside and focus on the investment in language
by the firm.

The payoff to a senior manager is computed as

$$\pi_i = \max \left\{ \bar{u}, \theta \left[ y_i + \Psi_{k \in \mathcal{L}} \frac{G(n)}{N} \right] \right\}. \quad (2)$$

As such, each manager gets a fraction of their own productivity and an equal share of the
value that is created by language in the organization.\(^3\)

When $\theta [y_i + \Psi_{k \in \mathcal{L}} \frac{G(n)}{N}] < \bar{u}$, the firm promises to supplement the compensation with cash
to meet the participation constraint. By inspection, the higher the quality of a manager and
the higher the potential for complementarities, the lower the the need for cash to supplement
an employee’s pay. For simplicity, we assume that $\bar{u} = 0$ so that cash is never required as a
supplement and no manager would voluntarily quit if the firm promoted him. We do this,
though, keeping in mind that if $\bar{u} > 0$ that an employee would quit if they were not offered

\(^2\)We make this assumption for analytical simplicity. Relaxing this assumption will make the firm’s reten-
tion decision more realistic, but will also make the analysis more complicated, which will not add much to
studying their investment in organization capital.

\(^3\)Our production sharing scheme describes an arrangement in which managers receive some compensation
based on their own revealed quality and a fraction of firm level bonus that is divided equally among all
$N$ employees. Following Stole and Zwiebel (1996), one might argue that a manager’s compensation must
depend on his or her *marginal* contribution to the firm’s total output. As long as $G$ is concave, such a
compensation scheme would be feasible in our model, and our central results would be preserved.
enough compensation to remain at the firm. Since the employment offer, in our model, is at the firm’s discretion, it is without loss of generality to consider that \( \bar{u} = 0 \) and that the firm chooses whether to retain or fire certain managers.

With this in mind, we denote the decision to retain incumbent managers by \( d_R \in \{0, 1\}^N \), where \( d_R(i) = 0 \) means that the \( i \)th incumbent is fired and \( d_R(i) = 1 \) means that he is retained. Thus, \( n = \sum_1^N d_R(i) \) and the expected profit to the firm is computed as

\[
\Pi(L, d_R) = (1 - \theta) \left[ nE[\bar{y}_I|L, K, n] + \frac{L}{K}G(n) + (N - n)\bar{y} \right].
\] (3)

We are now ready to solve and characterize the two-stage game.

B. Equilibrium Characterization

The object of interest that is determined in equilibrium is the number of incumbent managers \( n \) the firm wishes to promote, which will then determine the number of managers to hire from outside of the firm. At the time that this decision is made, the particular task \( k \) has not been observed, and therefore the firm does not know whether the firm’s organization capital will be put to good use in enhancing production. The firm does, however, observe the quality levels of its incumbent managers and uses this to make a decision regarding retention. Not surprisingly, this leads to an optimal threshold policy in equilibrium, which we characterize in the following proposition.

**Proposition 1.** There exists an optimal threshold productivity level \( y^* \) such that the firm retains all incumbent managers with \( y_i \geq y^* \) and replaces the rest with outsiders. For all \( L > 0 \), \( y^* < \bar{y} \). The threshold \( y^* \) is strictly decreasing in \( \frac{L}{K} \).

The expected quality of senior managers is strictly decreasing in \( \frac{L}{K} \).

To understand the intuition of Proposition 1, consider first that \( L = 0 \), or that there is no potential for incumbents to have an advantage over outsiders who join the firm. In this case, since the firm can gain \( \bar{y} \) in expectation from hiring outsiders, it will only retain incumbent managers with higher quality, that is, \( y_i \geq \bar{y} \). When \( L > 0 \), there is more to gain from keeping incumbents since there is a positive probability \( \frac{L}{K} \) that organization capital
can be put to good use. When the firm considers whether to retain one more incumbent, they compare the expected productivity from the incumbent with the expected productivity from an outsider. Specifically, they retain the incumbent if and only if

\[ y_i \geq \bar{y} - \frac{L}{K} \Delta G(n), \]

where \( \Delta G(n) = G(n) - G(n - 1) \). By inspection, the higher the organization capital, \( L \), relative to the span of opportunities that the firm may be confronted with, \( K \), the smaller is the employee turnover.

Proposition 1 has several empirical implications. First, firms with larger organization capital will experience lower turnover. Indeed, with higher \( L \), firms should be more averse to replacing managers with outsiders because employees who know the firm's language produce effectively within the firm and provide advice for other employees. Second, firms with higher organization capital should have more frequent insider CEO succession. Denis and Denis (1995) find that only 15 percent of firm top executive appointments are made to external candidates, which underscores the probable importance of organization capital in most firms. Consistent with this is the observation by Parrino (1997) that outside succession occurs most frequently in commodity industries in which organization capital is likely to be less important.

Proposition 1 also implies that the average quality of managers should decrease as organization capital \( L \) increases. As \( L \) rises, the bar that must be met to be promoted decreases. We can calculate the average quality of incumbent managers as

\[ E[\bar{y}_i | L, K, n] = \int_{y^*}^{y} ydF(y) \frac{1}{1 - F(y^*)}, \]

which implies that the average quality of all managers in the firm is

\[ E[y_i] = \frac{1}{N} \left[ n \int_{y^*}^{y} ydF(y) \frac{1}{1 - F(y^*)} + (N - n)\bar{y} \right]. \]

By inspection, it is clear that \( E[y_i] \) is decreasing in \( L \) and increasing in \( K \).

This has two important implications. The first is that organization capital improves productivity but is also associated with lower intrinsic manager qualities. For a set of managers with given qualities, higher communication within the firm due to the presence of incumbents leads to more sharing of ideas and greater productivity. At the same time,
as language becomes more important within the firm, the average quality of employees decreases because the bar that is required for promotion is lower. Therefore, when the firm operates, it must take into account both forces and weigh the tradeoffs that organization capital introduces.

The second implication is that organization capital affects the diversity of managers within the firm. As \( L \) increases, the support from which incumbents are drawn increases, which affects the difference in quality between employees. This, in turn, affects the expected amount of wage dispersion that exists in the organization. The following proposition characterizes the relationship between language and diversity and wage dispersion.

**Proposition 2.** The expected diversity and wage dispersion among incumbent senior managers is strictly increasing in \( \frac{L}{K} \).

According to Proposition 2, the variance of quality levels and the variance of expected wages increase as language plays a larger part within the firm. To gain intuition for this result, consider two levels of \( \frac{L}{K} \), namely \( \frac{L_1}{K_1} \) and \( \frac{L_2}{K_2} \), such that \( \frac{L_1}{K_1} > \frac{L_2}{K_2} \). By (4), it is clear that \( y^*(\frac{L_1}{K_1}) < y^*(\frac{L_2}{K_2}) \). The firm chooses incumbent managers from two distributions, which we can call \( H_1(\tilde{y}_1) \) and \( H_2(\tilde{y}_2) \), where \( \tilde{y}_1 \) and \( \tilde{y}_2 \) are random variables as defined in the text. The key observation to be made is that the distribution \( H_2(y) \) is a truncation of \( H_1(y) \). Therefore, it follows that \( Var(\tilde{y}_2) \leq Var(\tilde{y}_1) \), or that the variance of quality among incumbents is higher when language is more important to the firm. This leads to more diversity in the organization. Finally, since wages are linked to performance (through the fraction \( \theta \)), as language becomes more important in the organization, this leads to a higher variance of wages among incumbent managers. Empirically, then, Proposition 2 implies that the difference between the top incumbent wage earner and the average incumbent in a firm should be higher when organization capital is higher.

It is important to point out that junior managers would prefer to work at firms with higher language, holding all else equal. The following proposition formalizes this result.

**Proposition 3.** The expected payoff is higher for incumbents who begin the second period in firms with greater organization capital.
Proposition 3 may be appreciated as follows. Before a junior manager becomes informed about his type, he may compute the expected wage that he will receive at the firm in the second period. If his quality turns out to be $\tilde{y} < y^*$, then he will not be retained and will earn zero. If $\tilde{y} \geq y^*$, then he will expected to earn $E[\pi|\tilde{y} \geq y^*]$. Therefore, his expected wage is

$$E[\pi] = Pr(\tilde{y} \geq y^*)E[\pi|\tilde{y} \geq y^*] = \frac{[1 - F(y^*)]}{1 - F(y^*)} \theta \int_{y^*(L)}^{y^*} (y + \frac{L}{K}G(n(L))) \, dF(y)$$

or

$$E[\pi] = \theta \int_{y^*(L)}^{y^*} \left( y + \frac{L}{K}G(n(L)) \right) \, dF(y),$$

which is increasing in $L$. Therefore, as the language increases within a firm, junior managers have a higher expected wage in the future.

A simple extension of our model might set wages for juniors such that the total expected two-period compensation is equal to some reservation value. In such a model, salaries for juniors would be lower in firms with large organization capital, while the seniors in these firms would be well paid. In such a model, our theory would predict that the gap in compensation between juniors and seniors would be greater in firms with large organization capital. In other words, firms with a strong organization capital would exhibit greater steepness in their managerial wage profiles.

C. Empirical Implications

Many of the empirical predictions that follow from Propositions 1-3 are novel and have yet to be tested. It is worth discussing, however, how such implications might be analyzed. Testing our model would require a good proxy for language. One candidate is the density of social networks and the quality of relationships within those networks (e.g. Burt and Schott 1985; Raider and Krackhardt 2001). Indeed, the more intertwined managers are within an organization, both at work and outside of the firm, the more readily do they engender language and observe informal work routines. Along the same lines, another direct measure of a firm’s language is the importance and frequency of employee interactions as reported by the employees themselves. While collecting this data may be cumbersome, there is an increasing number of intrafirm studies on the value of communication (e.g. Ichniowski and
Shaw, 2003). In fact, given the increased reliance of firms on written emails, this difficulty of collecting relevant data has decreased since written communication may be analyzed by content analysis (e.g. Holsti 1969; Tetlock 2007).

An empirical proxy for the span of opportunities that the firm may be confronted with, \( K \), might be the frequency with which firms in a given line of business or industry change over time. Industries in volatile environments (larger \( K \)) are likely to see more firm exits and entry whereas industries in stable environments (smaller \( K \)) are likely to see the same firms operating in the industry or line of business over the years. Empirically, one could compute this volatility by examining the change in the composition of firms that are ranked in the top five for sales in a particular SIC code over the years.

Using these proxies, then, our model predicts that the density of intrafirm social networks should be positively correlated with wage steepness, compensation dispersion, and internal CEO promotion, and should be negatively correlated with employee turnover. Along the same lines, firms in less volatile industries should also have more wage steepness, more dispersed wages, and less employee turnover.

II. Investment in Organization Capital

In practice, firms are not simply endowed with organization capital as assumed in Section I. Rather, they cultivate it over time, which requires investment and the allocation of resources. In this section, we consider a firm’s optimal investment in organization capital through its employee retention policy, and characterize its evolution over time. We model the firm as an infinitely-lived entity and embed the static model of Section I into a dynamic setting. The key driver in this analysis is that the firm’s retention and hiring policies not only determine the current productivity of the firm, but also its ongoing stock of organization capital that it may draw upon in the future.

In each period \( t = 1, 2, \ldots \), the firm makes a retention decision \( d_R \in \{0, 1\}^N \), which is similar to that defined in Section I. The language of the firm, however, evolves depending on whether the current task at hand is included in the firm’s current language, and whether the firm’s language is transmitted from seniors to juniors.

Formally, for any task \( k \) that is performed, it becomes part of the firm’s language because
junior managers experience this task first hand. The probability that the rest of the current language is transmitted is given by the function \( p(n) \) such that \( p \) takes values within \([0, 1]\) and is strictly increasing in the number of senior incumbents \( n \) who know the firm’s language. This assumption captures the idea that language transmission is more likely if the firm invests in retaining incumbent managers. Also, if \( p_2(n) \geq p_1(n) \) and

\[
p_2(n + 1) - p_2(n) \geq p_1(n + 1) - p_1(n)
\]

for all \( n \leq N \), we call the language transmission process \( p_2 \) more responsive than \( p_1 \).

Therefore, three states of the world may be realized at the end of each period \( t \). If \( k \in L \), and the current language is transmitted, then the language going forward will be the same set \( L \) with count \( L \). If \( k \notin L \), and the rest of the current language is transmitted, then the language going forward will be a larger set of count \( L + 1 \). If the rest of the current language is not transmitted, then the language moving forward only includes a singleton, that is, the specific task just experienced by the junior managers. Thus, by construction, our model incorporates both organizational learning and organizational forgetting (e.g. Benkard 1999).

As a result, while the firm’s juniors do not produce themselves, they do observe the functioning of the seniors. It is indeed often difficult to imbibe a firm’s language, but once it is possessed it is very easy to assimilate the application of this language to the various tasks covered by the language. As a result, a manager learns either the entire language with all its applications or nothing of the language at all. We assume that juniors have identical language-learning skills and that they communicate amongst themselves, so that either all the juniors learn the firm’s language or none at all. It is intuitive that the probability that the language is transmitted depends on the number of incumbent seniors at the firm who speak the firm’s language.

The firm solves the following problem

\[
V(L, \{y_i\}_{i=1}^N) = \max_{d_R \in \{0, 1\}^N} \left[ \Pi(L, d_R, \{y_i\}_{i=1}^N) + \delta H(V, L, d_R) \right]
\]

As a result, while the firm’s juniors do not produce themselves, they do observe the functioning of the seniors. It is indeed often difficult to imbibe a firm’s language, but once it is possessed it is very easy to assimilate the application of this language to the various tasks covered by the language. As a result, a manager learns either the entire language with all its applications or nothing of the language at all. We assume that juniors have identical language-learning skills and that they communicate amongst themselves, so that either all the juniors learn the firm’s language or none at all. It is intuitive that the probability that the language is transmitted depends on the number of incumbent seniors at the firm who speak the firm’s language.

The firm solves the following problem

\[
V(L, \{y_i\}_{i=1}^N) = \max_{d_R \in \{0, 1\}^N} \left[ \Pi(L, d_R, \{y_i\}_{i=1}^N) + \delta H(V, L, d_R) \right]
\]

\[4\]For example, if the probability of transmission were parameterized as \( \phi p(n) \) with \( \phi > 0 \) then a larger value of the parameter \( \phi \) would imply both a higher probability of transmission as well as a more responsive transmission process.

\[5\]This is analogous to the high cost of learning a language and the low marginal cost of reading a book or conducting a conversation in a language already understood.
where $V$ is the value function, the current period expected profits $\Pi$ are

$$
\Pi(L, d_R, \{y_i\}_{i=1}^N) = (1 - \theta) \left[ \sum_{i:dR(i)=1} y_i + \frac{L}{K} G(n) + (N - n)\bar{y} \right],
$$

and the value of future profits is

$$
H(V, L, d_R) = \int \left[ p(n) \left\{ \frac{L}{K} V(L, \{\tilde{y}_i\}_{i=1}^N) + \left( 1 - \frac{L}{K} \right) V(L + 1, \{\tilde{y}_i\}_{i=1}^N) \right\} \
+ \{1 - p(n)\} V(1, \{\tilde{y}_i\}_{i=1}^N) \right] dF^N(\{\tilde{y}_i\}).
$$

The factor $\delta \in (0, 1)$, which discounts profits from future periods, is tied to the firm’s cost of capital. As such, $\delta$ is higher for firms with lower asset betas.

The following proposition characterizes the solution to the firm’s problem.

**Proposition 4.** *(Investment in Organization Capital)* A firm with a higher discount factor $\delta$ and/or a more responsive transmission function $p(\cdot)$ will retain more incumbent managers.

Proposition 4 adds to our characterization of investment in organization capital in two ways. First, as the weighted average cost of capital falls for a firm, we expect that the retention of incumbents should increase. Second, as the effectiveness of language transmission rises in an organization, investment in language through retaining incumbent managers also increases.

This gives rise to several novel cross-sectional empirical implications. First, since the discount rate is higher for firms in high-risk industries (i.e. high asset beta), Proposition 4 predicts that firms with higher systematic risk should have more employee turnover and less wage dispersion among incumbents. In general this shows that investment in language through employee retention is higher in firms that value the future more. Thus the model predicts that firms that discount the future cash flows less because they have low systematic risk (low asset betas) should be more likely to promote senior managers from within.

Second, the character of the language transmission process may be affected by geographic dispersion of employees. Our notion of firm language tries to capture the idea in Prescott and Visscher (1980) that casual conversations transmit valuable information at a very low cost.
to productivity. With increasing use of technology, even though communication has become cheaper and easier, the use of casual conversations and social interactions is still likely to be quite limited when employees are geographically apart. In firms that are geographically dispersed, junior managers are less likely to have opportunities to interact with their senior managers informally thus reducing the probability of firm language being transmitted from one period to the next. Therefore, if transmission risk is correlated with geographic dispersion of personnel then we would expect firms in which its employees are geographically dispersed to have, ceteris paribus, lower retention of incumbents than firms concentrated in a single locale. Given our discussion in Section I, then, we would predict there should be less diversity in skill and less wage dispersion among incumbents in firms that are geographically dispersed.

III. Mergers

Our model provides a rationale for value-creating mergers. If one firm has developed a very rich language, this language may usefully be adopted by other firms performing similar tasks. Consider a merger between two firms of roughly the same size, one with significant organization capital and a second with very little organization capital. Assuming there is some overlap between the tasks of the two firms, the juniors at the newly merged firm will likely learn the rich language of the high organization capital firm. The value created by a merger is equal to the value of the merged firm minus the values of the two constituent firms. Since the organization capital of the firm whose language is not adopted is simply lost, the value created by the merger is greatest when one of the constituent firms has a lot of organization capital and the second has very little.

The transmission of organization capital in our model depends on the number of incumbents who speak the language. If a small firm with a high quality language merges with a large firm with a relatively low quality language, the language that gets transmitted to junior managers in the merged firm may be the language of the larger constituent. That suggests that significant value may be destroyed when a large firm purchases a much smaller company with significant organization capital, since this organization capital will likely be dissipated in the merger.

\[6\] Crémer, Garicano and Pratt (2007) analyze a model in which two firms may choose to adopt a common code at some cost.
These arguments indicate that the most efficient mergers are between large firms with substantial organization capital and smaller firms with little organization capital. If we were to consider that the market-to-book ratio proxies for quality of organization capital, our theory predicts that the value created by the merger is highest between firms with very different market-to-book ratios. Indeed, Lang, Stulz and Walkling (1989) and Servaes (1991) show that total returns on merger announcements are larger when target firms have low market-to-book ratios and bidders have high market-to-book ratios. Admittedly, though, many firm characteristics drive a firm’s market-to-book ratio and may be responsible for previous findings. To directly test our predictions regarding organization capital and value creation in mergers, it would be important to correlate the proxies for intrafirm language described in Section I.C with the total returns following merger announcements, holding other firm characteristics constant.

In the same way, our analysis also implies that the market-to-book ratio and retention rates of the merged firm should closely resemble those of one of the constituent firms, rather than reflecting an average over both constituent firms, since we have presumed that only one language will survive in the merged firm. Again, testing our theory directly would involve using the proxies for intrafirm language discussed in Section I.C.

In general, though, mergers will indeed reduce the probability of organization capital transmission. Exporting a rich language via a merger can be beneficial, but also presents the risk of loss. It is not the case that firms with large organization capital should engage in unbridled expansion.

**IV. Conclusion**

We present a model describing a firm’s language as its organization capital. We show that firms with richer languages retain more employees and are therefore more likely to promote senior managers from within. We demonstrate that firms with more organization capital will exhibit greater variability in the compensation levels of their managers. We also prove that compensation rises more quickly over time in firms with richer languages.

Our analysis of the transmission of organization capital and its dynamic evolution generates predictions that do not naturally arise in models of static firm-specific human capital
with complementarities among managers. In particular, our results that firms in industries with low asset betas and firms with higher geographic concentration will have higher wage dispersion, more skill diversity, lower employee turnover and are more likely to promote insiders, are driven by the dynamic effects we explore.

Our description of organization capital as an internal language of the firm meets two important criteria. First, the firm’s language cannot be carried from the firm by departing employees. Second, the firm’s language is difficult to imitate.

It is important that organization capital be tied to the firm, for otherwise it is difficult to explain why employees and assets must stay together. A coordinated *en masse* defection by all employees can typically be ruled out because of the coordination difficulty discussed in Klein (1988). Hart (1989) argues that a threat of simultaneous defection by all employees can be still be credible unless some physical assets are involved. In our model, the language of the firm is used to describe the firm’s particular tasks and is therefore linked to the precise equipment and production arrangement used by the firm.

For organization capital to have value, it must also be costly for competitors to replicate (Rumelt, 1987). Inimitability, in our model, arises because the knowledge of a firm’s language is possessed by the firm’s managers and is not accessible to rivals. Moreover, the language is related to the particular way the firm is structured. In our model, learning and experience are necessary for the development of each firm’s language. These features combine to make the acquisition of language within the firm time-consuming and difficult.

Our model of organization capital provides novel testable implications linking the density of firms’ social networks to central issues in corporate finance including firms’ market values, compensation practices and merger strategies. Recent empirical evidence has bolstered the view that organization capital plays a significant role in production (Atekson and Kehoe, 2005). It is therefore important to broaden our understanding of how it creates value within the firm.

---

7 Bahk and Gort (1993) empirically document, using individual plant data for one sample of 15 industries and another sample of 41 industries, that “organization learning appears to continue over a period of at least 10 years following the birth of a plant.”
Appendix

Proof of Proposition 1

The firm will retain an incumbent manager if the value they create is expected to be higher than the quality of an outside manager. The firm will choose to retain one additional incumbent if their quality satisfies

\[ y_i + \frac{L}{K} \Delta G(n) \geq \bar{y}, \]

or if

\[ y_i \geq y^*, \]

where \( y^* \equiv \bar{y} - \frac{L}{K} \Delta G(n) \). By simple differentiation, \( \frac{\partial y^*}{\partial L} < 0 \) and \( \frac{\partial y^*}{\partial K} > 0 \).

The expected quality of a senior manager is computed as

\[ E[y_i] = \frac{1}{N} \left[ n \int_{y^*}^{y} y dF(y) + (N - n) \bar{y} \right]. \]

Differentiation using Leibnitz’ Rule yields \( \frac{\partial E[y_i]}{\partial L} < 0 \) and \( \frac{\partial E[y_i]}{\partial K} > 0 \). ■

Proof of Proposition 2

Considering two values \( z_1 = \frac{L_1}{K_1} \) and \( z_2 = \frac{L_2}{K_2} \), such that \( z_1 > z_2 \). By (4), it is clear that \( y^*(z_1) < y^*(z_2) \). Define the two distributions from which the firm chooses \( n \) incumbent seniors from as \( H_1(\tilde{y}_1) \) and \( H_2(\tilde{y}_2) \), where \( \tilde{y}_1 \) and \( \tilde{y}_2 \) are random variables as defined in the text.

By inspection, the distribution \( H_2(y) \) is a truncation of \( H_1(y) \). Since \( F(\cdot) \) is log-concave, and that \( H_1(y) \) and \( H_2(y) \) are both truncations of \( F(\cdot) \), it follows that \( Var(\tilde{y}_2) \leq Var(\tilde{y}_1) \).

(See Burdett, 1996 and An, 1998). Finally, since all incumbent managers receive \( \frac{\theta G(n)}{N} \), the difference in their wages depends on their dispersion in quality. This implies that the variance of wages increases as \( \frac{L}{K} \) rises. ■

Proof of Proposition 3

An incumbent who is not retained receives a payoff of \( \bar{u} = 0 \), so the expected payoff for an incumbent is given by

\[ \theta \int_{y^*(L)}^{y} \left( y + \frac{L}{K} G(n(L)) \right) dF(y). \]
Proposition 1 shows that $y^*(L)$ is decreasing in $L$ and $n(L)$ is increasing in $L$. It immediately follows that the incumbent payoff is increasing in $L$.

**Proof of Proposition 4**

As candidate value functions we consider continuous functions mapping from $\{0, .., K\} \times [0, \mathcal{Y}]^N$ to $\mathbb{R}^+$. The space $D = \{0, .., K\} \times [0, \mathcal{Y}]^N$ is a product of compact spaces, and is therefore compact (Munkres, p. 167). We define $(C(D))$ to be the set of continuous functions mapping from $D$ to $\mathbb{R}^+$, with the sup norm $\rho$. It follows from the fact that $D$ is compact that $(C(D))$ is complete under the sup norm. Thus $(C(D), \rho)$ is a complete metric space.

We define

$$TV = \max_{d_R \in \{0,1\}^N} [\Pi(L, d_R, \{y_i\}_{i=1}^N) + \delta H(L, V)].$$ (A.1)

It is clear from its definition that $TV : (C(D)) \rightarrow (C(D))$ meets Blackwell’s sufficient conditions for a contraction (Stokey and Lucas, p.54). The Contraction Mapping Theorem (Stokey and Lucas, p. 50) then shows that $T$ has a unique fixed point $V^*$. This unique fixed point $V^*$ is the value function for the firm’s dynamic optimization problem (Stokey and Lucas, p. 256-258).

For given transmission functions $p_2$ and $p_1$, we will describe $p_2$ as more responsive than $p_1$ if $p_2(0) \geq p_1(0)$ and $p_2(n+1) - p_2(n) \geq p_1(n+1) - p_1(n)$ for all $0 \leq n \leq N$.

We first note that $T$ maps non-decreasing non-negative functions into non-decreasing non-negative functions. Since the space of non-decreasing non-negative functions is closed, $V^*$ is non-decreasing and non-negative (Stokey and Lucas, p. 52).

Formally, Proposition 4A states that for $\delta_2 \geq \delta_1$ and a fixed set $\{y_i\}_{i=1}^N$, if retaining $n$ incumbents is optimal under discount factor $\delta_1$ then retaining $m \geq n$ incumbents is optimal under discount factor $\delta_2$. We denote the value function and contraction mapping associated with discount factor $\delta_i$ by $V_i$ and $T_i$, respectively, for $i \in \{1, 2\}$.

We first show that $V_2 \geq V_1$ in the sense that $V_2(L, \{y_i\}_{i=1}^N) \geq V_1(L, \{y_i\}_{i=1}^N)$ for all $0 \leq L \leq K$. Let nonnegative nondecreasing $x, y \in (C(D))$ be given. If $x \geq y$, it is clear from (A.1) that $T_2(x) \geq T_1(y)$. It thus follows that for $s \geq 1$, $T_2^s(x) \geq T_1^s(x)$. By the Contraction Mapping Theorem we have $T_i^s(x) \rightarrow V_i$ for $i \in \{0, 1\}$, and the result follows.

The proof now proceeds by induction. For the base step, we will show that for all $L \geq 2$
and $L \leq K$

$$(T_2(x))(L, \{y_i\}_{i=1}^N) - (T_2(x))(L-1, \{y_i\}_{i=1}^N) \geq (T_1(x))(L, \{y_i\}_{i=1}^N) - (T_1(x))(L-1, \{y_i\}_{i=1}^N).$$

(A.2)

For notational simplicity, we will not explicitly write out the arguments describing the dependence of $x$ or $V$ on $\{y_i\}_{i=1}^N$. We denote the maximizing argument on the right side of (A.1) for $T_j(x)(c)$ by $d_R(j,c)$ and we define $n(j,c)$ to be the sum of retained incumbents $(n(j,c)= \sum_i d_R(j,c))$. Let us also rewrite $H(V,L,d_R)$ as

$$H(V,L,d_R) = \int \left[ p(n) \left\{ \frac{L}{K}(V(L) - V(1)) + \left( 1 - \frac{L}{K} \right)(V(L+1) - V(1)) \right\} + V(1) \right] dF^N(\{\tilde{y}_i\})$$

where

$$h(V,L) = \int \left\{ \frac{L}{K}(V(L) - V(1)) + \left( 1 - \frac{L}{K} \right)(V(L+1) - V(1)) \right\} dF^N(\{\tilde{y}_i\})$$

We first assume that $n(1,L) \geq n(2,L-1)$. Since $d_R(1,L)$ and $d_R(2,L-1)$ are feasible choices for all $(j,c)$ (and hence for $(2,L)$ and $(1,L-1)$, respectively), we have

$$(T_2(x))(L) + (T_1(x))((L-1)) - (T_1(x))(L) - (T_2(x))((L-1)) \geq$$

$$(\delta_2 - \delta_1) \{ p(n(1,L))h(x,L) - p(n(2,L-1))h(x,L-1) \} \geq 0$$

where the final inequality follows from the fact that $x$ is nondecreasing.

Suppose instead that $n(1,L) < n(2,L-1)$. We now note that $d_R(1,L)$ is a feasible choice for $(1,L-1)$ and $d_R(2,L-1)$ is a feasible choice for $(2,L)$. Thus

$$(T_2(x))(L) + (T_1(x))((L-1)) - (T_1(x))(L) - (T_2(x))((L-1)) \geq$$

$$\Pi(L,d_R(2,L-1)) + \Pi(L-1,d_R(1,L)) - \Pi(L,d_R(1,L)) - \Pi(L-1,d_R(2,L-1))$$

$$+ \delta_2 p(n(2,L-1))h(x,L-1) - \delta_1 p(n(1,L))h(x,L) \geq$$

$$\geq \frac{1}{K}(1 - \theta)[G(n(2,L-1)) - G(n(1,L))] \geq 0.$$

This completes the proof of the base step. For the induction step, suppose the result has been shown for some $s$. We will show that for all $L \geq 2$ and $L \leq K$

$$(T_2^{s+1}(x))(L) - (T_2^{s+1}(x))(L-1) \geq (T_1^{s+1}(x))(L) - (T_1^{s+1}(x))(L-1).$$

(A.3)

20
As before, we denote the maximizing argument on the right side of (A.1) for \((T_2^{s+1}(x))(c)\) by \(d_R(j, c)\). We first assume that \(n(1, L) \geq n(2, L - 1)\). Since \(d_R(1, L)\) and \(d_R(2, L - 1)\) are feasible choices for all \((j, c)\) (and hence for \((2, L)\) and \((1, L - 1)\), respectively), we have

\[
(T_2^{s+1}(x))(L) + (T_1^{s+1}(x))((L - 1)) - (T_1^{s+1}(x))((L)) - (T_2^{s+1}(x))((L - 1)) \geq \\
\delta_2 p(n(1, L)) h(T_2^s(x), L) + \delta_1 p(n(2, L - 1)) h(T_1^s(x), L - 1) \\
- \delta_1 p(n(1, L)) h(T_1^s(x), L) - \delta_2 p(n(2, L - 1)) h(T_2^s(x), L - 1) \geq 0
\]

where the final inequality follows from the induction step and from the fact that if \(a \geq b \geq d \geq 0, a \geq c \geq d \geq 0, a - b \geq c - d, \lambda_1 \geq \lambda_2 \geq \lambda_4 \geq 0, \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0 \) and \(\lambda_1 - \lambda_2 \geq \lambda_3 - \lambda_4\) then \(\lambda_1 a - \lambda_2 b - \lambda_3 c + \lambda_4 d \geq 0\).

Suppose instead that \(n(1, L) < n(2, L - 1)\). We now note that \(d_R(1, L)\) is a feasible choice for \((1, L - 1)\) and \(d_R(2, L - 1)\) is a feasible choice for \((2, L)\). We have

\[
(T_2^{s+1}(x))(L) + (T_1^{s+1}(x))((L - 1)) - (T_1^{s+1}(x))((L)) - (T_2^{s+1}(x))((L - 1)) \geq \\
\Pi(L, d_R(2, L - 1)) + \Pi(L - 1, d_R(1, L)) - \Pi(L, d_R(1, L)) - \Pi(L - 1, d_R(2, L - 1)) \\
+ \delta_2 p(n(2, L - 1)) \{h(T_2^s(x), L) - h(T_2^s(x), L - 1)\} \\
- \delta_1 p(n(1, L)) \{h(T_1^s(x), L) - h(T_1^s(x), L - 1)\} \\
\geq \frac{1}{K}(1 - \theta)[G(n(2, L - 1)) - G(n(1, L))] \geq 0
\]

where the second inequality follows from the induction step. This completes the induction proof. The Contraction Mapping Theorem and (A.3) together show that for all \(L \geq 2\) and \(L \leq K\)

\[
V_2(L) - V_2(L - 1) \geq V_1(L) - V_1(L - 1).
\] (A.4)

We denote a solution to the right-hand side of (A.1) for \(\delta = \delta_i\) by \(d_R^i\). For \(L = 0\), the solution of the maximization problem is independent of \(\delta\) and \(v\), so the retention policy is the same for \(\delta = \delta_1\) and \(\delta = \delta_2\). We next assume that \(L \geq 1\). Suppose that \(n(d_R^1) > n(d_R^2)\). For convenience, we denote the objective function on the right-hand side of (A.1) by \(Q(L, d_R, V, \delta)\). We note that

\[
Q(L, d_R^1, V_2, \delta_2) - Q(L, d_R^2, V_2, \delta_2) - (Q(L, d_R^1, V_1, \delta_1) - Q(L, d_R^2, V_1, \delta_1)) = \\
\{\delta_2 p(n(d_R^1)) - \delta_2 p(n(d_R^2))\} h(V_2, L) - \{\delta_1 p(n(d_R^1)) - \delta_1 p(n(d_R^2))\} h(V_1, L) \geq 0
\]
where the inequality follows from (A.4). We conclude that $d_R^1$ is an optimal retention policy under $\delta_2$ as well, which completes the proof for Proposition 4A. The proof of Proposition 4B follows from identical arguments, replacing $\delta_i p(n)$ in the above proof with $p_i(n)$. ■
References


