



Organization Capital and Intrafirm Communication

Bhagwan Chowdhry, UCLA Anderson School and The Indian School of Business
Mark J. Garmaise, UCLA Anderson School

April, 2005

Research Paper No: 03-05

Research Paper Series
Centre for Analytical Finance
Indian School of Business

Organization Capital and Intrafirm Communication*

Bhagwan Chowdhry
UCLA Anderson School and The Indian School of Business

and

Mark J. Garmaise
UCLA Anderson School

Current Draft: April 2005

JEL Classification: D23, G34

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

* Correspondence to Bhagwan Chowdhry or Mark Garmaise, UCLA Anderson School, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095, or e-mail: bhagwan@anderson.ucla.edu or mark.garmaise@anderson.ucla.edu. We thank Tony Bernardo, Matthias Kahl, Sheridan Titman, the Finance seminar participants at Stanford and UCLA, the participants at the ASSA 2004 meetings in San Diego, and the participants at the December 2004 Winter Research Conference at the Indian School of Business (ISB) for many useful comments on earlier drafts.

Organization Capital and Intrafirm Communication

Abstract

We present a dynamic model of production in which a firm's output increases when its managers share ideas. Communication of ideas depends on the quality of the firm's internal language. We prove that firms with richer languages, *i.e.*, more organizational capital, will have higher market values. Organizational capital generates static complementarities among incumbents which implies that firms with larger organization capital will experience greater employee retention, insider managerial succession and higher wages. Dynamic complementarities between inter-temporal investments in organization capital generate long-run persistence in firm market-to-book and turnover ratios. We demonstrate that the optimal compensation of incumbents includes an earnings-insensitive component that is larger in firms with larger organization capital. In a simple model of mergers, we show that the most value-creating mergers are those between firms with very different levels of organization capital.

Introduction

Economists often think of a firm as a collection of assets. These assets are further classified as tangible and intangible assets. Examples of tangible assets are physical capital such as plant and machinery. Examples of intangible assets are patents, brand names, R&D expertise and the knowhow that resides within the organization with its employees (Lev, 2001). Another classification, that is often invoked by scholars of the theory of the firm, distinguishes human capital from non-human capital (Hart, 1989). We propose a taxonomy that divides firms' assets into three classes: physical capital, human capital and *organization capital*. We argue that this classification provides insights into the nature of the firm and suggests a consistent and coherent explanation of many empirical observations about firms. Prescott and Visschler (1980) argue that elucidating the role of organization capital is central to understanding the function of the firm. Our focus in this paper is to understand what constitutes organization capital and how it relates to firm value, labor practices, compensation and mergers.

We present a model of production in which a firm's output depends on its physical assets, on the qualities (human capital) of its managers, and on their ability to communicate effectively with each other. While performing their job functions, managers acquire tacit knowledge that can be useful to their peers but that may be difficult to convey. We posit that when a firm undertakes a new project or task, its managers develop *informal* communication channels for talking about that task and sharing tacit knowledge (Crémer, 1993). Informal work routines, convenient technical jargon and a vocabulary of patterns are developed in the course of carrying out the task.¹ The richness of a firm's language is a measure of the breadth of the set of tasks covered by its communications channels and is the essential component of its organizational capital.

We develop a dynamic model of a long-lived firm with managers who work for two periods. In their first period, junior managers acquire their firm's language with some probability. In their second period, senior managers may quit or be fired from the firm. The set of senior managers who remain with the firm then produces output that depends on their own qualities, the richness of the firm's language and the number of managers who share the

¹Simon (1979) refers to subconsciously remembered patterns or vocabulary of patterns in describing the richness of knowledge created by past experiences.

firm's language. Managers generate insights into the functions of other managers and can, if communication is possible, provide performance-improving advice to their peers at no cost. A richer language, *i.e.*, more organizational capital, facilitates communication between senior incumbents, since it provides critical terms about more tasks.² We also assume that the larger the number of retained senior incumbents, the greater the probability that the language will be successfully transmitted to junior managers and thus that the firm's organizational capital will be preserved.

The basic trade-off in the model is faced by a firm considering whether to replace a low quality incumbent. An outside replacement manager may have a higher personal productivity, but he will be unable to share ideas with other managers, and he will not facilitate the transmission of the firm's language and preservation of firm's organization capital. Our model differs from other studies of intrafirm communication (Bolton and Dewatripoint, 1994, Harris and Raviv, 2002, Dessein and Santos, 2003) in our focus on the effect of the firm's labor management policy on the dynamic evolution of organizational capital. Organizational capital builds up slowly as the firm undertakes new projects and gains valuable experience. If enough senior managers are retained by the firm then the organizational capital is likely transmitted to next period's managers. Organizational capital in our model thus accumulates in part via learning-by-doing but the main subject of our analysis is the role of managerial turnover and intrafirm communication in preserving and exploiting this capital.

We demonstrate that, given a fixed set of assets, firms with larger organization capital have greater market values because their managers are more productive. Firms with larger organization capital are firms in which information sharing and teamwork are important. We present evidence that firms with those attributes have higher market-to-book ratios and profitability.

We show a key result that if good ideas are relatively scarce, then the benefits to the firm from having more managers who know its language are increasing and convex. This generates *static complementarities* between incumbents. The value to the firm of retaining an additional incumbent increases in the number of incumbents retained. Individual incumbents

²Arrow (1974) argues that one of the advantages of the organization is its ability to economize in communication through a common code. 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).

benefit from the presence of other incumbents who can communicate with them and provide advice. As a consequence, we can show that firms with richer languages in which these complementarities are stronger tend to experience fewer quits and more retention. This suggests that in rich language firms managers will have longer within-firm tenures and insider CEO succession should be more frequent. Organizational capital is likely less important in commodities industries, so we predict that insider CEO succession will be less common in those industries relative to human-capital-intensive industries, and we cite empirical evidence in support of that prediction (Parrino, 1997).

We demonstrate that the presence of static complementarities also implies that firms will experience cascades in quitting. We show as well that wages are higher in firms with larger organization capital. Our model implies a novel prediction that suggests that wages of any given employee are also increasing in the seniority of *other* managers.

Our description of organization capital is related to the idea of firm-specific human capital and, indeed, several of the results detailed above would arise in many static models of firm-specific human capital with complementarities among managers. Our model differs in its focus on the intertemporal transmission of organization capital. The level of a firm's organization capital, in our setting, affects not only its current production but also its incentives to preserve its organizational capital. A firm that retained many incumbents in the previous period is likely to have a larger organization capital. This gives the firm an inducement to retain many incumbents this period in a bid to maintain its valuable organizational capital. Firms with smaller organization capital will have little incentive to retain incumbents and thus will likely have smaller organization capital in the following period as well. We describe these inter-temporal effects as *dynamic complementarities* (Cooper and Johri, 1997). We show that dynamic complementarities imply that organization capital is persistent and hence so are market-to-book ratios and turnover rates. This provides an explanation for the persistence of these variables that is found in the data. We also predict that organization capital persistence will be highest for firms with the largest organization capital, and we relate some consistent evidence. We show that firms can obtain high current period cash flow by over-firing (*e.g.*, through dramatic cost-cutting measures that may include massive layoffs) even though that is harmful for firm value in the long run because it destroys organization capital. Dynamic complementarities imply that expected payoffs are

higher for workers who begin their careers as juniors in large-organization-capital firms, since the organization capital will likely be transmitted to them by the time they produce. These results on dynamic complementarities would not arise in generic static models of firm-specific human capital.

The importance of transmitting the firm's organization capital has implications for optimal compensation. We argue that firms will reward incumbents who transmit the organization capital, even though this does not generate observable profits this period. Consequently, incumbents should have compensation that is less sensitive to current earnings, and this effect should be stronger in large-organization-capital firms.

Our model also sheds some light on the issue of the clash of corporate cultures in mergers. We suggest that in choosing between the languages of its constituent firms, the merged firm will weight two criteria; richer languages will be preferred, as will the language understood by more employees. 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 results on the persistence of organization capital, earnings-insensitive compensation for incumbents and merger value-creation arise from the dynamic feature of our model and distinguish our findings from those likely to be derived in a static model of firm-specific human capital with complementarities between managers.

The notion of organization capital that we develop does not encompass some assets that are sometimes thought of as organization capital, such as patents and brand name. We view patents, intellectual property and other legal privileges as best classified as physical assets (broadening this category to include intangible assets). Legally protected intangible assets share many of the qualities of physical assets (*e.g.* the firm's shareholders exercise complete control over their use and may transfer them at will) and these assets may naturally be grouped together. Brand name and reputation capital (Kreps, 1990, Hermalin, 2001) we think of as primarily representing signals of the quality of a firm's physical or human capital. If reputation capital had independent value as a separate type of asset, we would expect to see indiscriminate umbrella branding across product lines, which is not commonly observed. Moreover, competition between the many firms with strong brand names in the provision of

reputation and commitment services should drive the returns to reputation quite low. In this paper, we focus attention on understanding the nature of a firm's language, a firm resource that is neither controlled by the firm's owner nor part of the human capital of any single manager. The firm's language resides in the body of managers and is functional only when the managers work with the assets of the firm.

The fact that intangible assets or some organizational capital exists is perhaps non-controversial. Atkeson and Kehoe (2002) document that in the U.S. organizational capital has approximately two-thirds of the value of the stock of physical capital. The central question that we must persuasively answer in this paper is whether the notion of organization capital we develop constitutes a significant part of firm's organization capital. Some empirical regularities, such as persistence market-to-book ratios and profit rates, may follow from the presence of many different kinds of intangible assets. We derive, however, several empirical implications, such as those regarding managerial turnover, insider CEO succession, managerial compensation and merger integration, that arise distinctly from our notion of organization capital.

The rest of the paper is organized as follows: Section I details the role of firm language and communication in the production process. Section II describes the game of manager quitting and firm firing and provides a model of organization capital transmission. In Section III the existence of an equilibrium is proved. The main results are discussed in Section IV. Section V contains an application of the model to an analysis of compensation, and Section VI considers the issue of merger integration. Section VII concludes. The proofs of several results are given in the Appendix.

I. Language, Communication and Production

We model an infinitely-lived firm that employs managers who work for two periods. In a manager's initial work period neither the manager nor the firm knows the manager's type (quality) y , and the manager is referred to as a "junior". In a manager's second work period, the types of all managers are revealed to both managers and firms. Managers are referred to as "seniors" in this period.

A firm consists of $2N$ workers, N seniors and N juniors. Each worker produces some

output by virtue of his own quality and may produce more through communication with other workers in a manner to be described. The firm's physical capital will be fixed throughout the analysis in the paper.

A critical feature of a firm is its internal language that facilitates communication among its managers (Crémer, 1993)³. A language summarizes informal work routines, convenient technical jargons and a vocabulary of patterns remembered from past experiences. The language is private to the firm and its terms cannot be easily codified or conveyed without a significant investment of time on the part of a manager. Better communication amongst their managers will lead to increased production.

The composition of a firm's management team will be determined by a process of quitting and firing which we will detail below. We begin by describing the production of a firm with a given set of managers. For simplicity we will assume that juniors produce nothing and concentrate on the production of seniors.

In any given period, the firm is faced with carrying out a project or a task $k \in \{1, 2, \dots, K\}$ for some $K > 1$ that arises randomly according to a uniform random variable $\tilde{\epsilon}_t$ that is i.i.d. across time periods (with associated cumulative distribution function F_ϵ). For example, these tasks may be thought of as representing different strategic environments. Each manager i directly produces an individual output y_i that depends only on his own quality. The qualities of senior managers are revealed both to firms and to the manager's themselves at the beginning of the period. Manager i also may generate enhanced production through the use of the ideas of managers in the firm. A manager who shares a common language with colleagues may seek their counsel on improving the efficiency of his performance. We note that any new senior manager hired by the firm from the outside will not know the firm's language. Production enhancing communication is only possible when the language includes the particular project or task at hand. The task will be included in the language if the firm has performed that task in the past and the informal wisdom about how best to perform that task has been successfully transmitted to firm's incumbents by its past employees. Let \mathcal{L} denote a subset of $\{1, \dots, K\}$ which describes all the tasks which are part of firm's language

³DeMarzo, Vayanos and Zwiebel (2001) and Garicano (2000) provide other models of communication in organizations. Crémer, Garicano and Pratt (2003) analyze the optimal design of a code within an organization and consider implications for integration across different groups of agents.

and let L denote the number of such tasks. We define L to be the level of firm's organization capital.

If the given task is part of the language and the language is known by both managers i and j , then manager i will solicit the advice of manager j . Each manager generates identically distributed ideas with a total measure normalized to one and a density f on $[0, X]$ ($X > 0$) about the function performed by each of the managers in the firm. We assume, for technical convenience, that $f(x) > 0 \forall x \in [0, X]$. Ideas associated with higher values have a higher quality. We further assume that $f'(x) < 0 \forall x \in [0, X]$ to represent the notion that good ideas are scarce; there are more bad ideas than good ones.

Providing ideas is costless and manager i will approach all the other managers who know the firm's language and then select the best advice.⁴ We assume that the manager implements the best $\frac{1}{n}$ of the ideas of each of the n managers (including himself); formally, he implements all ideas $x > x^*$ where

$$\int_{x^*}^X f(x)dx = \frac{1}{n}$$

and receives as his payoff the average quality over all the ideas he implements. Formally, the idea-driven production g of a manager who collects ideas from n managers is given by

$$g(n) = n \left[\int_{x^*}^X x f(x)dx \right].$$

If the language is known by the subset $\mathcal{N}^{\mathcal{L}}$ of managers, then the total expected production of manager i with quality y_i is given by

$$Rev_i(y_i, \mathcal{N}^{\mathcal{L}}, \mathcal{L}) = \left[y_i + \Upsilon^i(\mathcal{N}^{\mathcal{L}}) \frac{L}{K} g(N^L) + \left\{ 1 - \Upsilon^i(\mathcal{N}^{\mathcal{L}}) \frac{L}{K} \right\} g(1) \right], \quad (1)$$

where Υ^i denotes the indicator function for whether manager i is in a given set and N^L denotes the number of managers who possess the language. So, if the manager does know the language – i.e., $\Upsilon^i(\mathcal{N}^{\mathcal{L}}) = 1$ – the probability is $\frac{L}{K}$ that the task undertaken is part of language in which case the manager benefits from getting advice from all N^L managers who possess the language and enhances his production by $g(N^L)$. With complementary probability $(1 - \frac{L}{K})$, the task is not part of language and the manager can benefit just from his own advice which

⁴Prescott and Visscher (1980) argue that casual conversations can transmit valuable information at a very low cost to productivity.

enhances the production by $g(1)$. Of course, if the manager does not possess the language at all – i.e., $\Upsilon^i(\mathcal{N}^{\mathcal{L}}) = 0$ – the idea-driven production equals $g(1)$ with probability one. The total idea-driven production of the firm when the selected task is part of the language and the language is known by n managers is given by $G(n) = ng(n) + (N - n)g(1)$.

Lemma 1. *Suppose that the selected task is covered by the firm’s language. The idea-driven production g of an individual manager who knows the firm’s language is monotonically increasing and concave in the number of managers who know the firm’s language, and the idea-driven production of the firm G is monotonically increasing and convex in the number of managers who know the language.*

A proof is found in the appendix.

The lemma shows that individual managers benefit from sharing the firm’s language with more managers (they thereby receive more ideas), but that these benefits are concave. In firms in which many managers know the language, the ideas of one more manager will not be that useful to any given manager. From a firm-wide perspective, however, the benefits to having more managers know the language are convex. As more managers know the language, not only does each individual manager benefit from more ideas, but more managers have high idea-driven production. The convexity relies upon our assumption that good ideas are scarce ($f' < 0$). It is the scarcity of good ideas that makes the addition of managers with firm’s language so valuable. If good ideas were common ($f' > 0$) then little added benefit would accrue from the ideas of new managers who know the firm’s language, and the firm-wide benefits would not be convex in the number of managers who share the firm’s language. We argue that most ideas are of fairly low quality, so that adding new managers with firm’s language can continue to generate significant benefits even when the pool of managers with firm’s language is already large.

We presume that a manager’s production is not verifiable and that the manager may, in an unverifiable manner, completely spoil or negate the value of his production. The firm, however, has a legal claim to any output generated from its physical assets. The firm and the manager must therefore bargain over the division of the output. In this Nash bargaining game, the firm pays the manager a fraction $\theta \in (0, 1)$ of the value of the production in

exchange for the remainder.⁵ The manager's expected payoff is then given by

$$\pi_i(y_i, \mathcal{N}^L, \mathcal{L}) = \theta \left[y_i + \Upsilon^i(\mathcal{N}^L) \frac{L}{K} g(N^L) + \left\{ 1 - \Upsilon^i(\mathcal{N}^L) \frac{L}{K} \right\} g(1) \right], \quad (2)$$

and the firm's total expected payoff this period is given by

$$\begin{aligned} \pi(\{y_i\}, \mathcal{N}^L, \mathcal{L}) &= \sum_{i=1}^N (1 - \theta) Rev_i(y_i, \mathcal{N}^L, \mathcal{L}) \\ &= (1 - \theta) \left[\sum_{i=1}^N y_i + \frac{L}{K} N^L [g(N^L) - g(1)] + N g(1) \right]. \end{aligned}$$

II. Quitting, Firing and Transmission of Organization Capital

The makeup of the set of managers who remain with the firm until the period end to engage in the production process described in the previous section is determined by a game of quitting and firing. At the beginning of the period, the firm's junior managers from the previous period are tentatively assigned the senior management slots. Their qualities $\{v_i\}$ are revealed to both themselves and the firm. We assume that $\tilde{v}_i \in [0, V]$ has associated pdf f_v (not dependent on i). A new batch of junior managers is then randomly selected. Each senior manager applies for a position at an outside firm and waits to receive an offer. If a manager accepts an offer then he leaves the firm. At the end of this quitting process, the firm receives applications for all senior manager positions and decides which incumbents to replace and which to retain. When the senior management team is in place, production begins.

Junior managers are interchangeable and indistinguishable, so firms will retain their initial set of juniors and hence no junior will leave the firm. The firm's new senior managers

⁵Following 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. In our model, however, a bargaining scheme based on manager's incremental contribution to firm's total output can generate a total wage bill for the firm that exceeds the total output of the firm. This is because the total output of the firm is convex in number of managers who know the firm's language. In our model, the direct output by managers is additive and thus linear in the number of managers which does not lend itself to determination of optimal number of productive employees N in the firm. We can conceive of an extension of our model in which the total direct output was sufficiently concave in the number of productive employees N which would then allow us to determine the optimal N . In such a model, it would be feasible to implement the bargaining scheme suggested in Stole and Zwiebel. For simplicity, we make use of our exogenously specified split of the manager's output.

each apply to one outside firm and await offers. Outside offers arrive randomly and are immediately accepted or rejected. We presume that no two offers arrive at the same time, so that all decisions are sequential. No manager receives more than one offer. For managers without any outside offer, the opportunity cost of working for the firm is presumed to be zero. Managers know the identities and decisions of all managers who have received prior offers.

A senior manager's strategy specifies the manager's decision to stay or leave the firm given the amount and timing of the offer, the prior decisions of other managers and the quality of the other managers (which will affect the likelihood of their remaining in the firm), and the level of the firm's organization capital. We denote manager i 's strategy by d^i , where we represent the decision to stay by $d^i = 1$ and the decision to leave by $d^i = 0$.

Some incumbent managers will choose to stay with the firm and others will not be given outside offers. Let \mathcal{I} denote the set of managers who choose to stay. The firm's strategy specifies which of these incumbent managers will be replaced. The firm's replacement decision will depend on the qualities $\{v_i\}$ of the incumbents, the qualities $\{v'_i\}$ of the replacements and the level of the organization capital. We assume that $\tilde{v}'_i \in [0, V]$ has associated pdf f_v (not dependent on i). Formally, given a set \mathcal{I} of remaining incumbents, applicant qualities $\{v'_i\}$, incumbent qualities $\{v_i\}$ and level L of the organization capital, we denote the firm's strategy by

$$\mathcal{I}^{\mathcal{R}}(\mathcal{I}, \{v'_i\}, \{v_i\}, L)$$

which specifies the set $\mathcal{I}^{\mathcal{R}}$ of incumbent managers that the firm chooses to retain.

The firm's juniors do not produce, but they do observe the functioning of the seniors. We presume that it is 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 (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). 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. The probability that this language is transmitted is dependent on

the number of seniors who know the language. The greater the number of managers with the language, the more opportunities the juniors will have for observation and learning and hence the greater the probability of transmission. Formally, the probability of transmission is given by p where p is increasing and convex. We let $\tilde{\xi}$ be a Uniform(0,1) random variable with associated cumulative distribution function F_{ξ} . Transmission takes place if and only if $\tilde{\xi} \leq p(N^L)$, where N^L is the set of seniors who know the firm's language. For example, the probability of transmission may depend on the number of manager pairs who know the language, since this determines the number of conversations that can take place. In that case we would have $p(N^L) = \frac{\binom{N^L}{2}}{\binom{N}{2}}$. Our model thus suggests that both organizational learning and organizational forgetting (Benkard, 1999) are possible. The transmission of language depends on the presence of incumbents. We are using the word "language" here as a convenient term for tacit knowledge that facilitates communication and that can be only be acquired by experience. In the terminology of Rajan and Zingales (1998), only incumbent senior managers have access to the language of the firm.

In all cases, the firm's current juniors observe the firm's production given the selected task. If the juniors learn the firm's language, and this task is covered by the language, then the juniors' language is identical to that of the seniors. If the juniors imbibe the firm's language and the task is new, then the juniors will develop terms for this task that are based on the language of the firm and augment the firm's language to cover this new task as well. If the juniors do not imbibe the firm's language, they develop a new language which will provide terms only for the one task observed. The firm may, if it wishes, choose new versions of any task. The firm's language will not provide terms for any new version selected in this way.⁶

Summary of the timing:

1. A firm's junior managers from the previous period are initially assigned to be the senior managers of the firm. The qualities of the new senior managers are revealed. New junior managers are randomly selected.
2. Each senior manager applies to an outside firm and waits to receive an outside offer.

⁶That is, we essentially assume that there is free disposal of the firm's organization capital. An alternative approach would be to model potentially destructive aspects of organizational capital but our focus is on the value creation arising from intrafirm communication.

Offers are accepted or rejected.

3. The firm receives applications for the senior manager positions and decides which incumbents to replace and which to retain.
4. A task arrives.
5. Production and bargaining over output division takes place.
6. The firm's managers attempt to transmit the firm's language, *i.e.*, its organization capital.

III. Equilibrium

Before the beginning of each period, the managers' types, the future offers to managers, the applications from replacements and the task are all unknown to the firm. The expected payoffs of the managers and the firm depend only on the firm's language, so we may write the firm's expected one-period profit as

$$Pr(\mathcal{L}) = E \left[\pi(\{\tilde{y}_i\}, \tilde{\mathcal{N}}^{\mathcal{L}}, \mathcal{L}) \right],$$

where $\tilde{y}_i = \tilde{v}_i$ if $i \in \tilde{\mathcal{N}}^{\mathcal{L}}$ and $\tilde{y}_i = \tilde{v}_i'$ otherwise. The firm plays the production game repeatedly and discounts the future at a discount factor $\delta \in (0, 1)$. The firm's total payoff is given by

$$E \left[\sum_{t=0}^{\infty} \delta^t Pr(\mathcal{L}_t) \right]. \quad (3)$$

In general, the equilibrium of the game may consist of the firm and managers playing strategies that depend on time and possibly on the entire history of the firm's and managers' decisions. We assume for simplicity that the exogenous variables are mutually independent and identically distributed over time. Our first result is to show that there exists an equilibrium of the game in which the strategies of the firm and managers are Markov, *i.e.*, stationary with respect to time. We focus our attention on Markov Perfect Equilibria (MPE) of the game.

Since the per-period firm payoff π is bounded, the state space of languages is finite and the exogenous variables are independent over time, the firm's total expected payoff may be

written as a value function $w(L_t, \{d^i\}, \mathcal{I}^R)$. To find an MPE we begin by associating every feasible value function w with a profile of Markov strategies $(\{d^i\}(w), \mathcal{I}^R(w))$ by considering the optimal one-period strategies of the agents given that the continuation payoffs are described by w .

We first determine the firm's optimal strategy. We let the set of remaining incumbents be given by \mathcal{I} (with cardinality I) and the tasks covered by the firm's language by \mathcal{L}_t . The firm's continuation payoff (the discounted expected value of all its future payoffs) depends only on the number of incumbents retained. For a fixed number of retained incumbents, the firm maximizes its payoff by retaining the incumbents with the highest values of $(v_i - v'_i)$, since this maximizes the current payoff. Reorder the managers such that $(v_{i+1} - v'_{i+1}) \geq (v_i - v'_i)$ for all $i \in \{1, \dots, I\}$. To select its optimal strategy the firm maximizes

$$\begin{aligned} \hat{\psi}(\mathcal{I}, \mathcal{L}_t, w) =: \max_{m \in \{0, 1, \dots, I\}} & \\ \delta E \left[p(m) \max_{\mathcal{L}^1 \subseteq \mathcal{L}_t \cup \{\tilde{k}_t\}} w(L^1) + \{1 - p(m)\} \max_{\mathcal{L}^2 \subseteq \{\tilde{k}_t\}} w(L^2) \right] & \quad (4) \\ + (1 - \theta) \left[\frac{L_t}{K} m [g(m) - g(1)] + Ng(1) + \sum_{n=1}^{I-m} v'_n + \sum_{n=I-m+1}^I v_n \right]. & \end{aligned}$$

The second line in (4) reflects the present value of future payoffs. Future payoffs depend only on whether the language is transmitted. The future payoffs to having a specific language are given by w . If the firm decides to retain m managers, the current language will be successfully transmitted with probability $p(m)$ and in addition the firm's managers will develop terms for talking about the task that is currently undertaken so that the next period the set of tasks for which there will be developed terms is given by $\mathcal{L}_t \cup \{\tilde{k}_t\}$. With probability $\{1 - p(m)\}$, the firm will fail to transmit the language and the next period the managers will have developed terms in the language to communicate only about the task the firm undertakes this period. The firm may choose a smaller subset of any transmitted language by selecting new versions of any task. The third line in (4) gives the current period payoff. We note that the symmetry across tasks shows that $\hat{\psi}(\mathcal{I}, \mathcal{L}_1, w) = \hat{\psi}(\mathcal{I}, \mathcal{L}_2, w)$ for all $\mathcal{L}_1, \mathcal{L}_2$ such that $L_1 = L_2$. For $l \in \{0, 1, \dots, K\}$, we can therefore define $\psi(\mathcal{I}, l, w) := \hat{\psi}(\mathcal{I}, \mathcal{L}, w)$ for some \mathcal{L} such that $L = l$.

Given this optimal firm strategy, we can determine the optimal strategies of the managers. Details are provided in the appendix in the proof of Result 1. Essentially managers evaluate

the probability that they will be retained and the expected number of other incumbents who will be retained in order to calculate the value of remaining with the firm. They then compare this value to their outside offer and choose the better option.

The managers and the firm are playing a dynamic game. Within each period the managers' strategies are constructed to be best responses to the strategies of the firm and the other managers. To show that the firm's strategy is a best response, we show that for fixed manager strategies the firm's optimization problem may be viewed as a single-agent dynamic programming problem, and the firm's strategy described above solves that problem.

These arguments show:

Result 1. *There is a Markov Perfect Equilibrium of the game.*

Details of the proof are given in the appendix.

We now seek to characterize these equilibria. The first question of interest is whether a larger organization capital generates a higher firm value.

We require two results.

Lemma 2. *A history of incumbent decisions to remain makes it likelier that other incumbents will remain and that the firm will choose to retain them.*

A formal statement of Lemma 2 and its proof is given in the appendix.

This effect arises from the complementarity between incumbent managers that is described in Lemma 1. Individual incumbents derive more benefit from remaining in the firm when other incumbents stay since they can then receive advice from more managers. Moreover, since firm-wide idea-driven production is convex in the number of incumbents, the firm is also more inclined to retain incumbents when there are more of them. If firm-wide idea-driven production were concave in the number of incumbents, then this lemma would not necessarily hold. Low quality incumbents might be discouraged by a history of remain decisions by other incumbents, concerned that they would be expendable. This might lead to more quit decisions. It is the scarcity of good ideas and the resulting complementarity of incumbents (from the perspective of both the individual managers and the firm) that underlies this lemma.

Lemma 3. *A larger organization capital induces incumbents to remain with the firm and encourages the firm to retain them.*

A formal statement of Lemma 3 and its proof is given in the appendix.

It is clear that if the organization capital is more valuable, then both the incumbent and the firm benefit more from the incumbent producing in the firm. Moreover, this effect will make it likelier that the first incumbents to receive offers will choose to remain. Lemma 2 then shows that the incumbents who receive later offers are then even more likely to remain.

We now obtain the following result:

Result 2. *In any MPE, firms with larger organization capital will have higher firm values.*

A proof is given in the appendix.

For a given set of managers, it is clear that the firm's production is increasing in the level of its organization capital. Lemmas 2 and 3 also show that firm's with larger organization capital retain more incumbents, leaving them with a larger pool of potential employees from which to pick. These effects together generate Result 2. It is worth pointing out that a firm may end up retaining an incumbent whose quality v_i is less than the quality v'_i of a feasible replacement because the incumbent helps enhance the production of other incumbent employees who share the firm's language and aids in language transmission.

A natural question is how to empirically measure organization capital. We will discuss two empirical proxies here and an additional two after presenting Result 3. The most direct measure of the level of a firm's organization capital in our model is the importance and frequency of employee interactions as reported by the employees themselves. While collecting such data is not always straightforward, there is an increasing number of intrafirm studies on the value of communication (Ichniowski and Shaw, 2003). A second proxy is a firm's geographical dispersion. Firms with scattered work sites are likely to have smaller organization capital since communication is more difficult across great distances.

Result 2 provides us with a prediction relating a firm's organization capital to its market value. A firm's market-to-book ratio will thus be associated with the quality of its language. Huselid (1995) and Ichniowski (1990) show that firms with more workplace communication and training have higher market-to-book ratios, which directly supports Result 2. Ichniowski, Shaw and Prensushi (1997) and Macduffie (1995) provide plant- and production line-level evidence that group problem solving and teamwork are associated with greater productivity.

Of course, a firm's market value or its market-to-book ratio may capture the value of its other intangible assets (such as reputation or a brand) as well. One needs to control for industry characteristics as well as firm characteristics such as the amount of physical capital,

growth prospects, firm age, level of R&D, firm beta etc., in detecting the link between market value and our notion of organization capital empirically.

IV. The Results

We now derive other results of our theory for firm behavior. We select and fix a value function w associated with an MPE that is consistent with Result 2. The convexity of firm production in the number of managers knowing the language, as described in Lemma 1, generates static complementarities between the firm's set of incumbents. We begin by considering the implications of these static complementarities for employee retention and wages.

A. Static Complementarities

We obtain the following result:

Result 3. *Firms with larger organization capital will experience fewer quits and greater employee retention.*

A formal statement of Result 3 and its proof is found in the appendix.

In firms with larger organization capital employees are reluctant to quit and the firm is averse to replacing them with outside applicants because employees who know the firm's language produce effectively within the firm, provide advice to other employees, and help to transmit the firm's language to the current juniors.

Result 3 suggests two additional empirical proxies for organization capital. First, firms in which managers have higher average within-firm tenures are likely to be firms with larger organization capital. Second, firms with more frequent insider CEO succession will typically have larger organization capital.

It is a basic implication of our model of organizational capital that the hiring of outsiders to senior positions should be relatively rare, since outsiders do not speak the language of the firm. This is consistent with the evidence of Denis and Denis (1995) that only 15 percent of firm top executive appointments are made to external candidates.

Result 3 also provides some other, sharper predictions. Organization capital is likely to be most important in industries in which human capital and communication are central

to production. Result 3 therefore suggests that insider CEO succession should be most common in these industries. Consistent with that prediction, Parrino (1997) finds that outside succession occurs most frequently in commodity industries in which organizational capital is likely to be less important.

Results 2 and 3 together suggest that market value and employee retention rates should be positively correlated. This prediction is supported by Hanka (1998) who finds that firms with high market-to-book ratios have fewer employment reductions in the following year, controlling for current and historic firm cash flows. In our model firms with high market-to-book ratios are firms with large organization capital (Result 2), and these firms are predicted to have high employee retention rates (Result 3), as was found by Hanka. We note that if high market-to-book ratios were driven only by other intangible assets, such as a brand, it would have no direct relation to employee retention rate or turnover.

This prediction is also consistent with the work of Bassi and Van Buren (2000), who find that firms who retain key personnel have higher (contemporaneous and future) values of (market capitalization - book value) than industry peers. They also find that the ability to *attract* talented employees is not correlated with (market value- book value), which suggests that the correlation of retention rates and market values is not driven purely by the appeal of working for a firm with a large market value.

It is a novel and untested implication of Result 3 that firms with high market-to-book ratios should have high rates of internal CEO succession.

For firm A, we define $I_t^R(L_A)$ to be the number of managers the firm retains in period t (the optimal m in (4)). The expected number of quits that Firm A with language level L_A will experience in period t is given by $q_t(L_A) := N - E_z [I_A(w, L_A, \tilde{z})]$.

Result 4 analyzes the effect of one manager's quitting on the decisions of other managers.

Result 4. *Firms will experience cascades in quitting.*

A formal statement of Result 4 and its proof is found in the appendix.

Result 4 predicts that firms that experience an unusually high (respectively low) number of quits in the first part of the period are likely to experience an unusually high (respectively low) number of quits in the rest of the period. Since incumbents benefit from the presence of other incumbents, a number of incumbent quits will lead other incumbents to quit. This prediction will also arise in an alternative model in which quitting signaled bad prospects

about firm's future performance.

We define a manager's wage premium to be the difference between his wage and the wage he would receive with no help from other managers. Our theory has several implications for inter-firm wage differentials.

Result 5. *a) The realized wage of a manager with a fixed observable quality increases in the level of the firm's organization capital, his own tenure and the tenure of other managers.*

b) The expected wage premium averaged across managers is higher for managers in firms with larger organization capital.

A formal statement of the result and its proof is given in the appendix.

Wages increase with organization capital and the tenure of the managers because richer languages facilitate the exchange of productive ideas, and communication is only possible amongst managers who served in the firm as juniors and share a knowledge of its language. Since managers receive a share of output, their wages reflect productivity improvements. Result 5a) provides empirical predictions in the case that the tenure of all managers is observed. We predict that wages will be higher in high market-to-book firms, since the large organization capital of these firms will benefit both managers and the firm as a whole. We predict that wages will rise with tenure in the firm (Becker, 1993, Topel, 1991), and we make the new and untested prediction that wages will rise with the tenure of other managers in the firm. Result 5a) is consistent with the finding of Dustmann and Meghir (2002) that wages increase with firm tenure, but not sector tenure.

Result 5b) provides a prediction about wages that is not conditional on managerial tenure. It suggests that average wages should be higher in firms with larger organization capital (i.e. firms with more communication, high book-to-market ratios, low turnover).

B. Dynamic Complementarities

The results on static complementarities that we described in the previous subsection might arise in many static models of firm-specific human capital with complementarities among managers. We now turn to the more specific predictions arising from dynamic aspect of our model. In our setting there are dynamic complementarities between investments in organization capital in different time periods. If a firm invested heavily in its organizational capital in the previous period (by retaining many incumbents), then it will likely have a large

organization capital this period, which provides it with a strong incentive to again invest in preserving its organization capital. The results in this section explore the implications of these dynamic complementarities.

The first implication is that organization capital is persistent; firms that have large organization capital this period tend to have large organization capital in future periods. A firm that has a large organization capital will value both the current and future benefits generated by this organization capital, and this firm will therefore will be more willing to retain incumbents who speak the firm's language but who have low personal qualities. Incumbents will also prefer to remain in this firm, where they are relatively productive. This increased retention of incumbents makes organization capital transmission to the juniors more probable, so the firm will likely have a large organization capital next period as well.

Result 6. *a) Organization capital is persistent: for all $s > 0$, $E[L_{t+s}|L_t = L]$ is increasing in L .*

b) Firm value and retention policies are persistent: for all $s > 0$, $E[w(L_{t+s})|L_t = L]$ and $E[I^R(L_{t+s})|L_t = L]$ are increasing in L .

c) Greater retention this period leads to greater expected firm value and retention in future periods: for all $s > 0$, $E[w(L_{t+s})|I^R(L_t), L_t = L]$ and $E[I^R(L_{t+s})|I^R(L_t), L_t = L]$ are increasing in $I^R(L_t)$.

Result 6 part c) does not imply that it is optimal for firms to retain all their incumbents. The firm may want to fire low-quality incumbents this period to enjoy higher current cash flows in exchange for lower future cash flows.

A basic empirical implication of Result 6b) is that firms' book to market ratios should be very persistent, a fact that is strongly supported by evidence (*e.g.* Fama and French, 1995). Cohen, Polk and Vuolteenaho (2003) show that, controlling for industry effects, book to market ratios are highly persistent at leads of 15 years. They also show that low book to market firms (firms with low organization capital in our model) are significantly less profitable at 15-year leads, which is consistent with our interpretation that large organization capital firms are more valuable because they will produce greater profits. Mueller (1990) and Maruyama and Odagiri (2002) show that long-run profit rates (relative to book value of assets) are highly persistent, even at leads of more than 30 years. This evidence that firms are able to maintain comparative advantages for long periods of time is consistent with our

view that organizational capital is valuable and that it is transmitted from one generation of managers to their successors. The evidence is, however, also consistent with other models of organization capital, such as brand name, that may be persistent over time.

Lane, Isaac and Stevens (1996) provide evidence that is consistent with the second part of Result 6 b). They show that there is significant variation in turnover rates across firms and that firm turnover rates are persistent, a finding that they regard as not well explained by the existing theoretical literature. They also find that, controlling for changes in size of workforce, firms that experience persistently high turnover are more likely to fail. This is consistent with our model's prediction that firms with small organization capital (presumably the firms most at risk for failure) are likely to experience high turnover rates (Result 3). Models with other intangible assets as organization capital do not naturally generate implications for turnover rates.

Result 7 shows that the large levels of organization capital are highly persistent. Very small levels of organization capital may also be very persistent. Defining P to be the underlying probability measure, we state Result 7.

Result 7.a) *Among firms with some level of organization capital, a larger level of organization capital is more persistent: for all $L \geq 2$, $P(L_{t+1} = L | L_t = L)$ is increasing in L .*

b) Loss of organization capital is less likely at better firms: for all $L \geq 0$, $P(L_{t+1} = 1 | L_t = L)$ is decreasing in L .

c) Very small level of organization capital can be persistent: if $p \leq \frac{K}{K+1}$ then $P(L_{t+1} = 1 | L_t = 1) \geq P(L_{t+1} = 2 | L_t = 2)$.

A high level of organization capital is very persistent because firms with large organization capital are more likely to transmit their organization capital and are more likely to face a new task that is covered by their language. Large organization capital is transmitted with greater probability because incumbents are highly valuable in firms with large organization capital both in current production and in transmission of organization capital and are therefore more often retained. Result 7 c) suggests that the probability of organization capital persistence may be u-shaped. A firm with a very small organization capital will have little interest in retaining its incumbents and may therefore languish in a state of low organization capital.

Levonian (1994) and Mueller (1986) report that the persistence of profit rates is highest

for the most profitable firms. That is, the most profitable firms exhibit the slowest reversion to industry-mean profit rates, a finding that is consistent with Result 7 a) and b).

Result 8. *Firms can generate higher than optimal current period cash flows by over-firing, but not by under-firing.*

A formal statement of Result 8 and its proof is found in the appendix.

The intuition for Result 8 is that over-firing can generate higher current profits by bringing high-quality replacements into the firm. The cost of over-firing is that it makes organization capital transmission less likely, thereby effectively reducing the capitalized value of the firm's organization capital. The firm is essentially ignoring the dynamic complementarities generated by its organizational capital. Under-firing invests too much in preserving the firm's organization capital. The retention of low quality incumbents will result in a lower profit this period.

Result 8 has the corporate governance implication that CEO compensation should not be based solely on current profits and measures of physical capital, since this neglects the organization capital of the firm. Sophisticated board monitoring is required to disentangle the net effect of layoffs on current profits and on the value of the firm's organization capital.

Result 8 also suggests if CEOs maximize current profits and not total firm value (or if they neglect organization capital in making firing decisions), then, controlling for current profits, the extent of current firings should be positively related to short-term future profits and negatively related to long-term future profits.

Hallock (1998) and Chen *et al.* (2001) find that the share price reaction to layoff announcements is negative. Chen *et al.* also find, however, that earnings and profit margins increase significantly in the first three years after a layoff announcement.⁷ This pattern of layoffs generating increased short-term profits while destroying firm value is consistent with Result 8 and an assumption that CEOs overweight current earnings in making labor decisions.

The organization capital of a firm affects both its current seniors and prospective managers.

Result 9. *The expected payoff is higher for managers who begin the period as juniors in*

⁷Firms' improved short-run profitability following the announcement suggests that layoffs do not signal short-term negative news but rather some destruction of long-term value.

firms with larger organization capital.

A formal statement of the result and its proof are given in the appendix.

Result 9 shows that there is a value to being a junior in a firm with a large organization capital, since the organization capital will likely be large in the following period. A simple extension of our model would 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 (Result 5a)). Our theory thus predicts that the gap in compensation between juniors and seniors will be greater in firms with large organization capital.⁸

V. Earnings-Insensitive Pay

We now extend the model to consider the case in which transmitting the firm's organization capital requires additional effort on the part of seniors. We will assume that seniors must receive fixed compensation u offsetting their effort costs to induce them to attempt to transmit the organization capital to the juniors. We will refer to this payment as *earnings-insensitive* since it is unrelated to the physical output generated by the seniors this period. Since the earnings-insensitive compensation leaves the seniors indifferent about attempting to transmit the organization capital, it will have no effect on their decisions to quit or remain with the firm. The firm can choose how many seniors to retain and it may then choose how many retained seniors will receive the payment u . Replacement seniors will clearly not receive any performance insensitive payment since they cannot facilitate transmission. The convexity of the transmission function p shows that either all or none of the retained incumbents will receive u . The firm's problem is to maximize

$$\hat{\psi}_u(\mathcal{I}, \mathcal{L}_t, w) =: \max_{m \in \{0, 1, \dots, I\}} \max_{i \leq m} \left\{ \delta E \left[p(i) \max_{\mathcal{L}^1 \subseteq \mathcal{L}_t \cup \{\tilde{k}_t\}} w(L^1) + \{1 - p(i)\} \max_{\mathcal{L}^2 \subseteq \{\tilde{k}_t\}} w(L^2) - iu \right] \right\} \quad (5)$$

⁸The results on static and dynamic complementarities that we present suggest large organization capital firms will have the following characteristics: persistently low turnover (even during economic downturns), high market-to-book ratios, relatively high wages, wages steeply increasing in within-firm tenure, and well-developed mentoring/coaching programs. Press and trade journal accounts suggest that some examples of such firms are Microsoft, Southwest Airlines, Marriott International, McKinsey, and PricewaterhouseCoopers (*Christian Science Monitor*, September 8, 2003, Features, p. 18 and *Consulting Magazine*, August 2001.)

$$+(1 - \theta) \left[\frac{L_t}{K} m [g(m) - g(1)] + Ng(1) + \sum_{n=1}^{I-m} v'_n + \sum_{n=I-m+1}^I v_n \right].$$

Result 10 describes the firm's optimal policy in paying earnings-insensitive pay.

Result 10. a) *Replacement seniors will not receive any earnings-insensitive payment.*

b) *The probability that earnings-insensitive compensation is paid is increasing in organization capital level L .*

c) *Organization capital transmission occurs with greater probability when earnings-insensitive compensation is paid.*

A proof of the result is found in the appendix.

Firms with larger organization capital have a greater interest in paying earnings-insensitive compensation to incumbents in order to ensure the transmission of their valuable organization capital. When no earnings-insensitive wages are paid, incumbents will not work to transmit the organization capital and transmission is therefore less likely.

Result 10 a) suggests that the compensation of incumbents should have a relatively low weight on earnings. This is consistent with Barro and Barro (1990) who find that compensation growth is less sensitive to firm accounting performance for CEOs with greater tenure. Bushman, Indjejikian and Smith (1996) find that individual performance evaluation plays a larger role in determining the CEO bonus in firms with high market-to-book ratios. Since compensating managers for organization capital transmission requires individual performance evaluation and not simply using accounting measures, their finding accords with Result 10 b). Bushman, Indjejikian and Smith also find that individual performance evaluation is important for CEO's with greater tenure, as suggested by Result 10 a).

We denote the total expected wages paid by the firm this period by

$$Wg(\mathcal{L}) = E \left[\sum_{i=1}^N \pi_i(\tilde{y}_i, \tilde{\mathcal{N}}^{\mathcal{L}}, \mathcal{L}) \right].$$

Result 11 describes the relationship between the firm's revenues and the wages it pays.

Result 11. *The firm's expected per-period wages and profits satisfy the following relationship*

$$Wg(\mathcal{L}) = \left(\frac{\theta}{1 - \theta} \right) Pr(\mathcal{L}) + \rho(L),$$

where ρ is an increasing function.

A proof of the result is found in the appendix.

Consider regressing wages on earnings for high- and low-organization capital firms separately. Result 11 implies that the slopes in the two regressions will be the same, but the intercept will be higher in the high-organization capital firm sample. The $\left(\frac{\theta}{1-\theta}\right) Pr(\mathcal{L})$ term represents the fraction of output that managers receive from bargaining with the firm. Static (and dynamic) complementarities generate the result that $\left(\frac{\theta}{1-\theta}\right) Pr(\mathcal{L})$ is increasing in the cardinality of \mathcal{L} . The $\rho(L)$ term represents the wages that the firm is willing to pay to transmit its organization capital. These wages are higher for high- L firms because of dynamic complementarities. Transmission is valuable independent of current performance, and the firm will pay managers to maintain its organization capital.

VI. Mergers

We now consider a model of mergers between firms with overlapping tasks. The role of language (i.e. organization capital) transmission will be central to our analysis. For simplicity we will presume that all the tasks of the two firms are identical. The two firms are presumed to possess different languages. We assume for simplicity that mergers are arranged after quitting and firing and before task selection and production. In the transition period, the seniors of the two firms produce as they would in separate firms, and the juniors select which of the two languages they will learn. We will presume that coordination is critical to the language learning process. We model this process in the following way. The juniors are ordered randomly. They each select, in turn, the language to be learned. If all juniors select the same language and this language is known by $\mathcal{N}^{\mathcal{L}}$ seniors, then the language in all its richness is transmitted with probability $\hat{p}(\mathcal{N}^{\mathcal{L}}) \in [0, 1]$. We assume that \hat{p} is increasing and convex. If the juniors choose the same language and it is not transmitted, they still learn the terms for the task chosen this period. If the juniors do not all choose the same language, no language is transmitted.⁹

This 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. In our model, the choice of which language to adopt is left to the junior managers.¹⁰

⁹Crémer, Garicano and Pratt (2003) analyze a model in which two firms may choose to adopt a common code at some cost.

¹⁰Another model would be one in which the senior managers of the acquiring firm simply select their own

Since there is a payoff to learning a language, all juniors will follow the language choice of the first junior. The value created by a merger is equal to the value of the merged firm minus the values of the two constituent firms.

Result 12. *a) The probability that a constituent firm's language is adopted by the merged firm is increasing in the level of its organization capital and the number of its retained incumbents. b) Value creation is decreasing in the level of the organization capital of the constituent firm whose language is not adopted.*

A formal statement of the result and its proof is found in the appendix.

The juniors prefer the language of the firm with larger organization capital since it will more likely cover the task they have to perform. They prefer the language shared by more incumbents as it is more likely to be transmitted. Since the organization capital of the firm whose language is not adopted is simply lost, the value created by the merger is greater when this organization capital is small.

Result 12a) predicts that the probability that the merged firm will take on the characteristics of a given constituent firm increases in the market-to-book ratio and number of employees of the constituent firm. These characteristics include the market-to-book ratio and retention rates of the constituent firm.

For example, consider the merger between a small firm with a high retention rate and a large firm with a low retention rate. Result 12a) predicts that the probability that the merged firm will adopt the retention rate of the small firm is increasing in both the size and retention rate of the small firm. Result 12a) also predicts that the merged firm is more likely to adopt the high retention rate of the small firm for *lower* values of the retention rate of the large firm.

Result 12b) shows that the most efficient mergers are between firms with highly disparate levels of organization capital. The difference between the market value of the merged firm and the sum of the market values of the constituent firms should be largest when the merger is between firms with very different market-to-book ratios, as long as the firm with higher market-to-book ratio is sufficiently large. Lang, Stulz and Walkling (1989) and Servaes (1991) show that total returns on merger announcements are larger when target firms have

constituent firm's language as the language of the merged firm for their own benefit, ignoring considerations of firm value. In this case our results below will indicate the costs associated with this agency problem.

low market-to-book ratios and bidders have high market-to-book ratios. The bidders in their samples are much larger than the targets, so if one assumes that the merged firm adopts the language of the bidder, as seems reasonable, then their findings are consistent with Result 12 b).¹¹

It is also an implication of Result 12 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.

We will say that a merger is a failure if the organization capital of the merged firm is strictly lower than the organization capital of each of the constituent firms.

Result 13. *a) The probability that a merger is a failure is increasing in the organization capital of the smaller firm. b) If the level of organization capital of the larger firm is above one, the probability that a merger is a failure is decreasing in the level of organization capital of the larger firm.*

A formal statement of the result and its proof is found in the appendix.

As the level of the organization capital of the small firm increases, it becomes more likely that the merged firm will attempt to adopt the small firm language. This leads to a greater risk of failure, since the small firm language is adopted only with a relatively low probability. For a small firm with sufficiently large organization capital, this risk is worth taking.

If the larger firm has a very low level of organization capital (i.e. $L = 1$), the merger cannot be judged a failure, irrespective of its outcome, so increasing the level of organization capital of the large firm only reduces the risk of failure when the large firm level of organization capital is above one.

We note that, in general, mergers will 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.

¹¹The extent to which these gains are permanent is unclear (Rau and Vermaelen, 1998, Mitchell and Stafford, 2000, and Gregory and McCorriston, 2002).

VII. Conclusion

This paper models the organization capital of a firm as the quality of its internal language and shows that firms with larger organization capital will have higher market values. As a result of the static complementarities between incumbents induced by firm organization capital, our theory predicts that employee retention and insider managerial succession will be greater in firms with larger organization capital. We suggest that differences in organizational capital provide a rationale for inter-firm wage differentials between managers with identical observable qualities. We also show that firms that have preserved their organization capital in previous periods by retaining incumbents have a greater incentive to preserve their organization capital in the current period. This dynamic complementarity between inter-temporal investments in organization capital generates long-run persistence in market-to-book ratios and turnover. We argue that this persistence is likely to be strongest among firms with the largest organization capital. We demonstrate that the optimal compensation of incumbents will include an earnings-insensitive component, and that this component will be larger in firms with larger organization capital. In a simple model of mergers, we show that value creation is greatest in the merger of two firms with different organization capital, as long as the firm with the larger organization capital is sufficiently large.

Our model describes a firm's language as its organization capital. This description of organization capital 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. Our notion of organization capital represents a form of firm-specific human capital (Becker, 1993) since the firm's internal language has no value to a manager who leaves the firm. 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. Our analyses of the transmission

of organization capital and its dynamic evolution generate predictions, however, that do not naturally arise in models of static firm-specific human capital with complementarities among managers. In particular, our results on the persistence of organization capital, the long-term costs of over-firing, earnings-insensitive pay and the success of mergers are driven by the dynamic effects we explore.

For organization capital to have value, it must also be costly for competitors to replicate (Rumelt, 1987). Our description of the firm’s organization capital ties it to information possessed by the organization as a whole. We argue, consistent with Prescott and Visschler (1980), that information processing is a central function of the firm. Inimitability 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 (Arrow, 1962; Rosen, 1972).¹² These features combine to make the acquisition of language within the firm time-consuming and difficult.

Our model of organization capital provides a unified framework linking firm market value, labor practices, compensation and the outcome of mergers. The organization capital we describe is difficult to measure directly but has potentially important effects on some of the central characteristics of the firm. We argue that providing a structure for the sharing and transmission of advice and new ideas is a central purpose of organizations.

Appendix

Proof of Lemma 1:

We define $F(x) = \int_0^x f(x)dx$. We may rewrite $g(n) = n \int_{\frac{n-1}{n}}^1 F^{-1}(y)dy$. For convenience we define $H(y) = F^{-1}(y)$, giving $g(n) = n \int_{\frac{n-1}{n}}^1 H(y)dy$. It can quickly be checked that $H' > 0$ and $H'' > 0$ (since $f > 0$ and $f' < 0$). We find

$$g'(n) = \int_{\frac{n-1}{n}}^1 H(y)dy - \frac{H(\frac{n-1}{n})}{n} > 0,$$

¹²Bahk 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.”

where the inequality follows from the fact that $H' > 0$. Furthermore,

$$g''(n) = \frac{-H'(\frac{n-1}{n})}{n^3} < 0.$$

It is clear that G is increasing. We find that

$$G''(n) = 2 \left(\frac{-H(\frac{n-1}{n})}{n} + \int_{\frac{n-1}{n}}^1 H(y) dy \right) - \frac{H'(\frac{n-1}{n})}{n^2}.$$

The convexity of H shows that $\forall y \in [\frac{n-1}{n}, 1]$

$$H(y) \geq \left(y - \frac{n-1}{n}\right) H' \left(\frac{n-1}{n}\right) + H \left(\frac{n-1}{n}\right).$$

This implies

$$\int_{\frac{n-1}{n}}^1 H(y) dy \geq \frac{H'(\frac{n-1}{n})}{2n^2} + \frac{H(\frac{n-1}{n})}{n},$$

which shows that $G''(n) > 0$.

Proof of Result 1:

We denote the arrival time order of manager i 's offer o_i by $\tilde{T}_i \in \{1, \dots, N, \emptyset\}$ (if $\tilde{T}_i = \emptyset$ then manager i receives no offer). We assume that the $\{\tilde{T}_i\}$ have a joint probability distribution function given by $e_{\{\tilde{T}_i\}}$. We assume that $\tilde{o}_i \in [0, O]$ has an associated pdf f_o (not dependent on i). Given an outside offer o_i , offer arrival order time t , a set \mathcal{N}^O of managers who have previously received outside offers, a set $\mathcal{N}^A \subseteq \mathcal{N}^O$ of managers who have accepted outside offers, qualities $\{v_i\}$ of the incumbent managers and level L of the organization capital, we denote manager i 's strategy by

$$d^i(o_i, t, \mathcal{N}^O, \mathcal{N}^A, \{v_i\}, L).$$

We let $(\{d^i\}, \mathcal{I}^R)$ denote a profile of Markov strategies that are candidates for an MPE. Given that the firm's language includes tasks that are described by set \mathcal{L}_t in period t , current and future firm payoffs are a function of the exogenous variables

$$\{\tilde{z}_s\}_{s=t}^{s=\infty} := \left(\{\tilde{v}_{is}\}, \{\tilde{T}_{is}\}, \{\tilde{o}_{is}\}, \{\tilde{v}'_{is}\} \{\tilde{\epsilon}_{is}\}, \{\xi_{is}\} \right)_{s=t}^{s=\infty}$$

and the strategies $\{d^i\}$ and \mathcal{I}^R .

We presume that the exogenous variables $(\{\tilde{v}_{is}\}, \{\tilde{o}_{is}\}, \{\tilde{v}'_{is}\} \{\tilde{\epsilon}_{is}\}, \{\xi_{is}\})$ are all mutually independent and are independent of the $\{\tilde{T}_{is}\}$. Since the per-period firm payoff π is bounded

by $D := (1 - \theta)N(X + V)$, the state space of languages is finite and the exogenous variables are independent over time, the firm's total expected payoff may be written as a value function $w(L_t, \{d^i\}, \mathcal{I}^R)$ where $w(\cdot, \{d^i\}, \mathcal{I}^R) : \{0, 1, \dots, K\} \rightarrow [0, \frac{D}{1-\delta}]$.

We let a value function $w \in [0, \frac{D}{1-\delta}]^{K+1}$ be given (viewing w as a point in \mathfrak{R}^{K+1} rather than as a function). Any point in $[0, \frac{D}{1-\delta}]^{K+1}$ is a feasible value function. The firm's optimal strategy $\mathcal{I}^R(w)$ is described in the text.

Suppose there is some m such that $T_m = N$, *i.e.*, manager m knows that he is the last manager to receive an outside offer. Given a set of tasks \mathcal{L} covered by firm's language, incumbent qualities $\{v_i\}$ and a set \mathcal{I} of incumbents who have chosen to remain with the firm (including manager m), manager m calculates his expected payoff from remaining in the firm as a senior (recall that managers work only two periods) at

$$\begin{aligned} \gamma^m(\mathcal{I}) := \\ \theta \int \Upsilon^m \left(\mathcal{I}^R(\mathcal{I}, \{v'_i\}, \{v_i\}, L) \left[\frac{L}{K} g \left(I^R(\mathcal{I}, \{v'_i\}, \{v_i\}, L) \right) + \left\{ 1 - \frac{L}{K} \right\} g(1) + v_m \right] \Pi_i df_{v'}(v'_i) \right). \end{aligned} \quad (6)$$

If manager m is one of the retained incumbents, *i.e.*, $\Upsilon^m \left(\mathcal{I}^R(\mathcal{I}, \{v'_i\}, \{v_i\}, L) \right) = 1$, with probability $\frac{L}{K}$ he will produce by communicating with other retained managers an enhanced level of production to obtain a payoff of $\theta g \left(I^R(\mathcal{I}, \{v'_i\}, \{v_i\}, L) \right)$ in addition to θv_m but with probability $(1 - \frac{L}{K})$ the task will not be covered by the language and the enhanced production will only generate an additional payoff of $\theta g(1)$ for the manager. The manager does not observe $\{v'_i\}$ at the time he makes his quit/remain decision but knows the joint distribution of $\{v'_i\}$ and calculates his expected payoff from deciding to stay. If this value, given in (6), weakly exceeds his offer o then the manager remains with the firm and otherwise he leaves. (The offer o will equal the value in (6) with probability zero.) This defines the strategy of manager m when $T_m = N$, which we denote by $d_N^i(w)$.

We now proceed inductively. Given the strategies of the l th and all later managers to receive their offers, we calculate the strategy of the $(l - 1)$ th manager to receive his offer. We assume that $T_{l-1} = j$ and for simplicity we relabel manager j as manager $l - 1$. We let \mathcal{I}' denote the set of managers who have already elected to remain with the firm. The set of incumbents who will choose to remain with the firm depends only on \mathcal{I}' and $\{T_i, o_i\}_{i \in E^{l-1}}$, where $E^{l-1} := \{i : T_i > l - 1\} \cup \{i : T_i = \emptyset\}$. For a given \mathcal{I}' and realizations $\{T_i, o_i\}_{i \in E^{l-1}}$, we

denote the final set of incumbents who choose to remain with the firm by $\mathcal{I}(\mathcal{I}', \{T_i, o_i\}_{i \in E^{l-1}})$. Given a set \mathcal{I}' of incumbents who have elected to remain with the firm, manager $(l-1)$'s expected payoff to remaining with the firm is given by

$$c^{l-1}(\mathcal{I}') = \sum_{\{T_i: i \in E^{l-1}\}} e(\{T_i : i \in E^{l-1}\} | T_i > l-1) \int \gamma^{l-1}(\mathcal{I}(\mathcal{I}', \{T_i, o_i\}_{i \in E^{l-1}})) \prod_{i \in E^{l-1}} f_o(o_i) d(\{o_i\}_{i \in E^{l-1}}). \quad (7)$$

Manager $l-1$ remains with the firm if and only if his outside offer o_{l-1} is below the expected payoff given in (7). This describes the strategy of manager j when $T_{l-1} = j$ which we label $d_{\Omega, l-1}^j(w)$. We have now detailed the strategies of each manager for all outcomes of the $\{\tilde{T}_i\}$.

For any proposed value function w , the strategies $(\{d^i\}(w), \mathcal{I}^{\mathcal{R}}(w))$ generate a distribution over the set of incumbents who choose to remain with the firm. We define a mapping $\Omega : [0, \frac{D}{1-\delta}]^{K+1} \rightarrow [0, \frac{D}{1-\delta}]^{K+1}$ by

$$\Omega_l(w) = E_{(\{d^i\}(w), \mathcal{I}^{\mathcal{R}}(w))} [\psi(\tilde{\mathcal{I}}, l, w)] \quad (8)$$

for $l \in \{0, 1, \dots, K\}$.

We claim that the mapping Ω defined by (8) is a continuous mapping. We let $s \in [0, \frac{D}{1-\delta}]^{K+1}$ be given and consider a sequence $\{s^n\}$, $s^t \in [0, \frac{D}{1-\delta}]^{K+1} \forall t$, such that $s^n \rightarrow s$. It is clear that the maximization problem (4) governing the firm's strategy $\mathcal{I}^{\mathcal{R}}(s)$ has a unique solution for almost every $\{v'_i\}$, since the $\{v_i, v'_i\}$ are continuous random variables. The problem (4) is continuous in the elements of s . Therefore, for every $\{v'_i\}$ for which (4) has a unique solution, $\mathcal{I}^{\mathcal{R}}(s^n) \rightarrow \mathcal{I}^{\mathcal{R}}(s)$ pointwise. That is, $\mathcal{I}^{\mathcal{R}}(s^n) \rightarrow \mathcal{I}^{\mathcal{R}}(s)$ almost everywhere. This implies that

$$\begin{aligned} & \Upsilon^i \left(\mathcal{I}^{\mathcal{R}}(s^n)(\mathcal{I}, \{v'_i\}, \{v_i\}, \mathcal{L}) \right) \left[g(I^{\mathcal{R}}) \frac{L}{K} + v_i \right] \\ & \rightarrow \Upsilon^i \left(\mathcal{I}^{\mathcal{R}}(s)(\mathcal{I}, \{v'_i\}, \{v_i\}, \mathcal{L}) \right) \left[g(I^{\mathcal{R}}) \frac{L}{K} + v_i \right] \end{aligned}$$

pointwise for almost every $\{v'_i\}$. Both Υ^i and g are bounded, so Lebesgue's dominated convergence theorem (Billingsley, 1995, p.209) shows that

$\gamma_{\mathcal{I}^{\mathcal{R}}(L^n)}(\mathcal{I}) \rightarrow \gamma_{\mathcal{I}^{\mathcal{R}}(L)}(\mathcal{I})$. This in turn implies that $d_{\Omega, N}^i(s^n) \rightarrow d_{\Omega, N}^i(s)$ for almost every offer,

since the manager is indifferent between remaining and leaving with probability zero. We argue inductively that given that $d_{\Omega,j}^i(s^n) \rightarrow d_{\Omega,j}^i(s)$ for almost every o_i for all $j \geq l$, we have

$$\mathcal{I}(\mathcal{I}', \{T_i, o_i\}_{i \in E^{l-1}})(s^n) \rightarrow \mathcal{I}(\mathcal{I}', \{T_i, o_i\}_{i \in E^{l-1}})(s)$$

for almost every $\{o_i\}_{i \in E^{l-1}}$. Hence for all \mathcal{I}' , $c_s^{l-1}(\mathcal{I}') \rightarrow c_s^{l-1}(\mathcal{I}')$ and therefore $d_{\Omega,l-1}^i(s^n) \rightarrow d_{\Omega,l-1}^i(s)$ for almost every offer. We conclude that $\forall l \in \{0, 1, \dots, K\}$

$$\begin{aligned} \Omega_l(s^n) &= E_{(\{d^i\}(s^n), \mathcal{I}^{\mathcal{R}}(s^n))} [\psi(\tilde{\mathcal{I}}, l, s^n)] \\ &\rightarrow E_{(\{d^i\}(s), \mathcal{I}^{\mathcal{R}}(s))} [\psi(\tilde{\mathcal{I}}, l, s^n)] \rightarrow E_{(\{d^i\}(s), \mathcal{I}^{\mathcal{R}}(s))} [\psi(\tilde{\mathcal{I}}, l, s)] = \Omega_l(s). \end{aligned}$$

The first convergence follows from the fact that for all \mathcal{I}' , $d_{\Omega,j}^i(s^n) \rightarrow d_{\Omega,j}^i(s)$ for almost every o_i and hence for almost every $\{o_i\}$

$$\mathcal{I}(\mathcal{I}', \{T_i, o_i\}_i)(s^n) \rightarrow \mathcal{I}(\mathcal{I}', \{T_i, o_i\}_i)(s).$$

The second convergence follows from the continuity of (4) in the future payoffs. Since the per-period firm payoff π is bounded by D , it is clear that Ω maps from $[0, \frac{D}{1-\delta}]$ into itself.

The space $[0, \frac{D}{1-\delta}]^{K+1}$ is closed, bounded, convex subset of \mathfrak{R}^{K+1} . The Brouwer Fixed-Point Theorem (Stokey and Lucas, 1989, p.517) shows that Ω has a fixed point $w_0 \in [0, D]^{K+1}$. We claim that the associated profile of strategies

$$(\{d^i\}(w_0), \mathcal{I}^{\mathcal{R}}(w_0))$$

form an MPE.

The managers' strategies by construction are best responses. We fix the $(\{d^i\}(w_0))$ and verify that $\mathcal{I}^{\mathcal{R}}(w_0)$ is a best response for the firm. We set

$\tilde{z}'_s = (\{\tilde{v}_{is}\}, \{\tilde{T}_{is}\}, \{\tilde{o}_{is}\}, \{\tilde{v}'_{is}\})$ and denote its associated measure by Q . For fixed manager strategies the firm's optimization problem may be viewed as a single-agent dynamic programming problem. The function w_0 has been selected to solve the problem

$$w_0(\mathcal{L}) = E_{z'} \left[\sup_{\mathcal{I}^{\mathcal{R}} \subseteq \mathcal{I}(\mathcal{L}, z')} \left[\pi(\{y_i\}, \mathcal{N}^{\mathcal{L}}, \mathcal{L}) + \delta \int w_0(\phi(\mathcal{L}, \mathcal{N}^{\mathcal{L}}, \epsilon, \xi)) dF_\epsilon(\epsilon) dF_\xi(\xi) \right] \right] \quad (9)$$

where $\mathcal{I}(\mathcal{L}, z')$ is the set of remaining incumbents, $y_i = v_i$ if $i \in \mathcal{I}^{\mathcal{R}}$ and $y_i = v'_i$ if $i \notin \mathcal{I}^{\mathcal{R}}$ and

$$\phi(\mathcal{L}, \mathcal{I}^{\mathcal{R}}, \epsilon, \xi) = \begin{cases} \max \left[\operatorname{argmax}_{0 \leq L^1 \leq \#(\mathcal{L} \cup \epsilon)} w(L^1) \right] & \text{if } p(I^{\mathcal{R}}) > \xi \\ \max \left[\operatorname{argmax}_{0 \leq L^2 \leq 1} w(L^2) \right] & \text{otherwise.} \end{cases}$$

We define $\tilde{q}_s = (\{\tilde{\epsilon}_{i,s-1}\}, \{\tilde{\xi}_{i,s-1}\}, \tilde{z}'_s)$ and

$$v(\mathcal{L}, q_s) = \sup_{\mathcal{I}^{\mathcal{R}} \subseteq \mathcal{I}(\mathcal{L}, z'_s)} \left[\pi(\{y_i\}, \mathcal{I}^{\mathcal{R}}, \mathcal{L}) + \delta \int w_0(\phi(\mathcal{L}, \mathcal{I}^{\mathcal{R}}, \epsilon_s, \xi_s)) dF_\epsilon(\epsilon_s) dF_\xi(\xi_s) \right] \quad (10)$$

The firm's strategy $\mathcal{I}^{\mathcal{R}}(w_0)$ by construction corresponds to the optimal policy solving the problem (10). We have $w_0(\mathcal{L}) = E_{z'_{s+1}} [v(\mathcal{L}, q_{s+1})]$ which implies that

$$\begin{aligned} v(\mathcal{L}, q_s) &= \sup_{\mathcal{I}^{\mathcal{R}} \subseteq \mathcal{I}(\mathcal{L}, z'_s)} \\ &\left[\pi(\{y_i\}, \mathcal{I}^{\mathcal{R}}, \mathcal{L}) + \delta \int v(\phi(\mathcal{L}, \mathcal{I}^{\mathcal{R}}, \epsilon_s, \xi_s), q_{s+1}) dF_{z'_{s+1}}(z'_{s+1}) dF_\epsilon(\epsilon_s) dF_\xi(\xi_s) \right] \\ &= \sup_{\mathcal{I}^{\mathcal{R}} \subseteq \mathcal{I}(\mathcal{L}, z'_s)} \left[\pi(\{y_i\}, \mathcal{I}^{\mathcal{R}}, \mathcal{L}) + \delta \int v(\phi(\mathcal{L}, \mathcal{I}^{\mathcal{R}}, \epsilon_s, \xi_s), q_{s+1}) dF_{q_{s+1}}(q_{s+1}) \right] \end{aligned}$$

Since the firm's profit is bounded from below at zero and the function w_0 is bounded, the Principle of Optimality (Stokey and Lucas, 1989, p.256-258) shows that w_0 is the firm's value function and that $\mathcal{I}^{\mathcal{R}}(w_0)$ describes the firm's optimal choice in the dynamic programming problem. This verifies that $\mathcal{I}^{\mathcal{R}}(w_0)$ is a best response.

We can show that the value function w associated with any candidate MPE profile $(\{d^i\}, \mathcal{I}^{\mathcal{R}})$ satisfies condition (9). Since the firm's payoffs are non-negative and bounded, the sum of the firm's discounted future payoffs may be calculated in each state. We define this sum $w_{\mathcal{I}^{\mathcal{R}}}$ to be the firm's value function. Since strategies are Markov, the value function is a function of only the current state. The strategy $\mathcal{I}^{\mathcal{R}}$ is required to be a measurable function of the state and must maximize the firm's expected payoff at each state given the strategies of the managers. In the language of Stokey and Lucas (1989, p.254), $\mathcal{I}^{\mathcal{R}}$ is a global plan, so the associated value function is measurable. Minor modifications to the argument given on Stokey and Lucas (1989, p.252-253), show that w must satisfy equation (9) and that the Markov strategy $\mathcal{I}^{\mathcal{R}}$ must coincide almost everywhere with an optimal policy associated with (9).

Proof of Lemma 2:

We define a *history* of received offers $\tilde{h}^{1,j} := \{\{\tilde{T}_i : \tilde{T}_i = l\}\}_{l=1}^j$. For a given realization $h^{1,j}$ and $h^{2,j} \in \{0, 1\}^j$, we define a *history* $h^j := (h^{1,j}, h^{2,j})$. The term $h^{2,j}$ represents a record of the remain/quit decisions made by the first j managers to receive offers. If $h^{1,j}(i) = \emptyset$ then we fix $h^{2,j} = 1$. We define $(\tilde{E}^j | h^{1,j}) = \{i : (\tilde{T}_i | h^{1,j}) > j\} \cup \{i : (\tilde{T}_i | h^{1,j}) = \emptyset\}$ to be the

set of managers who have not received an offer by the time the j th offer is made, and we set

$$(\tilde{z}_+^j | h^{1,j}) := (\{(\tilde{T}_i | h^{1,j})\}_{i \in (\tilde{E}^j | h^{1,j})}, \{\tilde{o}_i\}_{i \in (\tilde{E}^j | h^{1,j})}, \{\tilde{a}_i\}, \{\tilde{\epsilon}_i\}, \{\tilde{\xi}_i\})$$

to be the exogenous variables not yet realized by the time the j th manager makes his decision (e.g. whether the other managers will receive outside offers and what those offers will be). We denote by $\mathcal{I}(w, L, (z_+^j | h^j))$ the set of incumbents who elect to remain in the firm given a history h^j , a realization z_+^j of the exogenous variables and the strategies of the managers.

Formally, Lemma 2 states the following:

Let two histories h_a^j and h_b^j be given such that $h_a^{1,j} = h_b^{1,j}$ and $\forall i \in \{1, \dots, j\} h_a^{2,j}(i) \leq h_b^{2,j}(i)$. For any realization z_+^j of $(\tilde{z}_+^j | h_a^{1,j})$

$$a) \mathcal{I}(w, L, (z_+^j | h_a^j)) \subseteq \mathcal{I}(w, L, (z_+^j | h_b^j)).$$

$$b) \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L, (z_+^j | h_a^j))) \subseteq \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L, (z_+^j | h_b^j))).$$

Reorder the managers such that $(v_{i+1} - v'_{i+1}) \geq (v_i - v'_i)$ for all $i \in \{1, \dots, I\}$. Base case: $j = N$. In this case, $\mathcal{I}(w, L, (z_+^j | h_a^j)) \subseteq \mathcal{I}(w, L, (z_+^j | h_b^j))$ by the definition of a history. Consider manager $i \in \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L, (z_+^j | h_a^j)))$. Suppose that manager i has the a th highest index in $\mathcal{I}(w, L, (z_+^j | h_a^j))$. That he was not fired by the firm implies that

$$\begin{aligned} & [p(a) - p(a-1)] \delta E \left[\max_{\mathcal{L}^1 \subseteq \mathcal{L}_t \cup \{\tilde{k}_t\}} w(L^1) - \max_{\mathcal{L}^2 \subseteq \{\tilde{k}_t\}} w(L^2) \right] \\ & + (1 - \theta) \frac{L_t}{K} [G(a) - G(a-1)] \geq v'_i - v_i. \end{aligned} \quad (11)$$

Suppose that manager i has the b th highest index in $\mathcal{I}(w, L, (z_+^j | h_b^j))$. It must be that $b \geq a$. It is clear that

$$I^{\mathcal{R}}(\mathcal{I}(w, L, (z_+^j | h_a^j))) \subseteq I^{\mathcal{R}}(\mathcal{I}(w, L, (z_+^j | h_b^j))),$$

since retaining a given number of incumbents is always more appealing to the firm when it draws from a larger pool. The convexity of p and G and inequality (11) then show that $i \in \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L, (z_+^j | h_b^j)))$.

Induction: Assume the statement holds for $j = t$. Let two histories h_a^{t-1} and h_b^{t-1} be given such that $h_a^{1,t-1} = h_b^{1,t-1}$ and $\forall i \in \{1, \dots, t-1\} h_a^{2,t-1}(i) \leq h_b^{2,t-1}(i)$. Let any realization z_+^{t-1}

of $(\tilde{z}_+^{t-1}|h_a^{1,t-1})$ be given. Under this realization consider i such that $T_i = t$ (we will later consider the case in which there is no such i).

We let $\hat{\mathcal{N}}^{\mathcal{O}} = \{i : T_i \leq t-1\}$, $\hat{\mathcal{N}}^{\mathcal{A}_a} = \{i : \exists j \leq t-1, T_i = j, h_a^{2,t-1}(j) = 0\}$ and $\hat{\mathcal{N}}^{\mathcal{A}_b} = \{i : \exists j \leq t-1, T_i = j, h_b^{2,t-1}(j) = 0\}$. If

$$d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}_a}, \{v_i\}, L) \leq d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}_b}, \{v_i\}, L)$$

or if no such i exists, then the induction assumptions apply and give the result. Moreover, it cannot be that

$$d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}_a}, \{v_i\}, L) = 1 > 0 = d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}_b}, \{v_i\}, L)$$

since the payoff to manager i to quitting is the same under both histories (he receives his outside offer), while the expected payoff to remaining is higher under h_b^{t-1} , by the induction assumption (part a) and (6).

Proof of Lemma 3:

Formally, Lemma 3 states the following:

Let $L^a \leq L^b$ be given. For any history h^j and any realization z_+^j of $(\tilde{z}_+^j|h^{1,j})$

$$a) \mathcal{I}(w, L^a, (z_+^j|h^j)) \subseteq \mathcal{I}(w, L^b, (z_+^j|h^j)).$$

$$b) \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L^a, (z_+^j|h^j))) \subseteq \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L^b, (z_+^j|h^j))).$$

Base case: $j = N$. We have $\mathcal{I}(w, L^a, (z_+^j|h^j)) = \mathcal{I}(w, L^b, (z_+^j|h^j))$. In this case the result follows directly from (4). Under L^b the firm has a greater incentive to retain managers both for current production and to achieve a higher future language level.

Induction: Assume the statement holds for $j = t$. Let a history (h^{t-1}) and any realization z_+^{t-1} of $(\tilde{z}_+^{t-1}|h^{1,t-1})$ be given. Under this realization consider i such that $T_i = t$ (we will later consider the case in which there is no such i). We let $\hat{\mathcal{N}}^{\mathcal{O}} = \{i : T_i \leq t-1\}$ and $\hat{\mathcal{N}}^{\mathcal{A}} = \{i : \exists j \leq t-1, T_i = j, h^{2,t-1}(j) = 0\}$. If

$$d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}}, \{v_i\}, L^a) = d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}}, \{v_i\}, L^b)$$

or if no such i exists, then the induction assumptions apply and give the result. If

$$d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}}, \{v_i\}, L^a) \leq d^i(o_i, t, \hat{\mathcal{N}}^{\mathcal{O}}, \hat{\mathcal{N}}^{\mathcal{A}}, \{v_i\}, L^b)$$

then Lemma 2 and the induction assumption show that the result holds. It cannot be that

$$d^i(o_i, t, \hat{\mathcal{N}}^O, \hat{\mathcal{N}}^A, \{v_i\}, L^a) = 1 > 0 = d^i(o_i, t, \hat{\mathcal{N}}^O, \hat{\mathcal{N}}^A, \{v_i\}, L^b)$$

since the payoff to manager i to quitting is the same, while the expected payoff to remaining is higher under L^b , by the induction assumption (part a) and (6).

Proof of Result 2:

We let $w \in [0, \frac{D}{1-\delta}]^{K+1}$ be given. The firm's maximization problem (4) shows that

$$L^a \leq L^b \Rightarrow \psi(\mathcal{I}, L^a, w) \leq \psi(\mathcal{I}, L^b, w).$$

We denote the solution to (4) by $m(\mathcal{I}, L, w)$. We have $m(\mathcal{I}, L^b, w) \geq m(\mathcal{I}, L^a, w)$, implying that any manager retained when $L = L^b$ is also retained when $L = L^a$. For a given value function w , language richness L and outcome z of the exogenous variables, we define $\mathcal{I}(w, L, z)$ to be the set of incumbents who elect to remain with the firm. Lemma 3 (part b) shows that if $L^a \leq L^b$, then $\mathcal{I}(w, L^a, z) \subseteq \mathcal{I}(w, L^b, z)$. Since $\mathcal{I}_1 \subseteq \mathcal{I}_2 \Rightarrow \psi(\mathcal{I}_1, L, w) \leq \psi(\mathcal{I}_2, L, w)$, this shows that $\Omega_{L^a}(w) \leq \Omega_{L^b}(w)$. So $\Omega([0, \frac{D}{1-\delta}]^{K+1}) \subseteq \mathcal{U}$ where

$$\mathcal{U} = \left\{ \{u_i\}_{i=0}^K \in [0, \frac{D}{1-\delta}]^{K+1} : u_j \leq u_{j+1} \forall j \in \{0, \dots, K\} \right\}$$

which demonstrates that in any possible equilibrium firm value is increasing in the richness of language.

Proof of Result 3:

Formally, Result 3 states the following:

If $L_2 \geq L_1$ then for every realization z of the exogenous variables

$$a) \mathcal{I}(w, L_1, z) \subseteq \mathcal{I}(w, L_2, z)$$

$$b) \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L_1, z), z, L_1) \subseteq \mathcal{I}^{\mathcal{R}}(\mathcal{I}(w, L_2, z), z, L_2).$$

That is, each manager is less likely to quit and more likely to work for the firm this period under language \mathcal{L}_2 than under language \mathcal{L}_1 .

Lemma 3 proves Result 3.

Proof of Result 4:

Formally Result 4 states the following:

Consider Firms A and B with identical managers and histories h_A^j and h_B^j , respectively. If the firms have a common history of received offers, $h_A^{1,j} = h_B^{1,j}$, but Firm A has experienced more quits than Firm B this period, $\forall i \in \{1, \dots, j\} h_A^{2,j}(i) \leq h_B^{2,j}(i)$, then for any realization z_+^j of the exogenous variables ($\tilde{z}_+^j | h_A^j$), if manager m remains in firm A, $m \in \mathcal{I}(w, L, (z_+^j | h_A^j))$, then he also remains in firm B, $m \in \mathcal{I}(w, L, (z_+^j | h_B^j))$.

Result 4 follows directly from Lemma 2.

Proof of Result 5:

Given a set \mathcal{J}^R of retained incumbents and language \mathcal{L} , we define the realized wage premium λ_i of manager i by $\lambda_i(y_i, \mathcal{J}^R, \mathcal{L}) = \pi_i(y_i, \mathcal{J}^R, \mathcal{L}) - \theta(y_i + g(1))$, where $y_i = v_i$ for $i \in \mathcal{J}^R$ and $y_i = v'_i$ for $i \notin \mathcal{J}^R$.

Formally, Result 5 states that given incumbent and replacement qualities $\{v_i\}$ and $\{v'_i\}$ respectively, a set \mathcal{J}^R of retained incumbents and a firm language that covers a set of tasks \mathcal{L} :

a) $\pi_i(y_i, \mathcal{J}^R, \mathcal{L})$ is increasing in L and J^R and $\pi_i(b, \mathcal{J}^R, \mathcal{L}) \geq \pi_j(b, \mathcal{J}^R, \mathcal{L})$ if $i \in \mathcal{J}^R$ and $j \notin \mathcal{J}^R$.

b) If $L^a \leq L^b$, then $\frac{\sum_i E_z[\lambda_i(\tilde{I}_i, \tilde{\mathcal{J}}^R, \mathcal{L}^a)]}{N} \leq \frac{\sum_i E_z[\lambda_i(\tilde{I}_i, \tilde{\mathcal{J}}^R, \mathcal{L}^b)]}{N}$.

a) Follows directly from (2). b) Given a set $\mathcal{N}^{\mathcal{L}}$ of retained incumbents and a set of tasks covered by language \mathcal{L} , the realized average wage premium is

$$\frac{\sum_i \lambda_i(y_i, \mathcal{I}, \mathcal{L})}{N} = \frac{\theta \left(\frac{L}{K} \right) I[g(I) - g(1)]}{N}.$$

Lemma 3 shows that for any outcome z of the exogenous variables, $\mathcal{I}(w, L^a, z) \subseteq \mathcal{I}(w, L^b, z)$.

Proof of Result 6:

Given that $L_t = L$, the distribution over future language states is given by

$$L_{t+1} = \begin{cases} L + 1 & \text{with probability } \left(\frac{K-L}{K} \right) \int p(I_R(I(w, L, z))) dQ \\ L & \text{with probability } \left(\frac{L}{K} \right) \int p(I_R(I(w, L, z))) dQ \\ 1 & \text{with probability } (1 - \int p(I_R(I(w, L, z))) dQ) \end{cases}$$

By Lemma 3, $\int p(I_R(I(w, L, z))) dQ$ is increasing in L , so the conditional distribution of L_{t+1} is increasing in L in the sense of first-order stochastic dominance (FOSD). The transition from the firm's language state this period to its state next period is governed by a Markov transition matrix $\pi = \{\pi_{ij}\}_{i,j=0}^{i,j=K}$. The work above shows that π has the FOSD property:

$\sum_{r=0}^j \pi_{ir}$ is weakly decreasing in i for all $j \leq K$. We denote the language state at period u by c_u , a $1 \times (K + 1)$ row vector. If the state in period u is L , then the $(L + 1)$ st element of c_u is equal to one and all other elements are zero. For all $s > 0$ we have

$$c_{t+s} = c_t \pi^s. \quad (12)$$

We suppose that a and b are two $(K + 1) \times (K + 1)$ matrices each with the FOSD property. We will show that ab has the FOSD property. We let $i_1, i_2, j \in \{0, \dots, K\}$ be given such that $i_1 \leq i_2$. We have $\sum_{r=0}^j (ab)_{i_1 r} = \sum_{r=0}^K a_{i_1 r} \left(\sum_{s=0}^j b_{rs} \right)$. We may view $\{a_{i_1 r}\}_{r=0}^K$ and $\{a_{i_2 r}\}_{r=0}^K$ as distributions over the $(K + 1)$ states, and by the FOSD property, $\{a_{i_2 r}\}_{r=0}^K$ FOSD dominates $\{a_{i_1 r}\}_{r=0}^K$. Since b satisfies the FOSD property, $\left(\sum_{s=0}^j b_{rs} \right)$ may be viewed as a decreasing function of r . We have

$$\sum_{r=0}^j (ab)_{i_1 r} = \sum_{r=0}^K a_{i_1 r} \left(\sum_{s=0}^j b_{rs} \right) \geq \sum_{r=0}^K a_{i_2 r} \left(\sum_{s=0}^j b_{rs} \right) = \sum_{r=0}^j (ab)_{i_2 r},$$

which shows that ab has the FOSD property. This shows that π^{t-s} has the FOSD property, so the conditional distribution of L_{t+s} is increasing in L in the sense of FOSD. This shows part a). Results 2 and 3 show that firm value and retentions are increasing in the language state which shows part b).

Given $I_R(L_t)$, the distribution over future language states is given by

$$L_{t+1} = \begin{cases} L + 1 & \text{with probability } \left(\frac{K-L}{K} \right) p(I_R(L_t)) \\ L & \text{with probability } \left(\frac{L}{K} \right) p(I_R(L_t)) \\ 1 & \text{with probability } (1 - p(I_R(L_t))) \end{cases}$$

The conditional distribution of L_{t+1} is increasing in $I_R(L_t)$ in the sense of FOSD. Given two $1 \times (K + 1)$ row vectors c^1 and c^2 such that c^1 dominates c^2 in the sense of FOSD and given that $(K + 1) \times (K + 1)$ matrix a has the FOSD property, we will show that $c^1 a$ FOSD dominates $c^2 a$. We have that for any j , $\sum_{r=0}^j a_{rs}$ is decreasing, so

$$\sum_{r=0}^j (c^1 a)_r = \sum_{r=0}^k c_r^1 \sum_{s=0}^j a_{rs} \leq \sum_{r=0}^k c_r^2 \sum_{s=0}^j a_{rs} = \sum_{r=0}^j (c^2 a)_r.$$

The result follows from noting that $c_{t+s} = c_{t+1} \pi^{s-1}$.

Proof of Result 7:

Part a) follows from noting that $\left(\frac{L}{K}\right) \int p(I_R(I(w, L, z)))dQ$ is increasing in L . Part b) follows from the argument given in the proof of Result 6. For part c),

$$\begin{aligned} P(L_{t+1} = 1|L_t = 1) &= 1 - \left(\frac{K-1}{K}\right) \int p(I_R(I(w, 1, z)))dQ \\ &\geq \left(\frac{2}{K}\right) \int p(I_R(I(w, 2, z)))dQ = P(L_{t+1} = 2|L_t = 2), \end{aligned}$$

where the inequality follows from $p \leq \frac{K}{K+1}$.

Proof of Result 8:

For this result we will consider the firm's current profits and future value separately. Define

$$\begin{aligned} a_c(\mathcal{I}, w, \mathcal{L}_t, m) &\equiv (1 - \theta) \left[\frac{L_t}{K} m [g(m) - g(1)] + Ng(1) + \sum_{n=1}^{I-m} v'_n + \sum_{n=I-m+1}^I v_n \right], \\ a(\mathcal{I}, w, \mathcal{L}_t, m) &\equiv \delta E \left[p(m)w(\#(\mathcal{L}_t \cup \{\tilde{k}_t\})) + \{1 - p(m)\}w(\#(\{\tilde{k}_t\})) \right]. \end{aligned}$$

Formally, Result 8 states the following:

Suppose m^* maximizes firm value (4). If m' is such that $a_c(\mathcal{I}, w, \mathcal{L}_t, m') > a_c(\mathcal{I}, w, \mathcal{L}_t, m^*)$ then $m^* > m'$.

We will show:

- i) If $\hat{m} \geq m^*$ then $a(\mathcal{I}, w, \mathcal{L}_t, \hat{m}) \geq a(\mathcal{I}, w, \mathcal{L}_t, m^*)$.
- ii) Suppose m' is such that $a_c(\mathcal{I}, w, \mathcal{L}_t, m') > a_c(\mathcal{I}, w, \mathcal{L}_t, m^*)$. Then $m^* > m'$.

Part i) follows from the fact that w and p are increasing. For Part ii), suppose $m' \geq m^*$. By part i), this implies that $a(\mathcal{I}, w, \mathcal{L}_t, m') \geq a(\mathcal{I}, w, \mathcal{L}_t, m^*)$. If $a_c(\mathcal{I}, w, \mathcal{L}_t, m') > a_c(\mathcal{I}, w, \mathcal{L}_t, m^*)$, this would show that m^* does not maximize (4), which is a contradiction. Result 8 follows from Part ii).

Proof of Result 9: For a fixed realization z of the exogenous variables, and given the strategies of the firm and the managers and the richness L of the firm's discourse \mathcal{L} , for all i such that $T_i \neq \emptyset$, we can write manager i 's quit/remain decision as $d^i(z, L)$. We define the total payoff τ_i of manager i who began the period as an incumbent by

$$\tau_i(w, L, z) = \begin{cases} o_i & \text{if } T_i \neq \emptyset \text{ and } d^i(z, L) = 0 \\ \gamma^i(\mathcal{I}(w, L, z)) & \text{otherwise} \end{cases}$$

Formally, Result 9 states:

If $L^1 \leq L^2$, then $\delta E[\tau_i(w, L_{t+1}, \tilde{z})|L_t = L^a] \leq \delta E[\tau_i(w, L_{t+1}, \tilde{z})|L_t = L^b]$.

First we show that if $L^a \leq L^b$, then

$$E[\tau_i(w, L_{t+1}, \tilde{z})|L_{t+1} = L^a] \leq E[\tau_i(w, L_{t+1}, \tilde{z})|L_{t+1} = L^b]. \quad (13)$$

We set $\tilde{\mathcal{E}}^j := \{i : \tilde{T}_i \leq j\}$ and define $\tilde{z}^j := (\{\tilde{v}_i\}, \{\tilde{T}_i\}_{i \in \tilde{\mathcal{E}}^j}, \{\tilde{o}_i\}_{i \in \tilde{\mathcal{E}}^j})$ to be the exogenous variables observed by the j th manager to receive an outside offer. For a given realization \tilde{z}^j of these exogenous variables, for $s \in \{a, b\}$, we further define a history $h_s(\tilde{z}^j) = (h_s^{1,j}(\tilde{z}^j), h_s^{2,j}(\tilde{z}^j))$ recursively in the following manner:

$$h_s^{1,j}(\tilde{z}^j) = \{\{T_i : T_i = l\}\}_{l=1}^j,$$

$$h_s^{2,j}(\tilde{z}^j) = d^m(o_m, 1, \emptyset, \emptyset, \{v_i\}, L^s)$$

for m such that $T_m = 1$. For $l \geq 1$,

$$h_s^{2,j}(\tilde{z}^{l+1}) = (h(\tilde{z}^l), d^m(o_m, l+1, \hat{\mathcal{N}}^O, \hat{\mathcal{N}}^A, \{v_i\}, L^s))$$

for m such that $T_m = l+1$, $\hat{\mathcal{N}}^O = \{i : T_i \leq l\}$ and $\hat{\mathcal{N}}^A = \{i : \exists j \leq l, T_i = j, h_s^{2,j}(\tilde{z}^l)(j) = 0\}$.

We note that Lemmas 2 and 3 show that $h_a^{2,j}(\tilde{z}^j) \leq h_b^{2,j}(\tilde{z}^j)$ for all \tilde{z}^j .

We let a realization z of the exogenous variables be given. If $T_i = \emptyset$, then Lemma 3 shows that $\pi^i(w, L^a, z) \leq \pi^i(w, L^b, z)$, so $\tau_i(w, L^a, z) \leq \tau_i(w, L^b, z)$. If $T_i = m$ for some $m \in \{1, \dots, N\}$, we have for all realizations z_+^m of $(z_+^m | z_-^m)$ $\mathcal{I}^R(\mathcal{I}(w, \mathcal{L}^a, (z_+^m | h_a^m(z_-^m)))) \subseteq \mathcal{I}^R(\mathcal{I}(w, \mathcal{L}^b, (z_+^m | h_b^m(z_-^m))))$ by Lemmas 2 and 3. Since manager i chooses the best option and quits or remains given his conditional expectation of the payoff to remaining,

$$E[\tau_i(w, L^a, \tilde{z}) | h_a^m(z_-^m)] \leq E[\tau_i(w, L^b, \tilde{z}) | h_b^m(z_-^m)].$$

Inequality (13) follows from the law of iterated expectations. The result then follows from the fact that the conditional distribution of L_{t+1} given $L_t = L_2$ FOSD dominates the conditional distribution of L_{t+1} given $L_t = L_1$, as shown in the proof of Result 6.

Proof of Result 10:

It is clear, for part a), that paying u to replacements yields the firm no benefits.

For parts b) and c), it is necessary to show that Results 1 and 2 and Lemmas 2 and 3 hold in the modified model. The proof of Result 1 follows directly, with a simple modification in the definition of ϕ . Set $w_1 = \max \left[\operatorname{argmax}_{0 \leq L^1 \leq \#(\mathcal{L} \cup \epsilon)} w(L^1) \right]$ and $w_2 = \max \left[\operatorname{argmax}_{0 \leq L^2 \leq 1} w(L^2) \right]$. Then

$$\phi_u(\mathcal{L}, \mathcal{I}^R, \epsilon, \xi) = \begin{cases} w_1 & \text{if } p(I^R) > \xi \text{ and } p(I^R)(w_1 - w_2) - I^R u \geq p(0)(w_1 - w_2) \\ w_2 & \text{if } p(I^R) \leq \xi \text{ and } p(I^R)(w_1 - w_2) - I^R u \geq p(0)(w_1 - w_2) \\ w_1 & \text{if } p(0) > \xi \text{ and } p(I^R)(w_1 - w_2) - I^R u < p(0)(w_1 - w_2) \\ w_2 & \text{otherwise.} \end{cases}$$

The proof of Lemma 2 is analogous to that given previously since $E[\max\{p(x)(w_1 - w_2) + w_2 - xu, p(0)(w_1 - w_2) + w_2\}]$ is convex in x . This follows from the convexity of p and the fact that the maximum function is increasing and convex.

For the proof of Lemma 3, we let \mathcal{L}^a and \mathcal{L}^b such that $L^a \leq L^b$ be given. We will show that more incumbents are retained under L^b . We define $w_1^b = \max_{\mathcal{L} \subseteq \mathcal{L}^b \cup k} w(L)$, $w_1^a = \max_{\mathcal{L} \subseteq \mathcal{L}^a \cup k} w(L)$ and $w_1^k = \max_{\mathcal{L} \subseteq k} w(L)$.

It is sufficient to show that for any m

$$\begin{aligned} & E \left[\max\{p(m)(w_1^a - w_2) + w_2 - um, p(0)(w_1^a - w_2) + w_2\} \right] - \\ & E \left[\max\{p(m-1)(w_1^a - w_2) + w_2 - u(m-1), p(0)(w_1^a - w_2) + w_2\} \right] \\ & \leq E \left[\max\{p(m)(w_1^b - w_2) + w_2 - um, p(0)(w_1^a - w_2) + w_2\} \right] - \\ & E \left[\max\{p(m-1)(w_1^b - w_2) + w_2 - u(m-1), p(0)(w_1^a - w_2) + w_2\} \right]. \end{aligned}$$

For any given k , if $p(m-1)(w_1^a - w_2) + w_2 - u(m-1) \geq p(0)(w_1^a - w_2) + w_2$ then the result follows from the convexity of p and $w_1^b \geq w_1^a$. If $p(m-1)(w_1^a - w_2) + w_2 - u(m-1) < p(0)(w_1^a - w_2) + w_2$, case by case analysis shows that the result holds.

The modified Result 2 follows directly from the modified Lemmas 2 and 3. For the proof of part b) of this result, consider \mathcal{L}^a and \mathcal{L}^b and let an outcome of the exogenous variables be given. We denote the choices of m under \mathcal{L}^a and \mathcal{L}^b by m^a and m^b respectively. The modified Lemma 3 shows that $m^a \leq m^b$. Performance insensitive pay is paid under \mathcal{L}^a if and only if

$$\begin{aligned} & (p(m^a) - p(0))(w_1^a - w_2) \geq um^a \\ \Rightarrow & (p(m^b) - p(0))(w_1^a - w_2) \geq um^b \end{aligned} \tag{14}$$

$$\Rightarrow (p(m^b) - p(0)) (w_1^b - w_2) \geq um^b$$

where the first implication follows the convexity of p and the second from the fact that $w_1^b \geq w_1^a$. This shows that performance insensitive pay is paid under \mathcal{L}^b as well.

For the proof of part c) of this result, we note that the language is transmitted with probability $p(0)$ when no performance insensitive pay is paid. When performance insensitive pay is paid, the transmission probability must be at least this high.

Proof of Result 11:

The expected wages generated by bargaining over output are given by $\left(\frac{\theta}{1-\theta}\right) Pr(\mathcal{L})$. For a given value function w , language L and outcome z of the exogenous variables, we denote the number of managers receiving performance insensitive pay by $I^{PI}(w, L, z)$. We have

$$I^{PI}(w, L, z) = I^R(\mathcal{I}(w, L, z), z, L)$$

if

$$(p(I^R(\mathcal{I}(w, L, z), z, L)) - p(0)) (w(L) - w(0)) \geq I_R(\mathcal{I}(w, L, z), z, L) u$$

and $I^{PI}(w, L, z) = 0$ otherwise. Modified Lemma 3 and (14) show that the amount of performance insensitive pay is increasing in L .

Proof of Result 12: We denote the two merging firms by Firm 1 and Firm 2, with languages that encompass tasks given by sets \mathcal{L}_1 and \mathcal{L}_2 and numbers of remaining incumbents N_1 and N_2 , respectively. The value functions for the values of the constituent firms are w_1 and w_2 , and the value function for the merged firm is w .

The language of Firm 1 will be selected if and only if

$$\begin{aligned} & \hat{p}(N_1) E [\tau_m(w, \#(\mathcal{L}_1 \cup \tilde{k}), \tilde{z})] + \{1 - \hat{p}(N_1)\} E [\tau_m(w, 1, \tilde{z})] \geq \\ & \hat{p}(N_2) E [\tau_m(w, \#(\mathcal{L}_2 \cup \tilde{k}), \tilde{z})] + \{1 - \hat{p}(N_2)\} E [\tau_m(w, 1, \tilde{z})] \\ & \iff \hat{p}(N_1) (E [\tau_m(w, \#(\mathcal{L}_1 \cup \tilde{k}), \tilde{z})] - E [\tau_m(w, 1, \tilde{z})]) \geq \\ & \hat{p}(N_2) (E [\tau_m(w, \#(\mathcal{L}_2 \cup \tilde{k}), \tilde{z})] - E [\tau_m(w, 1, \tilde{z})]). \end{aligned} \tag{15}$$

We denote the selected language by $s(\mathcal{L}_1, N_1, \mathcal{L}_2, N_2) \in \{1, 2\}$. Formally, if the language s is selected the value Δ created by the merger is

$$\begin{aligned}
\Delta(\mathcal{L}_1, N_1, \mathcal{L}_2, N_2) &= \hat{p}(N_s)E \left[w(\#(\mathcal{L}_s \cup \tilde{k}_s)) \right] + \{1 - \hat{p}(N_s)\}E \left[w(\#(\tilde{k}_s)) \right] \\
&\quad - \hat{p}(N_1)E \left[w_1(\#(\mathcal{L}_1 \cup \tilde{k}_1)) \right] - \{1 - \hat{p}(N_1)\}E \left[w_1(\#(\tilde{k}_1)) \right] \\
&\quad - \hat{p}(N_2)E \left[w_2(\#(\mathcal{L}_2 \cup \tilde{k}_2)) \right] - \{1 - \hat{p}(N_2)\}E \left[w_2(\#(\tilde{k}_2)) \right]. \quad (16)
\end{aligned}$$

Formally, Result 12 states that

a) If $s(\mathcal{L}_1, N_1, \mathcal{L}_2, N_2) = 1$ then $s(\mathcal{L}, N, \mathcal{L}_2, N_2) = 1$ for all $N \geq N_1$ and \mathcal{L} such that $L \geq L_1$ (and analogously for Firm 2).

b) Suppose $N_1 \geq N_2$. If $L_1 \geq L_{2a} \geq L_{2b}$, then $s(\mathcal{L}_1, N_1, \mathcal{L}_{2a}, N_2) = 1 = s(\mathcal{L}_1, N_1, \mathcal{L}_{2b}, N_2)$ and $\Delta(\mathcal{L}_1, N_1, \mathcal{L}_{2b}, N_2) \geq \Delta(\mathcal{L}_1, N_1, \mathcal{L}_{2a}, N_2)$. If $s(\mathcal{L}_{1a}, N_1, \mathcal{L}_2, N_2) = 2$ and $L_{1a} \geq L_{1b}$ then $\Delta(\mathcal{L}_{1b}, N_1, \mathcal{L}_2, N_2) \geq \Delta(\mathcal{L}_{1a}, N_1, \mathcal{L}_2, N_2)$

a) This follows from inspection of (15) and (13). b) For the first statement, it is clear from (15) that under both mergers the language that is adopted by the merged firm covers tasks given by set \mathcal{L}_1 . The merger with Firm 2a results in the loss of a more valuable language (last two terms in (16)). For the second statement, the language adopted covers tasks \mathcal{L}_2 , and the merger with Firm 1a results in the loss of a more valuable language (third and fourth terms in (16)).

Proof of Result 13:

We denote by \mathcal{L}_m the realized language of the merged firm. We will say that a merger is a *failure* if $L_m < \min\{L_1, L_2\}$.

Formally, Result 13 states that

- a) If $N_1 \geq N_2$ then $P(L_m < \min\{L_1, L_2\} | L_2 = a)$ is increasing in a .
b) If $N_1 \geq N_2$ then $P(L_m < \min\{L_1, L_2\} | L_1 = b_1) \geq P(L_m < \min\{L_1, L_2\} | L_1 = b_2)$ for all $b_2 \geq b_1 \geq 2$.

Part a): It will always be the case that $L_m \geq 1$, so if $a = 1$ then

$$P(L_m < \min\{L_1, L_2\} | L_2 = a) = 0.$$

We now let $a_2 \geq a_1 > 1$ be given.

If $N_1 = N_2$ then

$$P(L_m < \min\{L_1, L_2\} | L_2 = a_1) = 1 - \hat{p}(N_1) = P(L_m < \min\{L_1, L_2\} | L_2 = a_2).$$

We now suppose $N_1 > N_2$. If $s = 2$ when $L_2 = a_1$ then by result 12a, $s = 2$ when $L_2 = a_2$. In this case

$$P(L_m < \min\{L_1, L_2\} | L_2 = a_1) = 1 - \hat{p}(N_2) = P(L_m < \min\{L_1, L_2\} | L_2 = a_2).$$

If $s = 1$ when $L_2 = a_1$

$$P(L_m < \min\{L_1, L_2\} | L_2 = a_1) = 1 - \hat{p}(N_1)$$

$$\leq \min\{1 - \hat{p}(N_1), 1 - \hat{p}(N_2)\} \leq P(L_m < \min\{L_1, L_2\} | L_2 = a_2).$$

Part b): As in part a), if $N_1 = N_2$, then the probability of failure does not depend on b (for $b \geq 2$).

We now suppose $N_1 > N_2$. If $s = 1$ when $L_1 = b_1$ then by result 12a, $s = 1$ when $L_1 = b_2$. In this case

$$P(L_m < \min\{L_1, L_2\} | L_1 = b_1) = 1 - \hat{p}(N_1) = P(L_m < \min\{L_1, L_2\} | L_1 = b_2).$$

If $s = 2$ when $L_1 = b_1$

$$P(L_m < \min\{L_1, L_2\} | L_1 = b_1) = 1 - \hat{p}(N_2)$$

$$\geq \max\{1 - \hat{p}(N_1), 1 - \hat{p}(N_2)\} \geq P(L_m < \min\{L_1, L_2\} | L_1 = b_2).$$

References

1. Arrow, Kenneth. 1962. The Economic Implications of Learning by Doing. *Review of Economic Studies* 29, 155-173.
2. Arrow, Kenneth. 1974. *The Limits of Organization*. New York: Norton.
3. Atkeson, Andrew and Patrick Kehoe. 2002. Measuring Organization Capital, NBER Working Paper 8722.
4. Bahk, Byong-Hyong and Michael Gort. 1993. Decomposing Learning by Doing in New Plants. *Journal of Political Economy* 101, 561-583.
5. Barro, Jason and Robert Barro. 1990. Pay, Performance, and Turnover of Bank CEOs. *Journal of Labor Economics* 8, 448-481.
6. Bassi, Laurie and Mark Van Buren. 2000. Measuring What Matters: Core Measures of Intellectual Capital. ASTD report, July.
7. Becker, Gary. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, Third Edition, Chicago: University of Chicago Press.
8. Benkard, C. Lanier. 1999. Learning and Forgetting: The Dynamics of Aircraft Production. Working Paper. Stanford University Graduate School of Business.
9. Billingsley, Patrick. 1995. *Probability and Measure*, Third Edition, Wiley-Interscience.
10. Bolton, Patrick and Mathias Dewatripont. 1994. The Firm as a Communication Network. *Quarterly Journal of Economics*, 109, 809-839.
11. Bushman, Robert, Raffi Indjejikian and Abbie Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting and Economics*, 21, 161-193.
12. Chen, Peter, Vikas Mehrota, Ranjini Sivakumar and Wayne W. Yu. 2001. Layoffs, shareholders' wealth, and corporate performance. *Journal of Empirical Finance* 8, 171-199.

13. Cohen, Randolph, Christopher Polk and Tuomo Vuolteenaho. 2003. The Value Spread. *Journal of Finance forthcoming*.
14. Cooper, Russell and Alok Johri. 1997. Dynamic complementarities: A quantitative analysis. *Journal of Monetary Economics* 40, 97-119.
15. Crémer, Jacques. 1993. Language and Shared Knowledge. *Industrial and Corporate Change* 2, 351-386.
16. Crémer, Jacques, Luis Garicano and Andrea Pratt. 2003. Codes in Organizations. Working Paper.
17. DeMarzo, Peter, Dimitri Vayanos and Jeffrey Zwiebel. 2001. Persuasion Bias, Social Influence, and Uni-Dimensional Opinions. Working Paper.
18. Denis, David J. and Diane K. Denis. 1995. Performance Changes Following Top Management Dismissals. *Journal of Finance* 50, 1029-1057.
19. Dessein, Wouter and Tano Santos. 2003. The Demand for Coordination. Working Paper. Columbia University Graduate School of Business.
20. Dustmann, Christian and Costas Meghir. 2002. Wages, Experience and Seniority. Working paper, Institute of Fiscal Studies, UCL.
21. Fama, Eugene and Kenneth French. 1995. Size and Book-to-Market Factors in Earnings and Returns. *Journal of Finance* 50, 131-55.
22. Garicano, Luis. 2000. Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy* 108, 874-904.
23. Gregory, Alan and Steve McCorriston. 2002. Foreign Acquisitions by UK Limited Companies: Long-run Performance in the US, Continental Europe and the Rest of the World. Working Paper. School of Business and Economics, University of Exeter.
24. Hallock, Kevin F. 1998. Layoffs, Top Executive Pay, and Firm Performance. *American Economic Review* 88, 711-723.

25. Hanka, Gordon. 1998. Debt and the terms of employments. *Journal of Financial Economics* 48, 245-282.
26. Harris, Milt and Artur Raviv. 2002. Organization Design. *Management Science* 48, 852 - 865.
27. Hart, Oliver. 1989. An Economist's Perspective on the Theory of the Firm. *Columbia Law Review* 89, 1757-1774.
28. Hart, Oliver. 1995. *Firms, Contracts and Financial Structure*. Clarendon Press, Oxford.
29. Hermalin, Benjamin. 2001. Economics and language in Handbook of Organizational Culture and Climate, edited by S. Cartwright et al., Chichester, England: John Wiley & Sons, Ltd.
30. Huselid, Mark. 1995. The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance. *The Academy of Management Journal* 38, 635-672.
31. Ichniowski, Casey. 1990. Human Resource Management Systems and the Performance of U.S. Manufacturing Businesses. NBER Working Paper Series.
32. Ichniowski, Casey, and Kathryn Shaw. 2003. Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices. *Journal of Economic Perspectives* 17, 155-180.
33. Ichniowski, Casey, Kathryn Shaw and Giovanna Prennushi. 1997. The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *American Economic Review* 87, 291-313.
34. Klein, Benjamin. 1988. Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited. *Journal of Law, Economics and Organization* 4, 199-213.

35. Kreps, David M. 1990. Corporate Culture and Economic Theory in J. Alt and K. Shepsle, eds., *Perspectives on Positive Political Economy*. Cambridge University Press, Cambridge, England.
36. Lane, Julia, Allan Isaac and David Stevens. 1996. Firm Heterogeneity and Worker Turnover. Working Paper. American University.
37. Lang, Larry, Rene Stulz and Ralph Walkling. 1989. Managerial Performance, Tobin's Q, and the Gains from Successful Tender Offers. *Journal of Financial Economics* 24, 137-154.
38. Lev, Baruch. 2001. *Intangibles: Management, Measurement, and Reporting*. Brookings Institution Press, Washington, D.C.
39. Macduffie, John Paul. 1995. Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry. *Industrial and Labor Relations Review* 48, 197-221.
40. Maruyama, Nobuhiro and Hiroyuki Odagiri. 2002. Does the 'persistence of profits' persist?: a study of company profits in Japan, 1964-1997. *International Journal of Industrial Organization* 20, 1513-1533.
41. Mitchell, Mark L. and Erik Stafford. 2000. Managerial Decisions and Long Term Stock Price Performance. *Journal of Business* 73, 287-329.
42. Mueller, Dennis. 1986. *Profits in the Long Run*. Cambridge University Press, Cambridge, England.
43. Mueller, Dennis (Ed.). 1990. *The Dynamics of Company Profits: An International Comparison*. Cambridge University Press, Cambridge, England.
44. Parrino, Robert. 1997. CEO Turnover and Outsider Succession: A Cross-Sectional Analysis. *Journal of Financial Economics* 46, 165-197.
45. Prescott, Edward C. and Michael Visschler. 1980. Organization Capital. *Journal of Political Economy* 88, 446-461.

46. Rajan, Raghuram and Luigi Zingales. 1998a. Power in a Theory of the Firm. *Quarterly Journal of Economics* 113, 387-432.
47. Rajan, Raghuram and Luigi Zingales. 1998b. The Governance of the New Enterprise. Working Paper. Graduate School of Business, University of Chicago.
48. Raisel, Ethan and Paul Friga. 2002. *The McKinsey Mind*. McGraw Hill.
49. Rau, P. Raghavendra and Theo Vermaelen. 1998. Glamour, value and the post-acquisition performance of acquiring firms. *Journal of Financial Economics* 49, 223-253.
50. Rosen, Sherwin. 1972. Learning by Experience as Joint Production. *Quarterly Journal of Economics* 86, 366-382.
51. Rumelt, Richard. 1987. Theory, Strategy and Entrepreneurship. In *Competitive Challenge*. Ed. D. I. Teece. Cambridge, MA: Ballinger.
52. Servaes, Henri. 1991. Tobin's Q and the Gains from Takeovers. *The Journal of Finance* 46, 409-419.
53. Simon, H.A. 1979. Information Processing Models of Cognition. *Annual Review of Psychology* 30, 363.
54. Stokey, N. L. and R. E. Lucas, Jr. 1989. *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Massachusetts.
55. Stole, Lars A. and Jeffrey Zwiebel. 1996. Organizational Design and Technology Choice under Intrafirm Bargaining. *American Economic Review* 86, 195-222.
56. Topel, Robert. 1991. Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority. *Journal of Political Economy* 99, 145-176.