Organizational Form and the Market for Talent

Bård Harstad, *Northwestern University*

This article brings together the market for products, the market for talent, and firms' organizational form. While the organizational design determines the allocation of blame and fame within the firm, the value of a good reputation depends on the market structure. Consequently, the market structure dictates the optimal organizational design. If competition becomes tougher and the market thicker, transparent firms decentralize while nontransparent firms concentrate control, transparency itself is improved, corporations switch from unitary to multidivisional form, and the turnover of managers increases. The model rationalizes recent trends in both executive pay and organizational design.

I. Introduction

Reputation and talent are becoming increasingly important for individual careers. Kanter (1989, 310–12) observes that "the creation of star quality is a hallmark in the dynamics of the professional career," and "the key variable in success is reputation." These trends are reflected in executives' pay: the top chief executive officer (CEO) pay has recently increased dramatically relative to both average CEO pay and average

I am grateful for comments from Philippe Aghion, Heski Bar-Isaac, Patrick Bolton, Tore Ellingsen, Guido Friebel, Nicola Gennaioli, Mariassunta Giannetti, Oliver Hart, Jaime Ortega, Torsten Persson, and several seminar audiences. Kathryn Spier and James Dana suggested the title; the previous title was "Organizations and Careers." Christina Lönnblad provided editorial assistance, and, at the early stages of this work, financial support was provided from Tore Browaldh's research foundation. Contact the author at harstad@kellogg.northwestern.edu.

[*Journal of Labor Economics*, 2007, vol. 25, no. 3]

0734-306X/2007/2503-0002$10.00

wages.[1] The value of a top reputation, relative to a moderate reputation, is larger than ever. Frank and Cook (1995) argue that such "superstar" effects are increasingly important in a growing number of markets. A superstar effect arises whenever one producer can serve many customers, and Rosen (1981) showed how this effect depends on the market structure.

At the same time, the possibility of earning a good reputation is determined by organizational design. If all control was centralized, for example, the CEO would receive all blame for a fiasco and all fame for a success. When control is decentralized, the glory of success is instead shared by all managers who may influence the firm's performance. Since organizational design determines the creation of reputations, while the value of a reputation depends on market structure, market structure dictates the optimal organizational design.

This article brings together markets, reputations, and organizations to study how reputational concerns create a link between market structure and organizational design. The results rationalize recent trends in both executive pay and organizational design.

The analysis requires a model with three parts: the product market, the labor market, and organizational design. For the product market, I modify Salop's (1979) simple model of price competition in the circular city. A firm's unit cost of production is assumed to depend on the manager's talent. Thus, with a talented manager, the firm finds it profitable to increase its production. A firm's willingness to pay for a particular manager is thus an increasing and convex function of her expected ability (labeled "reputation"). If competition becomes tougher (more substitutable products) and the market, thicker (more consumers), demand becomes more elastic with respect to (w.r.t.) the price. The best firm is then able to capture a huge share of the market, and the willingness to pay for a good manager, as a function of her reputation, becomes steeper and more convex. This provides a mapping from the market structure to the value of reputation.

The labor market is assumed to be perfectly competitive. There is a large pool of young potential managers with unknown talent, but some of these earn a good reputation after the firms' performances are observed. Everyone can observe the firms' profits and use Bayesian updating to learn about managers' talents. Although some firms hire young managers in every period, these firms could alternatively bid for old managers with a good reputation. Thus, the equilibrium wage for an old good manager equals the amount by which a firm's profit is expected to increase. An-

[1] As documented by Murphy (1999), the average CEO earned about 30 times more than the average production worker in 1970. By 1996, the average CEO received a nearly 90 times greater cash compensation and a total realized compensation of 210 times the earnings for production workers.

ticipating this, young managers prefer jobs where they are likely to earn a good reputation. The firm internalizes the desire to create good reputations since this makes the firm an attractive employer, and current wages can be accordingly reduced. Hence, the firm's problem is to maximize "the value of exposure."[2]

Alternative organizational forms differ in how the market learns about the managers' talents. Since the value of a reputation depends on the market structure, so does the optimal organizational design. I find that if the market becomes thicker and competition tougher, nontransparent firms centralize control, transparent firms decentralize control, transparency itself is enhanced, corporations switch from unitary (U-form) to multidivisional (M-form), and the turnover of managers increases.

Although the argument requires a model of three parts, the model is as simple as possible. For example, I entirely abstract from incentives and effort: managers costlessly make decisions or carry out management tasks revealing their talent. This simplification is quite innocent, however: if the firm can offer a wage contract conditional on performance, this contract can generate incentives independent of organizational design. Alternatively, such performance-contingent wage contracts can perfectly insure the managers if they are risk averse: the firms can simply promise a financial "parachute," should they earn a bad reputation. Thus, there is no loss of generality by assuming managers to be risk neutral.

This article takes a first step in investigating how reputational concerns make the organizational form depend on the market structure. The combined model relies on several strands of literature, which are discussed in the next section. Section III presents the first two parts of the model: the structure of the product market—from which the value of a reputation is derived—and the labor market. When this framework is in place, the analysis of alternative organizational design easily follows in Section IV. As mentioned, the subsections study decentralization in nontransparent firms, decentralization in transparent firms, and the value of transparency and compare M-form and U-form organizations. These results are easiest to illustrate in a 2-period model, but Section IV.E introduces a multiperiod model to discuss equilibrium age and turnover of managers. While the contribution of this article is theoretical, Section V discusses some empirical evidence and suggests that the results may also shed some light on the puzzling correlation between firm size and wages. Concluding remarks are provided in Section VI, while the appendix discusses robustness.

[2] The result, that firms internalize the value of creating reputations, holds in much more general settings, as discussed in the appendix.

II. Related Literature

This section briefly discusses some of the related theoretical literature. Contributions more specifically connected to this article are discussed in the relevant subsections, while Section VI reviews the empirical literature.

That today's performance affects tomorrow's wage is the fundamental pillar of the career-concern literature initiated by Fama (1980) and formalized by Holmström (1999). Dewatripont, Jewitt, and Tirole (1999) survey some results in this literature, which mostly focuses on the incentives to provide effort. Quite a few articles combine career concerns with organizational design. Meyer and Vickers (1997) study how better (comparative performance) information affects incentives, and Kanemoto and MacLeod (1992) show that it may be hard to separate types, unless the labor market is sufficiently competitive. Ortega (2003) models centralization versus decentralization in a way similar to my "nontransparent" firm and finds that some centralization is typically optimal since the larger visibility of the manager encourages more effort. However, for "transparent" firms, where subtasks are observable, Dewatripont et al. (1999) find most effort to be generated by decentralization. Aghion and Tirole (1995) compare U-form and M-form organizations and suggest that the M-form encourages the most effort since individual contributions are easier to identify. Apart from their focus on incentives, all these contributions abstract from the market structure and assume future wages to be linear in expected abilities. With this assumption, there is no value of learning, per se.

A firm typically benefits from learning about its employees' talents if it must make important promotions. Then, the internal labor market breaks the "linearity assumption." In this way, a few recent articles emphasize the value of learning about talent, per se. My closest antecedent is probably the contribution of Demougin and Siow (1994). They analyze fast-track versus up-or-out hiring policies, where the equilibrium regime depends on, for example, the skill premium. Although they discuss the organization of *training* within the firm, this can also be interpreted as *learning* about the employees since some trainees turn out to be successful and others do not. In contrast to Demougin and Siow (1994), I explicitly model the market structure and discuss other aspects of organizational design. Other related articles are Meyer (1994), Jeon (1996), Ortega (2001), and Carrillo (2003). They analyze how to choose task allocation, team composition, and job rotation when the firm needs to learn about the employees' talents to make promotions or efficiently allocate labor. All these articles focus on how the internal labor market determines the value of learning. This article, in contrast, shows how the external labor and product markets interact in determining the value of a reputation and, thus, the optimal organizational design.

The standard assumption that the future wage is linear in reputation is strong. The superstar literature initiated by Rosen (1981) shows that when many consumers can buy from the same provider, the market structure typically generates a convex relationship between wages and reputations. Frank and Cook (1995) discuss how these arguments hold quite generally in many types of markets. My model of the market structure is similar to that of Raith (2003), who studies how the market structure determines explicit incentive contracts. In such a model, the market structure also determines the value of a reputation and, thus, the value of learning about talent. I isolate this effect to study how the market structure influences the optimal organizational design. Thus, the article complements the empirical study by Krueger (2005, 28), who concludes that "our understanding of labor markets will be incomplete unless we better appreciate the interactions among product markets, technology, and labor markets."

III. Markets for Products and Talent

The model consists of three parts: the product market, the labor market, and organizational design. This section presents the first two parts of the model, and they are solved in the following three subsections. Alternative organizational forms are introduced and analyzed in Section IV.

To formalize the product market, I borrow Salop's (1979) workhorse model of price competition in the circular city. The marketplace consists of a large circle with perimeter l. There are a large number of potential firms. The n firms that decide to operate in the market are automatically located equidistant from one another on the circle.[3] The firms simultaneously set prices, and each consumer buys exactly one product. The consumers are uniformly distributed on the circle with a density m. Since they face linear travel costs $t > 0$, each consumer buys from the producer that minimizes the sum of expenses: price and travel cost. Parameter t can literally be interpreted as consumers' travel costs, as the firms' monopolistic power, or, more generally, as measuring the extent to which goods are substitutable. Its inverse, $1/t$, will be referred to as the "toughness of competition." Parameter m can be interpreted as the size of the market, the number of consumers or their budget, and will be referred to as the "thickness" of the market. However, a larger m and a lower t will both increase competition in the sense of prices decreasing.

[3] Economides (1984) vindicates the equal-spacing assumption in a model with quadratic transport costs. In the present model, a firm i would be indifferent to its location between two existing firms: i's demand is given by eq. (6) whenever firms $i - 1$ and $i + 1$ are distanced by $l/2n$, notwithstanding i's location in between these. Thus, equidistant location is a best response.

As in Raith (2003), the manager's role is to reduce costs. Firm i's unit cost is measured by

$$c_i = c - y_i, \tag{1}$$

which is stochastic since the outcome of the cost reduction task,

$$y_i = a_i + \epsilon_i, \tag{2}$$

depends on the manager's ability a_i and some shock ϵ_i, and both of these are unknown to everyone. The shock ϵ_i is firm specific and normally distributed with mean zero and variance σ_ϵ^2:

$$\epsilon_i \sim N(0, \sigma_\epsilon^2).$$

I assume there is a large pool (excess supply) of young potential managers. Their reservation wage is normalized to one,[4] and their abilities are independently drawn from the following distribution:

$$a_i \sim N(\bar{a}, \sigma_a^2).$$

To capture the dynamics of careers, let there be 2 periods. Based on the firm's performance in the first period, its manager earns a reputation, defined as the expected ability relative to the average:

$$r_i \equiv E(a_i|y_i) - \bar{a}. \tag{3}$$

Manager i's reputation depends on y_i, which everyone can estimate after observing the firms' prices and profits. Public learning about talent is not only simplifying but arguably the appropriate assumption when studying individual reputations. By the law of large numbers, only half of the managers earn a positive reputation before the second period. Good managers are therefore in short supply.[5] By restricting the analysis to 2 periods, the value of reputation (in the second period) is separated from the value of learning (in the first). This makes the analysis clean and the effects clear. There is no need for discounting. Section IV.E introduces multiple periods to discuss the age and turnover of managers.

Figure 1 illustrates the timing in each period. First, each firm decides whether to operate in the market. Thereafter, the firms simultaneously make hiring decisions and set prices. Each consumer buys from the pro-

[4] Alternatively, the reservation wage could depend on a manager's reputation. For example, the results below would be identical if the reservation wage were given by $1 + \max\{0, zE(a_i - \bar{a})\}$, if just $z < \sqrt{m/t}$. Crucial for the results is that good (and only good) managers are retained as managers in the second period.

[5] This is clearly necessary if young managers are to be hired in every period. Moreover, even if all firms have two managers, Sec. IV.E shows managers with sufficiently good reputations to be in short supply in the multiperiod model. The appendix discusses how the results would be modified if good managers were in excess supply.

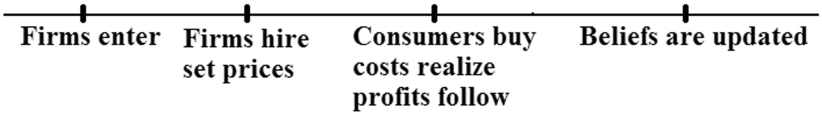

FIG. 1.—Timing of each period

ducer that minimized the generalized cost (price plus transport cost), and, then, the shocks and the unit costs are realized. Thus, prices are set before a firm knows the realization of its unit cost and before it knows the hiring decision of its neighbors. This timing simplifies the analysis a great deal, without qualitatively changing the results. Note that firms can fire or hire managers between periods, and managers are free to leave.[6] Although both firms and managers are assumed to be risk neutral, the appendix shows the results to hold also if managers are risk averse.

A. The Product Market

The following subsections solve the above model. In the product market, I let each consumer consider only the two closest firms and buy from the firm that minimizes the total costs: price plus transportation costs.[7] Thus, by reducing its price, a firm sells more products, and this is more valuable, the lower is its expected unit cost. The firm's surplus is therefore convex in the manager's ability: not only does a talented manager reduce the cost of all units but she also makes it profitable for the firm to reduce its price and increase its production. As Rosen (1982) describes, the CEO has a chain letter–like effect on the firm's performance since her influence increases with the firm's production.

Let $\bar{r}$ denote the average reputation of managers.

LEMMA 1. Expected surplus (eq. [4]) is a convex function of the manager's reputation r:

$$Es(r) = \frac{m}{t}\left(\frac{tl}{n} + \frac{r - \bar{r}}{2}\right)^2. \tag{4}$$

Proof. Consider a consumer at location x, between firm 1 (at location

[6] These assumptions are not crucial, however. The appendix argues that the results would hold even if managers "belonged" to their firms.

[7] Although standard, this is just an approximation of the firms' demand functions. In principle, one could image consumers between firms i and $i + 1$ to buy from firm $i + 2$, if the last is very competitive. The appendix discusses this possibility and argues that the results would then be reinforced.

0) and firm 2 (at location l/n). This consumer buys from firm 1 instead of firm 2 if

$$p_1 + tx \leq p_2 + t\left(\frac{l}{n} - x\right) \Rightarrow x \leq \frac{l}{2n} + \frac{p_2 - p_1}{2t}. \tag{5}$$

Considering both a firm's closest competitors and the density of consumers m, firm i's total demand becomes

$$\frac{ml}{n} + \frac{m}{t}\left(\frac{p_{i-1} + p_{i+1}}{2} - p_i\right). \tag{6}$$

Assuming positive demand, firm i's surplus is

$$s_i = m(p_i - c_i)\left[\frac{l}{n} + \frac{1}{t}\left(\frac{p_{i-1} + p_{i+1}}{2} - p_i\right)\right].$$

Let $\bar{p}$ denote the (expected) average price in the market, so that $Ep_j = \bar{p}$. If n is large, Ep_j and Ec_i are uncorrelated, so that the expectation of s_i becomes

$$Es_i = m(p_i - Ec_i)\left[\frac{l}{n} + \frac{1}{t}(\bar{p} - p_i)\right]. \tag{7}$$

Firm i thus chooses its price in order to solve

$$\max_{p_i} Es_i \Rightarrow p_i = \frac{Ec_i + \bar{p}}{2} + \frac{tl}{2n}. \tag{8}$$

Since all other firms do the same, $\bar{p} = \bar{c} + t/n$, where $\bar{c}$ is the expected average unit cost. By substituting for the prices, Es_i becomes

$$Es_i = \frac{m}{t}\left(\frac{tl}{n} + \frac{\bar{c} - Ec_i}{2}\right)^2. \tag{9}$$

Substituting with equations (1)–(3) completes the proof. QED

The findings in this section are stated as lemmas for two reasons. First, they will be instrumental for the main results of the article. Second, emphasizing them is worthwhile, although they are not novel. As anticipated, the profit function is convex, a firm's demand decreases in its price and more quickly so if t is low. Prices thus increase in t. To say more, the next subsection solves for the number of firms and takes into account the labor market.

B. The Market for Old Managers

While the product market is similar for every period, this subsection concentrates on the labor market in the last period. Good old managers with a positive reputation ($r > 0$) are in short supply, and some firms

must, in equilibrium, hire unknown managers with zero reputation ($r = 0$). These firms may bid for a good manager and are thus willing to pay the amount by which her reputation is expected to increase their surplus. This will determine the wage premium for a good manager, defined as the pay above the reservation wage of one. Since profit is convex in the managers' reputation, so are the managers' wages. If m increases and t decreases, demand becomes more elastic with respect to price, and firms with the best managers set the lowest price and capture a huge amount of the market. Firms are then willing to pay a great deal to hire the best managers, and, as a function of reputation, wages become steeper and more convex.

LEMMA 2. An old manager's wage premium (eq. [10]) is an increasing and convex function of her reputation r. This function becomes steeper and more convex if competition increases (t decreases) and the market becomes thicker (m increases):

$$w(r) = r\sqrt{\frac{m}{t}} + r^2\frac{m}{4t}. \tag{10}$$

Proof. As noted, some firms will hire managers with zero reputation even in the second period, and these enter until their surplus covers the reservation wage of one:

$$Es(0) - 1 = \frac{m}{t}\left(\frac{tl}{n_2} - \frac{\bar{r}}{2}\right)^2 - 1 = 0 \Rightarrow n_2 = \frac{tl}{\sqrt{t/m} + \bar{r}/2}, \tag{11}$$

where n_2 denotes the number of firms in the second period. Substituting into equation (4),

$$Es(r) = \frac{m}{t}\left(\sqrt{\frac{t}{m}} + \frac{r}{2}\right)^2. \tag{12}$$

How much more is a firm willing to pay for a manager with reputation $r > 0$? A firm is willing to pay the amount by which the expected surplus increases:

$$w(r) = Es(r) - Es(0) = r\sqrt{\frac{m}{t}} + r^2\frac{m}{4t}.$$

This is, indeed, the salary a firm has to pay to attract such a manager since a large number of firms bid for a smaller number of managers with a positive reputation. QED

This observation is in line with the superstar literature. Already Rosen (1981) showed that a decreasing transport cost increases both the variance and the skewness of the wage distribution relative to the distribution of

abilities. And, as discussed by MacDonald (1988), these prospects affect the young generation's willingness to expose their potential for the market.

Remark 1.—Note that the average reputation $\bar{r}$ enters neither in equation (10) nor in equation (12). Although the surplus in equation (4) decreases in $\bar{r}$, so does the equilibrium number of firms n_2, and the two effects cancel since the profit is pinned down when firms choose whether to enter (parameter l disappears for the same reason). This implies that there is no general equilibrium effect of learning and that the value of learning is not affected by the number of firms in period 1.

Remark 2.—Since $\bar{r}$ and l drop out, $w(\cdot)$ solely depends on m/t. By hiring a manager with high r, a firm is able to capture more consumers by reducing its price. The sensitivity of demand is, by equation (6), proportional to m/t: it increases in the density of consumers (m) and their willingness to switch suppliers ($1/t$).[8] Since the following analysis depends on the market structure only via $w(\cdot)$, all comparative static can focus on m/t.

C. The Market for Young Managers

After the first period, participants in the labor market observe firms' performances and use Bayesian updating to learn about managers' talents. Since one's future wage is a convex function of reputation, being exposed as a manager may be good for a young worker. If m increases or t decreases, the best firm captures a huge market share, and the value of a top reputation increases relative to the value of a moderate reputation. This makes the wage function more convex, and the value of exposure increases.

LEMMA 3. The future reputation of a manager is distributed according to equation (13). The value of exposure (eq. [14]) is increasing in the toughness of competition ($1/t$) and the thickness of the market (m).

$$r \sim N(0, \sigma_r^2), \text{ where } \sigma_r^2 \equiv \frac{\sigma_a^4}{\sigma_\epsilon^2 + \sigma_a^2}. \tag{13}$$

$$v(\sigma_r) = \sigma_r \sqrt{\frac{m}{2\pi t}} + \sigma_r^2 \frac{m}{8t}. \tag{14}$$

Proof. Remember that cost reduction y is determined by

$$y = a + \epsilon, \text{ where } a \sim N(\bar{a}, \sigma_a^2);\ \epsilon \sim N(0, \sigma_\epsilon^2). \tag{15}$$

[8] In addition, the revenues of existing demand increase since the production cost decreases. Existing demand is also increasing in m and decreasing in t since when t increases, firms become more profitable, n becomes larger, and each firm's market share becomes smaller.

After observing the outcome in the first period, the second-period belief about this manager's ability becomes (using standard Bayesian updating)

$$a|y \sim N\left(\lambda\bar{a} + [1-\lambda y], \lambda\sigma_a^2\right) = N\left(\bar{a} + [1-\lambda][y-\bar{a}], \lambda\sigma_a^2\right),$$

$$\text{where } \lambda = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_a^2}. \tag{16}$$

The expected posterior belief about the manager's ability (relative to $\bar{a}$) is her reputation, $r = (1-\lambda)(y-\bar{a})$. It follows from equations (15) and (16) that before a manager has undertaken any task, her second-period reputation r is distributed as

$$r \sim N(0, \sigma_r^2), \text{ where } \sigma_r^2 \equiv \frac{\sigma_a^4}{\sigma_\epsilon^2 + \sigma_a^2}. \tag{17}$$

Let $f(r)$ denote the probability distribution function of equation (13). Combined with the value of a certain reputation (eq. [10]), we can calculate a first-period manager's future expected wage premium, which may be called the value of exposure:

$$v(\sigma_r) = \int_0^\infty w(r)f(r)dr = \sigma_r\sqrt{\frac{m}{2\pi t}} + \sigma_r^2\frac{m}{8t}.$$

QED

The larger the value of exposure $v(\sigma_r)$, the more attractive it is to be a manager in the first period. The firm can thus pay $1 - v(\sigma_r)$, and a young manager is still willing to accept the job.[9] If possible, any firm would therefore like to maximize $v(\sigma_r)$. That the firm internalizes the managers' future benefits (here in the form of higher expected wages) is standard in the literature, and it follows from the theory of "equalizing differences" initiated by Adam Smith (1947) and surveyed by Rosen (1986).[10] If the salary of young managers for some reason was fixed, it would still seem likely that the firm would maximize $v(\sigma_r)$ after some (Coasian) bargaining, where the manager also had some bargaining power. Moreover, even if it was costly to switch jobs or the managers "belonged" to their firms, firms

[9] It might be surprising that, in the first period, the wage is reduced for the worker appointed to become a manager. But this hinges on the assumption that all workers have the same reputation in period 1. For a given reputation, it indeed seems reasonable that the wage decreases as the career prospects increase. The same result is found by MacDonald (1988) and Demougin and Siow (1994). Terviö (2006) shows that if young managers cannot pay up front for exposure, firms will employ too few young managers.

[10] The prediction is also empirically documented. For example, Olson (2002) finds that workers accept lower wages if the employer provides health benefits.

would be induced to maximize $v(\sigma_r)$ since that would maximize their own value of learning about their managers' talent. This is discussed in the appendix.

But how can the firm maximize $v(\sigma_r)$? While the organizational design above was very simple, that is, one single manager makes one single decision, alternative organizations will be analyzed in the next section. These organizations differ in the way their managers are exposed, that is, in σ_r. Since $v(\cdot)$ depends on the market structure (m/t), so will the optimal organizational design.

IV. Organizational Forms

It is time to analyze alternatives to the simple organizational design above. Most firms have multiple management tasks that must be undertaken, and these tasks can either be centralized to one manager or decentralized to several. For example, some decisions or tasks concern quality improvements and other cost reductions.[11] Equivalently, cost reduction itself may consist of two subtasks, A and B:

$$y = \frac{y_A + y_B}{2}. \tag{18}$$

The performance of each task depends on the decision maker's ability and some random task-dependent shock:

$$\begin{aligned} y_A &= a_A + \epsilon_A; \\ y_B &= a_B + \epsilon_B. \end{aligned} \tag{19}$$

Variable a_A (a_B) is the ability of the manager undertaking task A (B). For simplicity, the two task-dependent shocks have the same distribution and are uncorrelated. To make this model a generalization of the one above, the sum of the two shocks is simply ϵ:

$$\begin{aligned} \epsilon &= \frac{\epsilon_A + \epsilon_B}{2}; \\ \epsilon_A, \epsilon_B &\sim N(0, \sigma^2), \text{ where} \\ \sigma^2 &= 2\sigma_\epsilon^2. \end{aligned}$$

[11] In fact, y_i could be interpreted as a task increasing product quality. Suppose that firm i's expected quality Eq_i raises the consumers' willingness to pay accordingly. If we let k_i represent marginal cost, all equations would continue to hold if we just defined $Ec_i \equiv Ek_i - Eq_i$ and if consumers were risk neutral. Thus, higher expected quality has the same effect as lower expected cost since both allow the firm to raise its markup between price and expected cost, and y_i could be interpreted as increasing q_i or decreasing k_i.

In the previous section, both subtasks were centralized to one single manager (such that $a_A = a_B$). The next subsections will compare this to decentralization, where two managers divide the two tasks, first, for a "nontransparent" firm where only aggregate performance y is observable, then, for a "transparent" firm where the outcomes of both subtasks y_A and y_B are observable to everyone. The value of transparency itself is thereafter discussed.[12] The fourth subsection slightly modifies equation (18) to compare U-form and M-form corporations. Section IV.E introduces a multiperiod model to analyze the age and turnover of managers.

The potential benefit of all alternative designs comes from the value of exposure. The cost of decentralization is that another manager must be hired. Thus, in a static framework, as well as in the second period, the simple design with centralization is optimal. The value of reputation in the second period and exposure in the first are therefore still determined by equations (10) and (14).

A. Centralization in Nontransparent Firms

Decentralization implies that several managers share control. In a nontransparent firm, no one observes the different subtasks y_A and y_B, but everyone can calculate the firm's aggregate performance y. With two managers, 1 and 2, the observed performance is

$$y = \frac{a_1 + a_2}{2} + \epsilon,$$

and this outcome determines the reputations of both managers.[13]

In the present model, the advantage of having two managers is that we learn about two employees, not only one. A successful firm "graduates" two managers with a positive reputation. The advantage of concentrating control is that we learn more about one single manager, and we are more willing to believe that good performance is due to this particular manager's ability. Hence, the chance of earning an excellent reputation is better in the centralized firm. If competition increases and the market becomes thicker, the value of one excellent relative to two

[12] Note that for both transparent and nontransparent firms, all learning is public, and there is no asymmetric information. This definition differs from that allowing the firm to observe even if it is nontransparent. Andersson (2002) studies such transparency of explicit wage contracts.

[13] Ortega (2003) studied exactly this formalization of power concentration within firms. He found that some concentration of power was optimal since this maximizes the incentives to provide effort. He followed most of the literature by assuming wages to be linear in reputation, ignoring the market structure and, thus, the value of creating individual reputations, per se.

moderate reputations increases, which makes concentration of control superior.[14]

PROPOSITION 1. Nontransparent firms concentrate control if competition increases and the market becomes thicker (m/t increases).

Proof. In a decentralized firm, the posterior belief about each manager's ability becomes (using standard Bayesian updating)

$$a|y \sim N\big(\bar{a} + 2[1-\lambda)][y-\bar{a}], \lambda\sigma_a^2\big), \text{ where } \lambda = \frac{\sigma_a^2 + 4\sigma_\epsilon^2}{2\sigma_a^2 + 4\sigma_\epsilon^2}.$$

Before observing y, each of the managers' future reputation $r = 2(1-\lambda)(y-\bar{a})$ is distributed as

$$r \sim N(0, \sigma_{DN}^2), \text{ where } \sigma_{DN}^2 = \frac{\sigma_a^4}{2\sigma_a^2 + 4\sigma_\epsilon^2}. \tag{20}$$

The reputation of a manager in a decentralized nontransparent firm has the variance $\sigma_{DN}^2 < \sigma_{CN}^2 = \sigma_a^4/(\sigma_\epsilon^2 + \sigma_a^2)$, which is the variance of the manager's reputation in a centralized nontransparent firm (labeled σ_r^2 in the previous section). The benefit of concentrating control becomes

$$v(\sigma_{CN}) - 2v(\sigma_{DN}) + 1$$

$$= (\sigma_{CN} - 2\sigma_{DN})\sqrt{\frac{m}{2\pi t}} + (\sigma_{CN}^2 - 2\sigma_{DN}^2)\frac{m}{8t} + 1,$$

where the first parenthesis is negative and the second positive. As m/t is low, the first term dominates and the sum is negative (if the additional wage cost of one is relatively small). As m/t increases, however, the second term dominates, and the sum becomes positive. QED

The proposition says that for firms where the outcomes of various subtasks are indistinguishable, centralization is the optimal organizational design if the industry is competitive and the market is large. Centralizing control enhances transparency since individual contributions become observable. Therefore, it is unclear whether it is transparency or task allocation that drives proposition 1. To illuminate this, the next subsection analyzes decentralization in a transparent firm.

B. Decentralization in Transparent Firms

In a transparent firm, both subtasks y_A and y_B are observable. Still, centralization teaches us more about one manager, while decentralization teaches us about two managers. For the centralized firm, transparency is

[14] Another effect of centralizing control is that profit will be more volatile since it depends on one manager's talent only. This effect is, indeed, documented by Adams, Almeida, and Ferreira (2005).

of no importance since average performance y is a sufficient statistic for estimating the manager's ability a. For the decentralized firm, transparency allows us to learn more by separating individual performances. The value of more learning increases as m/t increases. This overturns the result of proposition 1.

PROPOSITION 2. Transparent firms decentralize control if competition increases and the market becomes thicker (m/t increases).

Proof. In a decentralized firm, each manager undertakes one task, and, according to equation (13), her future reputation has the distribution

$$r \sim N(0, \sigma_{DT}^2), \text{ where } \sigma_{DT}^2 \equiv \frac{\sigma_a^4}{\sigma_a^2 + \sigma^2} < \sigma_{CT}^2. \tag{21}$$

In a centralized firm, the manager's future reputation is distributed according to equation (17), just as before ($\sigma_{CT}^2 = \sigma_{CN}^2 = \sigma_r^2$). The benefit of concentrating control becomes

$$v(\sigma_{CT}) - 2v(\sigma_{DT}) + 1$$

$$= (\sigma_{CT} - 2\sigma_{DT})\sqrt{\frac{m}{2\pi t}} + (\sigma_{CT}^2 - 2\sigma_{DT}^2)\frac{m}{8t} + 1,$$

where both parentheses are negative. The expression is positive if m/t is small, because of the additional wage costs. As m/t increases, however, the negative terms dominate, and decentralization becomes optimal. QED

While a manager in a centralized firm carries out twice as many observable tasks as one in a decentralized firm, this does not imply that we learn twice as much about the manager in the centralized firm. After the first performance has been observed, the uncertainty about the manager's ability is reduced, as is the lesson we draw from observing the outcome of one additional task. Marginal learning is decreasing in the number of tasks, and allocating the two tasks to different managers maximizes the total amount of learning. The value of this, in turn, is larger when the market is thick and competition is tough.

Separating subtasks to another firm by outsourcing is a way of decentralizing in a transparent way, and the theory predicts this to occur more often in competitive and large markets. The contrast to proposition 1 reveals transparency to be a crucial characteristic of the firm when it comes to reputational concerns. The next subsection discusses how the value of transparency itself depends on the market structure.

C. The Value of Transparency

Compare the two decentralized firms in Sections IV.A and IV.B. By observing the outcome of each subtask y_A and y_B, we are able to make a more precise estimate about each manager's ability. We are more willing

to believe that good performance is due to one particular manager's talent. It follows that each manager's chance to earn an excellent reputation is larger in the transparent firm. If competition is tough and the market thick, the premium for an excellent reputation is large relative to that for a moderate reputation, which makes transparency more important.

PROPOSITION 3. Decentralized firms become more transparent if competition increases and the market becomes thicker (m/t increases).

Proof. In a decentralized firm, each manager's future reputation is distributed according to equation (20) or equation (21) if the firm is nontransparent or transparent, respectively. Since $\sigma_{DT}^2 > \sigma_{DN}^2$, the value of transparency is

$$2V(\sigma_{DT}) - 2V(\sigma_{DN}) = (\sigma_{DT} - \sigma_{DN})\sqrt{\frac{2m}{\pi t}} + (\sigma_{DT}^2 - \sigma_{DN}^2)\frac{m}{4t},$$

where both parentheses are positive. There might also be some technical or administrative cost associated with transparency, but as m/t increases, the terms above become very large and tend to dominate. QED

As noted in the previous subsection, transparency is of no importance if the firm is centralized. For the decentralized firm, proposition 3 actually follows as a corollary to propositions 1 and 2, and it may not be surprising at this point. Nevertheless, the result is explicitly stated since it is interesting in its own right. The proposition states that transparency is more important if the industry is competitive and its market large. There are many organizational reforms that may increase a firm's transparency, for example, outsourcing certain subtasks or clearly separating different responsibilities. The next subsection shows that this insight can be applied to compare U-form and M-form organizations.

D. U-form or M-form?

Variables A and B may be interpreted as two different products produced by the same firm. For each product, responsibility may be separated into two subtasks, 1 and 2, such that the production costs are given by

$$y_A = \frac{a_{1A} + a_{2A}}{2} + \epsilon_A;$$

$$y_B = \frac{a_{1B} + a_{2B}}{2} + \epsilon_B. \tag{22}$$

Variable a_{ij} is the ability of the manager responsible for decision $i \in \{1, 2\}$ for product $j \in \{A, B\}$. In an M-form corporation, responsibility is allocated according to product, such that one manager is responsible for all tasks associated with one product. Then, $a_{1j} = a_{2j} = a_j$ for $j \in \{A, B\}$. Equation (22) is then identical to equation (19). In a U-form cor-

poration, responsibility is allocated according to function, such that one manager is responsible for one particular task for all products. Then, $a_{iA} = a_{iB} = a_i$ for $i \in \{1, 2\}$ and equation (22) becomes

$$y_A = \frac{1}{2}(a_1 + a_2) + \epsilon_A;$$

$$y_B = \frac{1}{2}(a_1 + a_2) + \epsilon_B.$$

A classical example of the U-form was the early Ford Motor Company, which was organized into a number of functionally specialized departments: production, sales, purchasing, and so on. In contrast, General Motors is the prototypical M-form since it comprises a collection of fairly self-contained divisions, for example, Chevrolet, Oldsmobile, and Pontiac.

There is a large literature comparing U- and M-forms. The U-form corporation is typically assumed to provide some gains from specialization, κ, since each manager can concentrate on one type of task. The advantage of the M-form corporation is that it reveals more information about individual contributions.[15] This makes it less costly to encourage managers to provide effort by explicit contracts since conditional variance is reduced (Maskin, Qian, and Xu 2000). Aghion and Tirole (1995) suggest also that implicit incentives are larger in the M-form corporation since there is less moral hazard in teams. While my model abstracts from incentives, the emphasis on market competition provides a new rationale for the M-form. The advantage of more information is that the labor market can make a more precise estimate about each manager's ability. We are more willing to believe that good performance for a product is related to the manager's ability if she is solely responsible for this product. Thus, each manager's chance to earn an excellent reputation is larger in the M-form corporation. The M-form's advantage dominates as tougher competition and a thicker market boost the premium for excellent relative to moderate reputations.

PROPOSITION 4. Firms switch from U-form to M-form if competition increases and the market becomes thicker (m/t increases).

[15] Other advantages of the M-form have also been suggested. Williamson (1975) claims that the CEO becomes overloaded in the U-form. Milgrom and Roberts (1992) suggest that the M-form is better at coordinating finance and investment decisions. Qian, Roland, and Xu (2006) emphasize the M-form's larger flexibility w.r.t. experimentation. In all these articles, the benefit of the U-form is increasing returns to scale.

Proof. In the U-form corporation, y is a sufficient statistic for estimating $(a_1 + a_2)/2 \equiv \tilde{a}$. Thus, just as in the proof of lemma 3,

$$\tilde{a}\,|\,y \sim N\big(\bar{a} + [1 - \lambda][y - \bar{a}], \lambda\sigma_{\tilde{a}}^2\big), \text{ where } \lambda = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_{\tilde{a}}^2}$$

and $\sigma_{\tilde{a}}^2 = \sigma_a^2/2$. Since the market cannot identify individual contributions, $Ea_i = E\tilde{a}$. *Ex ante*, then, each manager's future reputation is distributed according to

$$r_i \sim N(0, \sigma_U^2), \text{ where } \sigma_U^2 = \frac{\sigma_{\tilde{a}}^4}{\sigma_{\tilde{a}}^2 + \sigma_\epsilon^2} = \frac{\sigma_a^4}{2\sigma_a^2 + 2\sigma^2}.$$

In the M-form corporation, future reputation is distributed according to $r \sim N(0, \sigma_M^2)$, where $\sigma_M^2 = \sigma_{DT}^2$ in equation (21); $\sigma_M^2 > \sigma_U^2$, and the value of switching from U- to M-form is

$$2v(\sigma_M) - 2v(\sigma_U) - \kappa = (\sigma_M - \sigma_U)\sqrt{\frac{2m}{\pi t}} + (\sigma_M^2 - \sigma_U^2)\frac{m}{4t} - \kappa,$$

where both parentheses are positive and κ is the U-form's benefit of specialization. If m/t is small, a positive κ makes the expression negative. As m/t increases, however, the positive terms dominate, and the M-form becomes superior. QED

E. Turnover of Managers

The above analysis detected two possible concerns for firms when hiring managers. In the second period, a good old manager allows the firm to profitably increase its production. In the first period, there is a value from learning about young unknown managers. In reality, these two concerns are often in conflict.

To illustrate the trade-off, consider a multiperiod model with overlapping generations of managers. Each generation lives 2 periods, and the discount factor is δ. If a firm hires (or keeps) a good old manager with a reputation $r \geq 0$, the surplus in this period increases by $[s(r) - s(0)]$. Alternatively, a firm may hire a young manager with unknown talent. The benefit from this is that since she is unknown, she is a potential superstar, and her value of exposure makes her cheap to hire: the firm only needs to pay $[1 - \delta v(\sigma_r)]$. The firm's willingness to pay for the old manager is thus the sum of these two brackets:

$$s(r) - s(0) + 1 - \delta v(\sigma_r).$$

Since the old manager's reservation wage is one, she is hired if and only if $r \geq \hat{r}$,[16] given by

$$s(\hat{r}) - s(0) = \delta v(\sigma_r).$$

Clearly, $\hat{r} > 0$. Thus, good old managers with positive reputations $r \in (0, \hat{r})$ are replaced by younger ones with less (expected) talent.[17] The possibility of the young managers being very talented dominates the value of a "fairly" good reputation, particularly when m/t is large. While a tougher market increases the value of a good reputation, there is a relatively larger increase in the value of an excellent reputation. The chance of discovering a superstar will then dominate the benefit of keeping a fairly good manager. Thus, when m/t increases, so does $\hat{r}$ and the turnover of managers.

Just as described in the previous sections, a firm may hire more than one manager when hiring young managers.[18] In any case, the following result holds.

PROPOSITION 5. Young managers replace old managers with reputations $r < \hat{r}$, where $\hat{r} > 0$. The threshold $\hat{r}$ increases in m/t, implying that turnover increases as competition becomes tougher and the market thicker.

Proof. In the multiperiod model, the expected surplus is equation (4), just as before. Firms hiring young managers each hire a number $h \in \{1, 2\}$ and pay each manager $1 - \delta v$. Such firms enter until their expected surplus equals the wages. Thus, the number of firms, n, is

$$Es(0) = \frac{m}{t}\left(\frac{tl}{n} - \frac{\bar{r}}{2}\right)^2 = h(1 - \delta v) \Rightarrow$$

$$n = \frac{tl}{\sqrt{th(1 - \delta v)/m} + \bar{r}/2}. \tag{23}$$

And $n < \infty$ requires $\delta v < 1$. Substituting equation (23) in equation (4) gives surplus as a function of r. Define

$$\Delta(r) \equiv Es(r) - Es(0) = r\sqrt{\frac{mh(1 - \delta v)}{t}} + r^2\frac{m}{4t}. \tag{24}$$

[16] The function v is different from that in the 2-period case. In particular, v depends on the equilibrium $\hat{r}$ since reputations $r < \hat{r}$ are worthless.

[17] These results are also found by Demougin and Siow (1994) in the case of trainees and by Carrillo and Mariotti (2001) in the case of political candidates. For an analysis of promotions and turnover where the firm has private information about its employees' talents, see Owan (2004).

[18] However, it is easily shown that a firm hiring an old good manager will centralize control and choose the U-form.

For the old manager, a firm is willing to pay the amount by which its expected profit increases, $\Delta(r) + h(1 - \delta v)$. The wage premium is thus

$$w(r) = \Delta(r) + (h - 1) - \delta h v.$$

The value of exposure v is implicitly defined by

$$v = \int_{\hat{r}}^{\infty} w(r)dF = \int_{\hat{r}}^{\infty} \left[\Delta(r) + (h - 1) - \delta h v\right]dF, \tag{25}$$

where F is the cumulative distribution function corresponding to f. The threshold $\hat{r}$ is given by

$$\begin{aligned} w(\hat{r}) = 0 \Rightarrow \Delta(\hat{r}) + (h - 1) - \delta h v = 0 \Rightarrow \\ G(\hat{r}, \sqrt{m/t}) \equiv \Delta(\hat{r}) + (h - 1) - \delta h v \\ = \Delta(\hat{r}) + (h - 1) - \frac{\delta h \int_{\hat{r}}^{\infty} \Delta(r) + (h - 1)dF}{1 + \delta h \int_{\hat{r}}^{\infty} dF} = 0. \end{aligned}$$

And

$$\frac{\partial \hat{r}}{\partial \sqrt{m/t}} = -\frac{G_2}{G1},$$

which is positive since, first,

$$G_1 = \frac{\partial \Delta(\hat{r})}{\partial \hat{r}} - \delta h \frac{\partial v}{\partial \hat{r}} = \frac{\partial \Delta(\hat{r})}{\partial \hat{r}} > 0$$

and, second, $G_2 < 0$, which can be seen by first defining

$$\begin{aligned} s(r) \equiv \frac{\partial \Delta(r)/\partial \sqrt{m/t}}{\Delta(r) + h - 1} \\ = \frac{r\sqrt{h(1 - \delta v)} - r\delta\sqrt{mh/t(1 - \delta v)}\partial v/\partial\sqrt{m/t} + r^2\sqrt{m/t}/2}{r\sqrt{mh(1 - \delta v)/t} + r^2 m/4t + h - 1}, \end{aligned}$$

which is increasing in r, and then writing G_2 as

$$\begin{aligned} G_2 &= s(\hat{r})[\Delta(\hat{r}) + (h-1)] - \frac{\delta h \int_{\hat{r}}^{\infty} s(r)[\Delta(r) + (h-1)]dF}{1 + \delta h \int_{\hat{r}}^{\infty} dF} \\ &= s(\hat{r})\left\{\Delta(\hat{r}) + h - 1 - \frac{\delta h \int_{\hat{r}}^{\infty} [s(r)/s(\hat{r})][\Delta(r) + (h-1)]dF}{1 + \delta h \int_{\hat{r}}^{\infty} dF}\right\} \\ &< s(\hat{r})\left\{\Delta(\hat{r}) + h - 1 - \frac{\delta h \int_{\hat{r}}^{\infty} [\Delta(r) + (h-1)]dF}{1 + \delta h \int_{\hat{r}}^{\infty} dF}\right\} \\ &= s(\hat{r})G(\hat{r}, \sqrt{m/t}) = 0. \end{aligned}$$

QED

The number of firms will be the same in every period in the steady state. Since the firms only compete for managers with a reputation $r > \hat{r} > 0$, less than half the managers end up with a sufficiently good reputation, if we rely on the law of large numbers. Thus, even if all firms hire two managers, good managers will be in short supply, and some firms hire young managers in every period, just as anticipated above.

V. Predictions and Evidence

The model provides several predictions for how the market structure affects executive wages and organizational design. These should also change over time if the market structure evolves. This section summarizes the main predictions of the article and relates them to some anecdotal evidence. Although I do not provide any data, it might seem plausible that, if anything, product markets have indeed become more "competitive" and "thicker." Consumers' transport costs and search costs have declined, and the variety of products has increased. The number of consumers as well as their budgets have increased, and the size of a typical market is larger than before. Thus, I will discuss some of the predictions of the model in this light. The first subsection discusses the predictions for executives' wages, as these appear in Section III. This is important since the model in Section III drives the organizational changes in Section IV. The next subsection summarizes the predictions for organizational change. Finally, it is shown that the theory can also shed some light on the puzzling correlation between firm size and wages.

A. Executive Pay

Lemma 2 in Section III makes two main predictions. First, since $w(\cdot)$ becomes steeper as m/t increases, top executives' salaries should increase as competition becomes tougher and the market thicker. Since these changes also make $w(\cdot)$ more convex, the second prediction is that the

variance and skewness in pay should increase in m/t. Lemma 3 shows that the value of learning also increases in m/t. A young executive pays up front for this future value, in the form of lower wages today. It follows that the wage gap between young and old executives should increase in m/t. Finally, since abilities are revealed over time, wage inequality within one cohort should increase with the cohort's age.

Executive pay also varies with the organizational changes discussed in Section IV. The general lesson from propositions 1–5 is that learning about the managers increases as m/t increases. This reveals more about the managers' abilities, and the variance in their reputations increases. This, in turn, increases the difference in their wages. Therefore, tougher competition and thicker markets increase the variance in CEO pay both directly and indirectly through organizational change.

All these predictions are in line with the evidence: Murphy (1999) showed how CEO pay had tripled relative to average pay over the preceding 3 decades, and he found the relative CEO pay to be particularly large in the United States (where the market might be thicker and competition tougher). The CEO pay has also become more skewed, in that top CEO pay has increased relatively more than average CEO pay. For the income distribution in general, Neal and Rosen (2000) document that both inequality and skewness have increased over time, and these are certainly larger in the United States than in most developed countries. They also find that average wages, wage inequality, and skewness within one cohort all increase with the cohort's age and that these differences among cohorts have increased over time. They notice that standard sorting or matching models of the labor market can explain how average wage, inequality, and skewness increase with a cohort's age, but these theories are less suitable for explaining the time trend or the explosion in CEO pay. Barth (1997) studies the positive relationship between seniority and wages within firms and finds the seniority effect not to be firm specific. Thus, it cannot be explained by human-capital accumulation, but it is consistent with lemmas 2 and 3.

B. Organizational Change

Section IV provides several predictions for organizational change. If m/t increases, propositions 1–3 state that nontransparent firms should centralize control, transparent firms should decentralize control, and transparency itself should increase. Proposition 4 claims that corporations should switch from U- to M-form, and proposition 5 suggests that the turnover of managers should increase in m/t.

The interesting contrast between propositions 1 and 2 might be difficult to test, however, since it is not clear how transparency should be measured. But since proposition 3 suggests that transparency itself should increase,

the overall effect may be that firms decentralize control as m/t increases. Such decentralization is, indeed, well documented (see, e.g., Caroli and Van Reenen 2001). Moreover, one certain way of decentralizing in a transparent way is to outsource tasks to different firms. Outsourcing and smaller firms are both parts of recent trends. Abraham and Taylor (1996), for example, discuss recent growth in outside contracting and alternative explanations.

The shift from U-form to M-form in the last century is well documented by Chandler (1962). Several rationales for this shift are provided in the literature (see, e.g., Williamson 1975; Milgrom and Roberts 1992; Aghion and Tirole 1995; Maskin et al. 2000). While these contributions clarify the benefits of the M-form, they typically fail to show that the benefits of the M-form have increased over time, relative to those of the U-form. This shift in relative benefits is the central point of proposition 4.

That the probability of forced CEO turnover increases with bad performance is empirically documented by Gibbons and Murphy (1990) and Huson, Parrino, and Starks (2001). Moreover, Murphy (1999) finds old average-performing CEOs to have a much higher departure probability (36.3%) than young executives realizing returns 30% below the industry average (8.5%). Huson et al. (2001) show that the frequency of forced turnover has increased in 1971–94, while Hadlock and Lumer (1997) find that turnover has increased since the 1930s. Fee and Hadlock (2000) find that turnover increases with competition in the newspaper industry. All these facts are consistent with proposition 5.

C. Firm Size and Wages

The positive correlation between firm size and wages is well documented. Although some explanations are summarized by Oi and Idson (1999), there is still no consensus on the answer.[19] The model in this article suggests three alternative reasons for a correlation between firm size and wages.

First, consider the second period. A firm that hires a manager with a good reputation will certainly have to pay a higher salary. At the same time, this firm will produce more and expect a larger profit due to lower costs. The positive correlation between size and executive pay is immediate.

Second, if industries differ in their market size m and competition t, then the willingness to pay (eq. [10]) for the best managers will be largest in the industries where m/t is largest. According to equations (6) and (12), these are also the industries where the production and profit per firm are

[19] After discussing several alternatives, Gibbons and Katz (1992, 530) conclude that "we know of *no* model that fits all the facts," while Brown and Medoff (1989, 1056) conclude that "our analysis leaves us uncomfortably unable to explain it."

largest. Once more, there is a positive correlation between size and executive pay. Such arguments for matching large firms and good managers are analyzed by Rosen (1982), and they lead to two additional predictions. Since m/t affects profit and production as well as salaries, the relationship between size and wages should not vary with m/t. Murphy (1999) indeed documents this relationship to be remarkably stable over time as well as across countries. Moreover, since proposition 5 suggests that turnover should be higher in industries where m/t is large, there should be a positive correlation between firm size and executive turnover. This correlation is also documented by Murphy (1999).[20]

However, studies by, for example, Brown and Medoff (1989) and Gibbons and Katz (1992) find that there is still a substantial size-wage effect, even after controlling for worker characteristics. In fact, even piece-rate workers earn more in larger firms. This suggests that wages might be higher to compensate for inferior working conditions. But, these authors argue, this argument fails to account for the correlation, in particular because working conditions appear to be similar across firms within the same industry.

The arguments of Section IV suggest an explanation. The performance of a large firm depends on a large number of workers. If the performance is good, these workers only marginally improve their individual reputation. The market is less able to learn about an employee if the firm is large.[21] Since salaries are convex in expected abilities, workers prefer to work for small firms where their talents are more easily recognized. Hence, larger firms must compensate for lower career possibilities by higher wages. Since the value of learning is increasing in m/t, the size-wage effect may increase over time and should be larger in countries where m/t is large. Oi and Idson (1999) do show that the size-wage effect is larger in the United States than in Europe and that the effect was larger in 1983 than in 1979.[22]

VI. Conclusion and Future Research

By combining models of market structure, reputations, and organizational design, this article shows how reputational concerns lead to a link

[20] Relatedly, Baker and Hall (2004) show that since the manager of a large firm affects the entire production, her incentives (and pay) must accordingly be stronger. Also, competitive assignment models generate a positive relation between firm size and CEO pay (Gabaix and Landier 2007; Terviö 2007).

[21] That learning takes longer in large firms is consistent with the finding (by Brown and Medoff 1989) that there is less worker turnover in large firms. Slower learning reveals bad matches between firms and employees at a later stage.

[22] However, Oi and Idson (1999) suggest that the last finding could be due to a recession in 1983, and this time span is too short anyway to "test" the theory. In addition, increased labor mobility over the last decades might have pushed toward a weaker size-wage relationship, offsetting the force suggested above.

between market structure and optimal organizational design. Section III showed that the executive's future wage, as a function of expected ability, is steeper and more convex, the tougher the competition and the thicker the market. This motivates firms to become more transparent, decentralize control, switch from U-form to M-form, and hire younger managers. These predictions are in line with Kanter's (1989, 307) observation of the modern firm: "Highly decentralized organizations with matrix structures or project assignments tend to provide general management responsibilities to many more people, much earlier in their careers." Apart from rationalizing recent organizational changes, the model may improve our understanding of executive wages and why these differ across firms.

However, this article is just a small step toward a deeper understanding of these issues. Future research may analyze other organizational designs than those above. My model of the market structure (borrowed from Salop 1979) is quite simple, and a more detailed study will lead to new results. Perhaps most important, the predictions of the model ought to be related to empirical evidence across industries, as well as over time.

Reputations are important also for other types of organizations. Instead of studying the organization within firms, the theory could be applied to studying the organization between firms. Tadelis (1999) suggests that the existence of firms' reputations induces firms to secretly trade their names, and he also shows how such trade might enhance the incentives to build a reputation in the first place (Tadelis 2002). Cabral (2000) applies these ideas to study the motivation for umbrella branding. Common for this literature—as for the literature on career concerns—is the assumption of a linear relationship between ability and its value. Section III above showed that this assumption typically fails if the product market is taken into account. Then, the value of a certain reputation will depend on the market structure, while the allocation of tasks determines how blame and fame are distributed between firms. This raises a host of questions related to the boundary of the firm. Which combination of tasks will a firm prefer to produce in house, and which should be delegated to other firms? Which set of products will be produced under the same brand? How do the answers to these questions depend on the firm's age, reputation, and—in particular—the market structure?

Appendix

The argument of this article requires models of market competition, learning, and organizational design. The combined framework could easily be intractable, unless simplifying assumptions are made. To check the robustness of the model, however, a number of these assumptions deserve a closer look.

A. The Problems of the Normal Distribution

It is assumed above that both talent and noise, and thus firms' unit costs, are normally distributed. "The assumption of normality is very convenient analytically but has the drawback that prices and quantities may take negative values. However, the possibility of these phenomena can be controlled by controlling the variances of the random variables" (Vives 1999, chap. 8 n. 6). These assumptions are thus quite standard.

Note that in the above model, it is the expected unit cost Ec_i that is important since a firm sets its price before the exact cost is realized. Moreover, since a firm would never hire a manager with a negative reputation, extremely large values of Ec_i are not relevant. For extremely low values of Ec_i, that is, when a firm hires a real superstar, Ec_i may be negative. This is not a problem analytically and could be interpreted as the firm's quality being very large (see n. 11).

A problem arises, however, when a firm i manages to capture the buyer at the turf of firm $i + 1$ (i.e., the buyer whose location coincides with firm $i + 1$). Then, firm i captures all consumers of firm $i + 1$, its demand jumps (discontinuously), and it starts competing with firm $i + 2$. This possibility is typically ignored by simply letting a firm compete with its two closest neighbors only (Tirole 1988, 283). Raith (2003, 1427) justifies this approximation by stating that "Large *random* cost differences can be ignored if the variance of the cost shocks σ^2 is sufficiently small."

If this possibility is not ignored, the results of this article appear to be reinforced. Note that firm i captures all consumers of firm $i + 1$ as soon as $p_{i+1} - p_i \geq tl/n$. Substituting from equations (5) and (8), this is expected to occur when $r_i \geq 2\sqrt{t/m} + 2\bar{r}$. Then, firm i's demand is expected to jump by the size of an average firm's "backyard," that is, $ml/2n = (\sqrt{m/t})/2 + (m/t)\bar{r}/4$. Thus, fixing $\bar{r}$, the discrete jumps in x_i are expected to be larger and come sooner (for a smaller r_i) when m/t is large. Compared to the above analysis, this makes $s(r)$ and $w(r)$ even more convex when m/t increases, thus reinforcing the results.[23]

B. The Supply and Number of Good Managers

While the next subsection argues that the results hold even if a firm can retain and hold up its good managers, the above analysis presumes a perfect labor market where a good manager extracts the entire value of her reputation. This is possible because the number of managers with a sufficiently good reputation is smaller than the number of firms. Clearly, this condition must be satisfied in any steady state where juniors (with zero reputation) are to be hired in every period. Moreover, it follows by

[23] However, since $\bar{r}$ depends on the organizational design and, thus, on the market structure, there may be multiple equilibria (for some values of m/t). The analysis of this case must thus be left to future research.

the law of large numbers in Section IV.E since, even if all firms have two managers, less than n managers end up with a reputation $r > \hat{r}$. Allowing for exit (i.e., assuming that a fraction of young managers die after their first period), this would hold also in the 2-period model. The exact number of good managers is of no importance, as long as the number of firms is larger.

However, with a smaller number of firms (where we cannot rely on the law of large numbers), a situation may occur where the number of successful managers from decentralized firms exceeds the number of firms. What would then happen? If $\tilde{r}$ denotes the reputation level such that the number of managers whose $r > \tilde{r}$ is identical to the number of firms, then, in equilibrium, $w(r) = 0$ for all $r \leq \tilde{r}$, while $w(r) = Es(r) - Es(\tilde{r})$ for $r > \tilde{r}$. If workers anticipated that only reputations $r > \tilde{r} > 0$ are valuable, they would prefer more exposure to increase this likelihood. Thus, when the expected $\tilde{r}$ increases, desire increases for more transparency, the M-form, and a larger turnover.

C. If Firms Have More Bargaining Power

The above analysis assumed a perfect labor market with no switching costs for managers. Kanemoto and MacLeod (1992) instead let the worker face some cost C if switching jobs. As they show, a worker's wage equals her outside option minus this fixed cost (if C is not too large). With small switching costs, the analysis thus ends up being identical to that above.

However, switching jobs may be ruled out if C is huge or if a manager's performance is private information to the firm. If the firm has all the bargaining power over the wage, it pays all workers their reservation wage and directly extracts the values of reputations $Es(r)$, without the need for requiring the workers to pay up front for this as juniors. Once more, the firm designs its organization to maximize the value of learning about its employees, just as studied above. However, if a firm has no desire to discover two successful managers (if it cannot use them in another plant or "sell" them to another firm), the case for decentralization is weakened.[24]

D. Risk Aversion

The assumption of risk neutrality might seem restrictive since alternative organizations have different implications for individual risk. Relaxing this assumption will not change the results, however.

A risk-neutral firm might be able to fix a performance-contingent salary for risk-averse employees. Since there is no moral hazard, the firm can

[24] In any case, switching costs may have declined over time. Frydman (2005) suggests that firm-specific skills may have become less important relative to more general skills and argues that this contributes to explaining the increases in CEO pay and executive turnover.

offer a perfect insurance to managers. This means that a manager's first-period salary becomes $1 - \delta w(r)$, that is, the salary is larger if the manager is unsuccessful. This compensates the manager for a lower future salary following a fiasco, and it appears as a financial parachute. Tougher competition and a thicker market increase this parachute relative to the manager's salary following good performance. Since financial parachutes seem to have increased over the last decades, this prediction seems reasonable.

Even if the firm is risk averse or if it cannot commit to performance-contingent salaries, the results appear to survive. Let each manager have the concave utility function $u[w(r)]$. The risk aversion adds a cost related to transparency and exposure. Learning is beneficial only if u is convex with respect to r, which requires $w(\cdot)$ to be "more" convex than $u(\cdot)$ is concave. As competition increases and the market becomes thicker, $w(\cdot)$ becomes more convex, and it becomes more likely that u is convex with respect to r. Therefore, a larger m/t makes exposure more beneficial (or less costly) relative to a lower m/t, which makes a firm more likely to hire new managers, adopt transparency and the M-form, centralize control if it is nontransparent, and decentralize control if it is transparent.

References

Abraham, Katharine G., and Susan K. Taylor. 1996. Firms' use of outside contractors: Theory and evidence. *Journal of Labor Economics* 14, no. 3:394–425.

Adams, Renée B., Heitor Almeida, and Daniel Ferreira. 2005. Powerful CEOs and their impact on corporate performance. *Review of Financial Studies* 18, no. 4:1403–32.

Aghion, Philippe, and Jean Tirole. 1995. Some implications of growth for organizational form and ownership structure. *European Economic Review* 39, nos. 3–4:440–55.

Andersson, Fredrik. 2002. Career concerns, contracts, and effort distortions. *Journal of Labor Economics* 20, no. 1:42–59.

Baker, George P., and Brian J. Hall. 2004. CEO incentives and firm size. *Journal of Labor Economics* 22, no. 4:767–99.

Barth, Erling. 1997. Firm-specific seniority and wages. *Journal of Labor Economics* 15, no. 3, pt. 1:495–506.

Brown, Charles, and James Medoff. 1989. The employer size-wage effect. *Journal of Political Economics* 97, no. 5:1027–59.

Cabral, Luis. 2000. Stretching firm and brand reputation. *RAND Journal of Economics* 31, no. 4:658–73.

Caroli, Eve, and John Van Reenen. 2001. Skill biased organizational change? Evidence from a panel of British and French establishments. *Quarterly Journal of Economics* 116, no. 4:1449–92.

Carrillo, Juan D. 2003. Job assignments as a screening device. *International Journal of Industrial Organization* 21:881–905.
Carrillo, Juan D., and Thomas Mariotti. 2001. Electoral competition and politician turnover. *European Economic Review* 45, no. 1:1–25.
Chandler, Alfred. 1962. *Strategy and structure*. New York: Doubleday.
Demougin, Dominique, and Aloysius Siow. 1994. Careers in ongoing hierarchies. *American Economic Review* 84, no. 5:1261–77.
Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999. The economics of career concerns, part II: Application to missions and accountability of government agencies. *Review of Economic Studies* 66:199–217.
Economides, Nicholas. 1984. The principle of minimum differentiation revisited. *European Economic Review* 24:345–68.
Fama, Eugene F. 1980. Agency problems and the theory of the firm. *Journal of Political Economy* 88, no. 2:288–307.
Fee, C. Edward, and Charles J. Hadlock. 2000. Management turnover and product market competition: Empirical evidence from the U.S. newspaper industry. *Journal of Business* 73, no. 2:205–43.
Frank, Robert H., and Philip J. Cook. 1995. *The winner-take-all society*. New York: Free Press.
Frydman, Carola. 2005. Rising through the ranks: The evolution of the market for corporate executives, 1936–2003. Working paper, Sloan School of Management, MIT.
Gabaix, Xavier, and Augustin Landier. 2007. Why has CEO pay increased so much? Working paper, Department of Economics, MIT.
Gibbons, Robert, and Lawrence F. Katz. 1992. Does unmeasured ability explain inter-industry wage differentials? *Review of Economic Studies* 59, no. 3:515–35.
Gibbons, Robert, and Kevin Murphy. 1990. Relative performance evaluation for chief executive officers. Special issue, *Industrial and Labor Relations Review* 43 (February): 30–51.
Hadlock, Charles J., and Gerald B. Lumer. 1997. Compensation, turnover, and top management incentives. *Journal of Business* 70, no. 2:153–87.
Holmström, Bengt. 1999. Managerial incentive problems—a dynamic perspective. Special issue, *Review of Economics Studies* 66, no. 1:169–82.
Huson, Mark, Robert Parrino, and Laura Starks. 2001. Internal monitoring mechanisms and CEO turnover: A long-term perspective. *Journal of Finance* 56, no. 6:2265–97.
Jeon, Seonghoon. 1996. Moral hazard and reputational concerns in teams: Implications for organizational choice. *International Journal of Industrial Organization* 14:297–315.
Kanemoto, Yoshitsugu, and W. Bentley MacLeod. 1992. The ratchet effect and the market for secondhand workers. *Journal of Labor Economics* 10, no. 1:85–98.
Kanter, Rosabeth Moss. 1989. *When giants learn to dance: Mastering the*

challenge of strategy, management, and careers in the 1990s. New York: Simon & Schuster.
Krueger, Alan B. 2005. The economics of real superstars: The market for rock concerts in the material world. *Journal of Labor Economics* 23, no. 1:1–30.
MacDonald, Glenn M. 1988. The economics of rising stars. *American Economic Review* 78, no. 1:155–66.
Maskin, Eric, Yingyi Qian, and Chenggang Xu. 2000. Incentives, information, and organizational form. *Review of Economic Studies* 67, no. 2:359–78.
Meyer, Margaret A. 1994. The dynamics of learning with team production: Implications for task assignment. *Quarterly Journal of Economics* 109, no. 4:1157–84.
Meyer, Margaret A., and John Vickers. 1997. Performance comparisons and dynamic incentives. *Journal of Political Economy* 105, no. 3:547–81.
Milgrom, Paul, and John Roberts. 1992. *Economics, organization, and management*. Englewood Cliffs, NJ: Prentice Hall.
Murphy, Kevin J. 1999. Executive compensation. In *Handbook of labor economics*, vol. 3B, ed. Orley Ashenfelter and David Card. Amsterdam: Elsevier Science.
Neal, Derek, and Sherwin Rosen. 2000. Theories of the distribution of earnings. In *Handbook of income distribution*, vol. 1, ed. Anthony B. Atkinson and Francois Bourguignon. Amsterdam: North Holland.
Oi, Walter Y., and Todd L. Idson. 1999. Firm size and wages. In *Handbook of labor economics*, vol. 3B, ed. Orley Ashenfelter and David Card. Amsterdam: Elsevier Science.
Olson, Craig A. 2002. Do workers accept lower wages in exchange for health benefits? *Journal of Labor Economics* 20, no. 2, pt. 2:S91–S114.
Ortega, Jaime. 2001. Job rotation as a learning mechanism. *Management Science* 47, no. 10:1361–70.
———. 2003. Power in the firm and managerial career concerns. *Journal of Economics and Management Strategy* 12, no. 1:1–29.
Owan, Hideo. 2004. Promotion, turnover, earnings, and firm-sponsored training. *Journal of Labor Economics* 22, no. 4:955–78.
Qian, Yingyi, Gérard Roland, and Chenggang Xu. 2006. Coordination and experimentation in M-form and U-form organizations. *Journal of Political Economy* 114, no. 2:366–402.
Raith, Michael. 2003. Competition, risk, and managerial incentives. *American Economic Review* 93, no. 4:1425–36.
Rosen, Sherwin. 1981. The economics of superstars. *American Economic Review* 71, no. 5:845–58.
———. 1982. Authority, control, and the distribution of earnings. *Bell Journal of Economics* 13, no. 2:311–23.
———. 1986. The theory of equalizing differences. In *Handbook of labor*

economics, vol. 1, ed. Orley Ashenfelter and David Card. Amsterdam: Elsevier Science.

Salop, Steven C. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10, no. 1:141–56.

Smith, Adam. 1947. *An inquiry into the nature and causes of the wealth of nations*. New York: Modern Library.

Tadelis, Steven. 1999. What's in a name? Reputation as a tradeable asset. *American Economic Review* 89, no. 3:548–63.

———. 2002. The market for reputations as an incentive mechanism. *Journal of Political Economy* 110, no. 4:854–82.

Terviö, Marko. 2006. Superstars and mediocrities: Market failure in the discovery of talent. Working paper, Haas School of Business, University of California, Berkeley.

———. 2007. The difference that CEOs make: An assignment model approach. Working paper, Haas School of Business, University of California, Berkeley.

Tirole, Jean. 1988. *The theory of industrial organization*. Cambridge, MA: MIT Press.

Vives, Xavier. 1999. Oligopoly pricing: Old ideas and new tools. Cambridge, MA: MIT Press.

Williamson, Oliver. 1975. *Markets and hierarchies*. New York: Free Press.