Executive Pay, Talent and Firm Size*

Jaeyoung Sung†
University of Illinois at Chicago and Ajou University, Korea

and

Peter L. Swan‡
Australian School of Business, University of New South Wales

This version: October 22, 2008

Abstract

We present an integrated agency model of career concerns and labor market equilibrium. Unlike the existing literature, our managerial reservation utility levels and thus their pay levels are endogenously determined, and managers with high expected talent levels are not necessarily hired by large firms. A number of our theoretical results are supported by our panel data for 1992-2006 exclusive of time factors, which strikingly suggest that the average talent level of large-firm CEOs is actually slightly lower than that for small-firm CEOs, but these large-firm CEOs are drawn from a much tighter talent distribution that helps to explain their far higher pay. We show that CEOs employed in larger firms are more productive due to scale economies in effort and more notably in talent. Talent is rewarded via both higher pay and CEO income from shareholdings. Finally, a sizeable portion of the increased real CEO pay levels over recent decades is explicable as compensation for higher risk.

Key words: executive pay, firm size, career concern, CEO talent, principal-agent, optimal contract.

JEL Classification: G34, J41, J44, L25

---

* We wish to thank the Australian Research Council (ARC) for financial support and Vijay Philip, Jasper Timm and Cybele Wong for valuable programming and related support. We thank seminar participants at UNSW, FIRN Research Day, Jeff Coulton, Alex Edmans, David Feldman, Denzil Fiebig and Ning Gong for useful comments.

† Contact details: Department of Finance, College of Business Administration, University of Illinois at Chicago, email: jsung@uic.edu; Tel: (312) 996-0720

‡ Contact details: Department of Banking and Finance, Australian School of Business, University of New South Wales, Email: peter.swan@unsw.edu.au; Tel: 61 (0)2 9385 5871.

1. Introduction

A strong positive correlation between firm size and executive pay has become one of the most highly documented facts in the area of executive compensation for many decades over many countries.[1] In particular, executive pay increases in the size of the firm with approximately a one-third higher pay level for each doubling of firm size. However, the reasons for this are not well understood. At one extreme, Rosen (1982) hypothesizes that higher pay is due to greater talent while at the other higher CEO pay associated with acquisitions and getting larger is simply a demonstration of managerial power (see, for example, Bebchuk and Fried (2004, pp. 127-30)).[2] One of our aims is to help explain these peculiar empirical findings and to show that neither the talent nor rent-seeking perspectives alone can explain the rapid rise in CEO pay. Rather, a combination of rising firm size and CEO aversion to risk appear to be the major contributors.

In this paper, we present an integrated agency model of career concerns and labor market equilibrium, in which both executive pay levels and incentive contracts are endogenously determined. Then, we provide empirical results supporting our model by using panel data for 1992-2006 exclusive of time factors.

[1] See, for example, Roberts (1956), Lewellen and Huntsman (1970), Cosh (1975), Murphy (1985), Baker, Jensen, and Murphy (1988), Kostiuk (1990), Barro and Barro (1990), Rosen (1992), Joskow et al (1993), Rose and Shepard (1997), and Frydman and Saks (2007). Their findings suggest that a 100 percent larger firm will pay its CEO about thirty-three percent more. Zhou (1999) finds a positive correlation between executive pay and firm size for Canadian firms. Kaplan (1997), Kato (1997) and Kato and Kubo (2006) find a similar result for Japanese firms; Cosh and Hughes (1997) and Conyon (1997), McKnight, and Tomkins (1999) and Conyon and Murphy (2000) for British firms; and Merhabi, Pattenden, Swan and Zhou (2006) for Australian firms.

[2] In fact, the President, George W. Bush, has blamed the sub-prime crisis on “excessive” executive pay with Richard Fuld, CEO of the failed investment bank, Lehman Brothers, allegedly receiving $480m.

It is well known in the existing agency literature that the expected pay of an agent is the sum of the agent reservation certainty equivalent wealth and compensation for effort production. However, the literature typically assumes the agent's reservation utility as an exogenous parameter, and thus it does not provide a model that can be used to compare different executive pay levels across firms with different sizes. We build a principal-agent model with reservation utility levels endogenously determined though labor market competition.

Our model is based on a two-period economy consisting of two firms of different sizes and two agents with unknown talent levels drawn from two different distributions. In each period, the two firms compete with each other to hire the better of the two agents, and consequently, agent reservation utility/pay levels are endogenously determined. For each period, each firm signs a contract with one agent chosen from two candidates with unknown abilities. After the first period, each firm again makes its hiring decision for the second period. The hiring decision will be made with each agent's past performance record taken into account. His past record would provide each agent with differing negotiating power for second-period contracting, and thus a different reservation utility level. In our framework we would expect to see quite different pay outcomes depending on firm size and "manager reputation" based on track record.

We argue that in executive labor-market equilibrium, reservation wealth is made up of compensation for wealth creation due to the agent's effort and talent had he been hired by the small (reference) firm plus compensation for any future job-market disadvantages he may face because of his working for the current firm. We believe ours to be the first estimable agency model to explicitly include managerial talent.

The endogenous reservation wealth levels enable us to characterize hiring decisions by the firms, and to explicitly compute the managerial pay differential between the small and large firms. Recall that the matching literature (e.g., Gabaix and Landier (2008)) a priori assumes that large firms (or firms with better production technologies) hire managers with superior abilities, and argues that there should be a positive relationship between pay and firm size because large firms hire managers with superior talent who thus "deserve" higher pay (e.g., Rosen (1982)). However, in our agency world, a large firm (with better production technology) may not always be willing to hire a high expected ability manager if there is too high a level of uncertainty in ability, because this uncertainty hurts work incentives. We argue that even when a large firm hires a low-ability manager, the expected pay for the low-ability manager can be higher than that for a high-ability manager who is hired by a small firm. This requires that the large firm's productivity and size are sufficiently higher and larger than those of the small firm. Thus large firms can in equilibrium hire low-ability CEOs but pay them as if they were high-ability CEOs. Our empirical findings support this hypothesis. In fact, we find that on average the CEOs of large firms are actually of slightly lower ability than those of small firms, holding the technology of the two firm types the same, yet are paid far more. Despite the similar means, larger firms benefit from reduced talent dispersion which in turn is particularly advantageous for large firms.

In order to empirically test our theoretical result, we estimate the stochastic CEO production function measuring the ability of CEOs to convert total assets under management at the beginning of each year into total claimant wealth at the end. Approximately 19,000 CEO-years from a sample of S&P 1500 firms over the period, 1992-2006, are evaluated on the basis of their performance when subjected to moral hazard. All performance measures are recorded in the dollars of 2006 and are thus in real terms. These “internal” wealth creation measures are based on both stock market and accounting numbers. Surprisingly, in 37% of CEO-years based on the market measure (19% based on the accounting measure), these internal contributions to productivity are negative in real terms; indicating that poor productivity performance is widespread, especially due to volatility in equity markets, and that much apparent company growth is externally funded.

The plan of the paper is as follows: Section 2 reviews the literature, the model is developed in Section 3 and the empirical estimation is in Section 4. Section 5 concludes.

2. Literature Review

Our paper is most closely related to the agency literature on pay-size relationship, pay sensitivity, and career concerns. Recently, Gabaix and Landier (2008) develop a calculable competitive assignment model of CEO pay, under the assumption that the best managers are paired with the largest firms. Based on the extreme value theory, the authors argue that a negligible difference in managerial talent of only 0.016% between the CEO ranked number 250 and the top CEO accounts for pay for the top manager that is 500% higher than for manager number 250. This raises the quandary as to why the market for executive talent does not appear to clear. In particular, if the alternative for the most talented executive assigned to the largest firm is to be employed by a smaller firm, say # 250, why does not the largest firm offer (say) just $1 more than firm #250 for a manager of almost precisely the same ability, rather than pay 500% more? Contrary to Gabaix and Landier, we find that the distribution of

CEO ability is significant and for the large firms lower than for of small firms such that the risk-adjusted mean (equivalent of the Sharpe ratio) is considerably higher.

Apart from the Gabaix and Landier assignment approach, there have been other attempts to empirically proxy managerial talent. For example, Rajgopal, Shevlin and Zamora (2006) employ a combination of newspaper publicity and return on assets.

Giannetti (2007) develops a theoretical model in which the growth in job hopping opportunities for risk-neutral CEOs leads to higher CEO pay. While we model the CEO labor market, job hopping *per se* does not affect the level of pay in our model. Moreover, within our dataset of slightly under 19,000 CEO years, job movements are surprisingly few.[3]

In contrast to recent decades, Frydman and Saks (2007) find only a weak relationship between compensation and firm size from the late 1940s to the mid-1970s. These findings for the earlier period suggest that technological advances in the last 35 years such as computers have increased the ability of able executives to manage very large companies successfully. Scale economies in effort and talent that we identify have most likely been augmented by the technological revolution and in particular the introduction of computers.

Hermalin (2005) develops a model of board monitoring in which all participants are risk-neutral, in contrast to our assumption that CEOs are risk averse. Since the CEO is

[3] There have been several other attempts to try and reconcile the pay-size premium with partial explanations of the phenomena put down to effects, such as, compensating differentials by Dunn (1986), union status by Lewis (1986), and efficiency wages by Krueger and Summers (1988). Idson and Oi (1999) and Bayard and Troske (1999) both find that workers in larger firms achieve higher labor productivity, based on some restrictive measures of labor productivity. Bebchuk and Grinstien (2006) find that there is an economically and statistically significant positive relationship between CEO compensation and the CEO's past decisions to increase firm size, by means of increasing the number of shares on issue.

not compensated for risk, it can be advantageous for firms to externally recruit CEOs of unknown talent with the intention of termination if some performance standard is not met.

The pay sensitivity issue has emerged as an important issue in the agency literature since Jensen and Murphy (1990) argued that there is a negative relationship between sensitivity and firm size. The Jensen and Murphy (1990) finding is modeled by Schaefer (1998) taking into account larger team sizes in bigger firms. More recently, Edmans, Gabaix and Landier (2007) have introduced a multiplicative specification for an agency model with both incentives and talent assignment that can explain why equity incentives fall with firm size.

Our model explicitly accounts for firm size effects on contracting and our empirical estimates show that pay-performance sensitivity optimally falls with size. Moreover, our modeling tells us how sensitivities are affected by executives' career paths. Given our estimated elasticities, our model predicts that a manager recruited from a small to a large firm will be given a lower pay-performance sensitivity that a manager recruited from a firm of the same size. This is because the large firm's performance provides a more reliable signal of the manager's ability with less risk being borne by the CEO and thus warrants the use of higher-powered incentives.

Baker and Hall (2004) estimate a form of a production function of CEO effort, and document that CEO efforts increase in pay-performance sensitivity of the manager. However, their production function ignores the impact of managerial talent on output. Their estimated effort elasticity based on market value ranges from as low as 37 percent up to 66 percent and thus overlaps with our estimate for large firms of 46%, after controlling for managerial ability.

Our third focus, apart from the estimation of pay sensitivity, and the pay-size relationship, is executive career concern. Fama (1980) provocatively argued in the absence of formal modeling that the managerial labor market could provide a perfect substitute for incentive pay by rewarding managers with high reputations for talent even though there is a moral hazard problem due to the unobserved nature of the manager's actions. Holmstrom (1999) formally modeled such career concern issues. He showed that the less is known about managerial ability the greater the incentive for the manager to supply effort. Gibbons and Murphy (1992) also model career concern issues to argue that explicit contracts should provide stronger incentives as executives approach retirement as the impact of the implicit incentives provided by the labor market decline with the prospect of retirement. As far as career concerns are concerned, our paper is closely related to Gibbons and Murphy. We argue that CEO career concerns not only affect the sensitivity but their negotiating power in the future labor market and thus the current pay size.

3. The Two-Period Career Concerns Model

There are two firms, S (small) and L (large), and two agents a and b with unknown abilities, θ^a and θ^b, respectively. The firms are risk neutral and both agents are risk averse with the same constant absolute risk aversion (CARA) coefficient, r. We assume that $\theta^{\hat{a}}$, for $\hat{a} \in \{a,b\}$, is normally distributed with mean $m_{\theta^{\hat{a}}}$ and variance $\sigma^2_{\theta^{\hat{a}}}$, and that θ^a and θ^b are independent of each other. There are thus two dimensions to managerial talent, namely, the mean and variance. One may say that agent $\hat{a}$ is more talented if his talent is drawn from a distribution with a higher mean, $m_{\theta^{\hat{a}}} > m_{\theta^{\hat{b}}}$, or a tighter distribution, $\sigma^2_{\theta^{\hat{a}}} < \sigma^2_{\theta^{\hat{b}}}$. The initial signal enabling expected ability levels to differ between the candidates could be résumés' indicating that one candidate has better educational attainments or track-record to date.

There are two periods with three dates, 0, 1, and 2. Contracting between the two firms and two agents occurs at time 0 and 1. At time $t-1$, for $t \in \{1,2\}$, firm i $i \in \{S, L\}$ hires agent $\hat{a}$, and the agent exerts effort $e_{t-1}^{\hat{a}}$ to produce outcome $Y_t^i(\hat{a})$, where the production function for CEO output takes the additively separable form:

$$Y_t^i(\hat{a}) = e_{t-1}^{\hat{a}} f(K^i) + \theta^{\hat{a}} g(K^i) + h(K^i)\varepsilon_t^i, \tag{1}$$

and $K^S < K^L$ where $K^i > 0$ indicates the ith's firm's capital stock and thus size; and ε_t^i, for $t \in \{1,2\}$ and $i \in \{S, L\}$, are independent normal random variables, each of which are distributed with a mean of zero and a standard deviation of σ. The ability (or talent) $\theta^{\hat{a}}$ may represent agent $\hat{a}$'s decision-making competency or information-gathering ability to identify better investment opportunities.

The functions $f(K^i)$ and $g(K^i)$, respectively, describe how firm size and scale economies affect the agent's marginal productivity of effort and ability, and $h(K^i)$ signifies the way risk (dollar volatility) varies with firm size. We assume that $f(K) = K^{\gamma_f}$, $g(K) = K^{\gamma_g}$, and $h(K) = K^{\gamma_h}$, for some $\gamma_f, \gamma_g, \gamma_h > 0$. Then, the first two terms of equation (1) imply that the expected outcome is given as the sum of two Cobb-Douglas production functions: the first in labor effort, $e_{t-1}^{\hat{a}}$, and capital, K, and the other in expected ability, $m_{\theta^{\hat{a}}}$, and capital K, reflecting the manner in which talent reaps scale economies in assets under management distinctively from effort. The last term of the equation captures the random element in the productive process with its volatility increasing in K.

At time $t \in \{0,1\}$, agent $\hat{a}$ exerts effort $e_t^{\hat{a}}$, incurring a personal (monetary) cost of $c(e) = (\kappa/2)e^2$ for some $\kappa > 0$. Hence the shadow cost of effort is independent of either ability or career profile. At time $t \in \{1,2\}$, agent $\hat{a}$ working for firm i is compensated by an amount $S_t^i(Y_t^i(\hat{a}))$, and the utility of the agent takes the form $-\exp\left(-r\sum_{t=1}^{2}\left\{S_t^i(\hat{a}) - c(e_{t-1}^{\hat{a}})\right\}\right)$. Without loss of generality and in the spirit of Holmstrom and Milgrom (1987) and Schattler and Sung (1993), we assume, for $t \in \{1,2\}$, that there is a linear pay schedule with a fixed and a performance component:[4]

$$S_t^i(Y_t^i(\hat{a})) = \alpha_{t-1}^i(\hat{a}) + \beta_{t-1}^i(\hat{a})Y_t^i(\hat{a})$$

At time 0, agent $\hat{a}$'s reservation utility is $-\exp\left(-rW_0^{\hat{a}}\right)$, where $W_0^{\hat{a}}$ is called the certainty equivalent reservation wealth. At time 1, since outcomes of agents' effort are realized, and provide better information about agents' abilities, firms compete for better agents, and as a consequence, the certainty equivalent wealth level of agent $\hat{a}$ who previously worked for firm k changes to $W_1^{\hat{a}}(Y_1^k(\hat{a}))$.

3.1 The Second-Period Contracting

In this section, contracting occurs twice: initial contracting at time 0, and re-contracting at time 1. That is, this dynamic contracting problem consists of the first and second-period contracting problems. We consider the second-period problem first so as to solve the overall problem recursively.

[4] In fact, one can show by using Kaman-Bucy filtering technique that in an analogous continuous-time setting with incomplete information, the optimal contract is linear as in this paper. Thus, all results in this paper can be interpreted as those in the continuous-time model.

Suppose that at time 0, agent $\hat{a}$ worked for firm k ($k \in \{S, L\}$) and at time 1, he is hired by firm i. The outcome of agent $\hat{a}$'s time-0 performance with firm k is realized at time 1 to be $Y_1^i(\hat{a})$, and firm i updates its belief on agent $\hat{a}$'s ability $\theta^{\hat{a}}$ based on the realized outcome $Y_1^i(\hat{a})$. Since both $Y_1^i(\hat{a})$ and $\theta^{\hat{a}}$ are normally distributed, and they are linearly related to each other through equation (1), $\theta^{\hat{a}}$ conditional on $Y_1^i(\hat{a})$ is normally distributed with its mean and variance given as follows:

$$E[\theta^{\hat{a}} \mid Y_1^k(\hat{a})] = m_{\theta^{\hat{a}}} + p^{\hat{a}k}\left(Y_1^k(\hat{a}) - e_0^{\hat{a}} f(K^k) - m_{\theta^{\hat{a}}} g(K^k)\right), \tag{2}$$

and

$$Var[\theta^{\hat{a}} \mid Y_1^k(j)] = \frac{\sigma_{\theta^{\hat{a}}}^2 \sigma^2 h^2(K^k)}{\sigma_{\theta^{\hat{a}}}^2 g^2(K^k) + \sigma^2 h^2(K^k)}, \tag{3}$$

where

$$p^{\hat{a}k} \equiv \frac{\sigma_{\theta^{\hat{a}}}^2 g(K^k)}{\sigma_{\theta^{\hat{a}}}^2 g^2(K^k) + \sigma^2 h^2(K^k)}. \tag{4}$$

That is, common beliefs on agent abilities are updated over time according to equation (2) with their conditional means and variances given by equations (2) and (3). This updating reduces the estimated variance of ability level from $\sigma_{\theta^{\hat{a}}}^2$ to that in equation (3). Equation (2) implies that each conditional mean is determined by a regression line between the ability level $\theta^{\hat{a}}$ and the realized outcome Y_1^k with an intercept of $m_{\theta^{\hat{a}}}$ and a slope of p^k.

As a result, the second-period contracting problem for firm i hiring agent $\hat{a}$ who worked for firm k for the first period can be stated as follows:

Problem 1 (The second-period contracting.) Choose pay contract $S_2^i(\hat{a})$ to maximize the expected profit to the shareholders in period 2 conditional on the agent's performance outcome in period 1:

$$E[Y_2^i(\hat{a}) - S_2^i(\hat{a}) \mid Y_1^k(\hat{a})], \text{ subject to}$$

(1) $Y_2^i(\hat{a}) = e_1^{\hat{a}}(i) f(K^i) + \theta^{\hat{a}} g(K^i) + h(K^i)\varepsilon_2^i$,

$Y_1^k(\hat{a}) = e_0^{\hat{a}}(k) f(K^k) + \theta^{\hat{a}} g(K^k) + h(K^k)\varepsilon_1^k$,

(2) $e_1^{\hat{a}}(i) \in \arg\max_{\hat{e}} E\left[-\exp\left\{-r\left(S_2^i(\hat{a}) - c(\hat{e})\right)\right\} \mid Y_1^k(\hat{a})\right]$

s.t. $Y_2^i(\hat{a}) = \hat{e} f(K^i) + \theta^{\hat{a}} g(K^i) + h(K^i)\varepsilon_2^i$,

$Y_1^k(\hat{a}) = e_0^{\hat{a}}(k) f(K^k) + \theta^{\hat{a}} g(K^k) + h(K^k)\varepsilon_1^k$,

(3) $E\left[-\exp\left\{-r\left(S_2^i(\hat{a}) - c(e_1^{\hat{a}}(i))\right)\right\} \mid Y_1^k(\hat{a})\right] \geq -\exp\left(-rW_1^{\hat{a}}\right)$.

The first constraint is simply the production functions for periods 1 and 2. The second constraint is the agent's effort incentive constraint conditional on the outcome in period 1, and the third constraint is the participation constraint given the agent's reservation utility in period 1. The first order condition (FOC) from the incentive constraint combined with the participation constraint implies that the second period pay schedule:

$$S_2^i(\hat{a}) = \underbrace{W_1^{\hat{a}} + c(e_1^{\hat{a}}) + \frac{r}{2}\left(\frac{c_e(e_1^{\hat{a}})}{f(K^i)}\right)^2 \left(\frac{g^2(K^i)\sigma_{\theta^{\hat{a}}}^2 \sigma^2 h^2(K^k)}{\sigma_{\theta^{\hat{a}}}^2 g^2(K^k) + \sigma^2 h^2(K^k)} + \sigma^2 h^2(K^i)\right)}_{\text{fixed compensation}} \tag{5}$$
$$+ \underbrace{\left(\frac{c_e(e_1^{\hat{a}})}{f(K^i)}\right)\left\{Y_2^i(\hat{a}) - \left(e_1^{\hat{a}} f(K^i) + E[\theta^{\hat{a}} \mid Y_1^k] g(K^i)\right)\right\}}_{\text{performance-based compensation}}.$$

Note that the sensitivity of the contract, or the sensitivity of the compensation to the realized outcome $Y_2^i(\hat{a})$, is $\frac{c_e(e_1^{\hat{a}})}{f(K^i)}$. The structure of equation (5) is well-known, consisting of two parts: fixed and performance-based compensations.

The first term of the fixed compensation is the agent's certainty equivalent reservation wealth, the second the cost of effort, and the third the compensation-risk premium. The performance-based compensation is in proportion $\frac{c_e(e_1^{\hat{a}})}{f(K^i)}$ to the unexpected outcome, $Y_2^i - E[Y_2^i \mid Y_1^k]$, which is realized minus expected outcomes. The performance-based compensation constitutes a compensation risk to the agent, on which the agent demands a risk premium. Aggrawal and Samwick (1999) show empirically that CEO pay is increasing in risk while Becker (2006) shows that CEOs with more wealth such that they are presumably less risk averse receive higher incentives.

By equation (5), the expected profit of firm i for the second period is

$$\pi^i(\hat{a}(k), W_1^{\hat{a}(k)}) := E[Y_2^i(\hat{a}) - S_2^i(\hat{a}) \mid Y_1^k(\hat{a})] = g(K^i)E[\theta^{\hat{a}} \mid Y_1^k(\hat{a})] + \Phi(K^k, K^i, \sigma_{\theta^{\hat{a}}}) - W_1^{\hat{a}},$$

where:

$$\Phi(K^k, K^i, \sigma_{\theta^{\hat{a}}}) = \max_e e(K^i)^{\gamma_f} - \frac{\kappa}{2}e^2 - \frac{r}{2}\left(\frac{\kappa e}{(K^i)^{\gamma_f}}\right)^2 \left(\frac{(K^i)^{2\gamma_g}\sigma_{\theta^{\hat{a}}}^2 \sigma^2 (K^k)^{2\gamma_h}}{\sigma_{\theta^{\hat{a}}}^2 (K^k)^{2\gamma_g} + \sigma^2 (K^k)^{2\gamma_h}} + \sigma^2 (K^i)^{2\gamma_h}\right) \tag{6}$$

Substituting the FOC with respect to effort e back into the RHS of equation (6) we have

$$\Phi(K^k,K^i,\sigma_{\theta^{\hat{a}}})=\frac{1}{2}e(K^i)^{\gamma_f}=\frac{1}{2}\frac{1}{\kappa(K^i)^{-2\gamma_f}+r\kappa^2\sigma^2\left(\frac{(K^i)^{2(\gamma_g-2\gamma_f)}\sigma^2_{\theta^{\hat{a}}}}{\sigma^2_{\theta^{\hat{a}}}(K^k)^{2(\gamma_g-\gamma_h)}+\sigma^2}+(K^i)^{2(\gamma_h-2\gamma_f)}\right)}. \tag{7}$$

3.1.1 Pay Sensitivity and Firm Size

Pay sensitivity has become an important issue in the agency literature since Jensen and Murphy (1990) argued that there is empirically a negative relationship between the sensitivity and firm size. The firm i 's problem in equation (6) enables us to relate the sensitivity to firm size.

The FOC for equation (6) also implies that for firm *i* hiring agent $\hat{a}$ who worked for firm *k,* the sensitivity of the contract for the second period to motivate the agent to exert effort is

$$\frac{\kappa\hat{e}}{(K^i)^{\gamma_f}}\equiv\beta^i(\hat{a}(k))\equiv\beta^{ki}=\frac{1}{1+r\kappa\sigma^2\left(\frac{(K^i)^{2(\gamma_g-\gamma_f)}\sigma^2_{\theta^{\hat{a}}}}{\sigma^2_{\theta^{\hat{a}}}(K^k)^{2(\gamma_g-\gamma_h)}+\sigma^2}+(K^i)^{2(\gamma_h-\gamma_f)}\right)}. \tag{8}$$

Note that the sensitivity expression, equation (8), does not directly depend on agent type $\hat{a}$, because of our assumption that the risk aversion and effort cost functions for both agents are identical, but it does depend on the distribution of agent talents, $\sigma^2_{\theta^{\hat{a}}}$, as this is updated by information derived from the initial work experience. Thus the sensitivity depends on the agent's work experience *k,* because the experience affects the volatility of the second-period outcome as the distribution of the agent ability level is updated.

Equation (8) immediately relates the sensitivity to the size of the firm as follows:

Proposition 1: *Suppose that K^k, the size of the firm for which the manager previously worked for is given. Then, holding other things constant:*

(i) *The sensitivity is inversely related to the firm size, i.e., $\partial\beta^{ki}/\partial K^i < 0$, if either the relative scale elasticities, $\gamma_h - \gamma_f \geq \gamma_f - \gamma_g > 0$ or $\gamma_g > \gamma_f > 0$ and $\gamma_h > \gamma_f > 0$.*

(ii) *The sensitivity is positively related to the firm size, i.e., $\partial\beta^{ki}/\partial K^i > 0$, if either $0 < \gamma_g - \gamma_f \leq \gamma_f - \gamma_h$ or $\gamma_f > \gamma_g > 0$ and $\gamma_f > \gamma_h > 0$.*

Proofs are given in the *Appendix*.

Proposition 1 suggests that, for example, the large firm offers a lower- (higher-) powered incentive contract than the small firm, when expected marginal effort-productivity $(K^i)^{\gamma_f}$ is sufficiently lower (higher) than expected ability-productivity $(K^i)^{\gamma_g}$ and volatility growth $(K^i)^{\gamma_h}$ over firm size. Substituting our empirically derived elasticities estimated in Table 4 below, we find, based on stock market productivity measures for the entire sample, that $\gamma_g = 0.98 > \gamma_h = 0.81 > \gamma_f = 0.38$. The same inequalities are satisfied for accounting measures of productivity and for both large and small firms. Hence condition (i) rather condition (ii) is satisfied and pay-performance sensitivity is optimally negatively related to firm size, as is shown in Table 3 below with a partial correlation coefficient of 5.3%.

This implication is in contrast with Baker and Hall (2004) who argued that the sensitivity is negatively related to the firm size because dollar volatilities of profits of large firms are higher than those of small firms. However, our Proposition 1 indicates that the relationship depends more on relative sizes of γ_f, γ_g and γ_h than it does on differences in dollar volatilities. For example, if $\gamma_f = \gamma_g = \gamma_h$, then sensitivities of both the large and small firms are identical, even though the total dollar volatility of the large firm can be considerably higher than that of the small firm.[5]

3.1.2 Career Path and Sensitivity

Equation (8) also tells us how sensitivities are affected by executives' career paths.

Proposition 2: *Suppose that agents a and b are hired for the first period by firms S and L, respectively, and agent b (a) comes from a tighter distribution than a (b). If $\gamma_g - \gamma_h \geq 0$ and $\sigma_{\theta^b} < \sigma_{\theta^a}$ (if $\gamma_g - \gamma_h \leq 0$ and $\sigma_{\theta^a} > \sigma_{\theta^b}$), then the second-period contract sensitivity for the agent who previously worked for the large firm is higher (lower) than the sensitivity for the agent who previously worked for the small firm.*

The statement comes directly from equation (8). To see this intuitively, note that equation (3) implies that if $\gamma_g - \gamma_h > 0$ and $\sigma_{\theta^b} > \sigma_{\theta^a}$, the conditional variance of the agent ability given his performance with the first-period firm is inversely (positively) related to the firm size. That is, if $\gamma_g - \gamma_h > (<)\, 0$, the informativeness of the agent's past performance about his ability increases (decreases) with the firm size. Thus, if $\gamma_g - \gamma_h > (<)\, 0$, then the firm hiring a manager coming from the large firm would have lower (higher) outcome volatility and consequently it provides its manager with a higher-powered (lower-powered) contract than the other firm hiring a manager coming from the small firm would.

[5] Note the Baker and Hall (2004) do not take account of managerial ability at all insofar as ability is implicitly assumed to be identical for all managers. Moreover, the volatility growth $h(K) = \sigma K^{\gamma_h}$ is only implicit in Baker and Hall (2004) and is not formally modeled in their article.

Since our empirical estimation set out in Table 4 below shows that $\gamma_g - \gamma_h > 0$ for all market and accounting productivity measures, irrespective of whether all CEOs are included or the two sized-based samples, the model predicts that the contract sensitivity for the manager who moves from one large firm to another, or for that that matter stays in a large firm, will have higher contract sensitivity than the manager who moves from a small firm to a large firm. There is only a relatively small sample of 125 CEO movers out of about 19,000 CEO productivity-year observations. We regress the change in $\hat{\beta}_{t-1}^{ki}$ sensitivity for the new hire relative to the incumbent on the difference between the firm size of the new hire relative to the incumbent for our sample of movers. As predicted, the sign is negative with a significant *t*-value of 1.83 at about the 6% level and an R^2 of 2.66%.

3.2 Pay and Firm Size: Labor Market Equilibrium

In this section, we examine relationships between expected executive pay and firm size over the two contracting periods. As can be inferred from the form of the salary function in equation (5), the main issue in computing the expected executive pay is to understand how the executive reservation certainty-equivalent wealth level $W_1^{\hat{a}}$ is determined. It will be seen that the wealth level can depend on labor market competition for agents which is based on each agent's ability estimated from his past performance. In the labor market, each firm assesses each agent's ability given his past performance, and makes a job offer. Then, he chooses from job offers made by the two firms. As a consequence, the agent's certainty reservation wealth is competitively determined.

For this, we model labor market competition between the two firms as follows. At time 0, agent a works for firm S, and agent b works for firm L. Then at time 1, there can be two possible cases: case (SS; LL) where agent a is rehired by firm S, and agent b is also rehired by firm L; and case (SL; LS) where agent a now works for firm L, and agent b now works for firm S.

We define the executive labor market equilibrium as follows. Each firm makes job offers to all agents on a first-come first-served basis. All job offers are structured in the form of equation (5). For example, a job offer made out to agent $\hat{a}$ by firm i is represented by a level of certainty equivalent wealth $W_0^{\hat{a}i}$ with the contract structure given in the form of equation (5). Thus, the two agents $(\hat{a},\hat{b})$ receive job offers $(W_1^{\hat{a}S},W_1^{\hat{b}S})$ from firm S and $(W_1^{\hat{a}L},W_1^{\hat{b}L})$ from firm L. If agent $\hat{a}$ takes the offer by firm *S* before agent $\hat{b}$ does, then agent $\hat{a}$ enjoys a certainty equivalent wealth level of $W_1^{\hat{a}S}$, and agent $\hat{b}$ is hired by firm *L*.

It should be clear that each firm would like to hire an agent who would produce an expected profit to the firm at least as great as the other agent would. However, each firm's decision can also affect/be affected by the other firm's decision. We examine the following type of executive labor market equilibrium.

Definition 1: *The executive job market is in equilibrium with agents* $\hat{a}$ *and* $\hat{b}$ *choosing to work for firms S and L, respectively, if job offers* $(W_1^{\hat{a}i},W_1^{\hat{b}i})$ *to agents* $(\hat{a},\hat{b})$ *by firm* i*, for* $i = S$ *and* L*, satisfy the following properties.*

(i) (Profit maximization.)

$$W_1^{\hat{a}S} \in \arg\max_W \pi^S(\hat{a},W) \ s.t. W \geq W_1^{\hat{a}L}, \ and \ \pi^S(\hat{a},W_1^{\hat{a}S}) \geq \max_W \pi^S(\hat{b},W) \text{ s.t. } W \geq W_1^{\hat{b}L}.$$

$$W_1^{\hat{b}L} \in \arg\max_W \pi^L(\hat{b},W) \ s.t. W \geq W_1^{\hat{b}S}, \ and \ \pi^L(\hat{b},W_1^{\hat{b}L}) \geq \max_W \pi^L(\hat{a},W) \text{ s.t. } W \geq W_1^{\hat{a}S}.$$

(ii) (Expected zero profit condition for the small firm.)

$\pi^S(\hat{a}, W_1^{\hat{a}S}) = \pi^S(\hat{b}, W_1^{\hat{b}L}) = 0.$

Condition (i) implies that each firm chooses an agent to maximize its expected profit. When the small and large firms hire agents $\hat{a}$ and $\hat{b}$, respectively, condition (i) implies that the small firm makes an offer to agent $\hat{a}$ by matching the offer by the large firm to the same agent such that $W_1^{\hat{a}S} = W_1^{\hat{a}L}$ and $\pi^S(\hat{a}, W_1^{\hat{a}S}) \geq \pi^S(\hat{b}, W_1^{\hat{b}L})$, and similarly that we have $W_1^{\hat{b}L} = W_1^{\hat{b}S}$ and $\pi^L(\hat{b}, W_1^{\hat{b}L}) \geq \pi^L(\hat{a}, W_1^{\hat{a}S})$. Condition (ii) suggests that agents' reservation certainty-equivalent wealth levels are determined by their job opportunities with the small firm, and that the small firm's expected profit is always driven to zero (perhaps by job/product market competition).

The next proposition sheds some light on equilibrium hiring decisions in the second period. First, let us define:

$$\begin{aligned}
&A(K^S, K^L; \sigma_{\theta^a}, \sigma_{\theta^b})\left(g(K^L) - g(K^S)\right) \\
&= \left\{\Phi(K^S, K^S, \sigma_{\theta^a}) - \Phi(K^L, K^S, \sigma_{\theta^b}) - \Phi(K^S, K^L, \sigma_{\theta^a}) + \Phi(K^L, K^L, \sigma_{\theta^b})\right\}.
\end{aligned}$$

Then $A(K^S, K^L; \sigma_{\theta^a}, \sigma_{\theta^b})\left(g(K^L) - g(K^S)\right)$ measures the comparative advantage of agent *b* over agent *a* in terms of the marginal effort contribution to the large firm's expected profit over that of the small firm. If $A(K^S, K^L; \sigma_{\theta^a}, \sigma_{\theta^b})\left(g(K^L) - g(K^S)\right)$ is positive, agent *b's* marginal effort-contribution to the expected profit of the large firm is relatively larger than that of agent *a*.

Proposition 3: *Suppose that agents a and b are hired for the first period by firms S and L, respectively. If $E[\theta^a - \theta^b \mid Y_1] \leq (>) A(K^S, K^L; \sigma_{\theta^a}, \sigma_{\theta^b})$, then in equilibrium, agents a (b) and b(a) are rehired (hired) for the second period by firms, S and L,* *respectively, with their second-period reservation certainty equivalent wealth levels given by* $W_1^a = g(K^S)E[\theta^a \mid Y_1^S(a)] + \Phi(K^S, K^S, \sigma_{\theta^a})$, and $W_1^b = g(K^S)E[\theta^b \mid Y_1^L(b)] + \Phi(K^L, K^S, \sigma_{\theta^b})$.

The small-firm manager moves to the large firm in the second period if and only if his expected ability conditional on his first period performance, $E[\theta^a \mid Y_1]$, turns out to be sufficiently large, such that $E[\theta^a \mid Y_1] > E[\theta^b \mid Y_1] + A(K^S, K^L, \sigma_{\theta^a}, \sigma_{\theta^b})$. In this sense, one may view the function *A* as a measure of executive job mobility: a high *A* means a low probability for small-firm managers to move to large firms. In other words, $E[\theta^b - \theta^a \mid Y_1] + A(K^S, K^L, \sigma_{\theta^a}, \sigma_{\theta^b})$ measures the comparative advantage of agent *b* over agent *a* in terms of contributions by both ability and effort to the expected profit of the large firm. Thus, the agent worked for the small firm can be hired by the large firm only when his expected ability level is large enough to get over the large firm manager's comparative advantage.

Unlike the matching literature in which more talented managers are automatically allocated to larger firms, agents hiring/moving decisions in this paper are based not only on perceived/expected ability levels but also on the volatilities of their ability levels due to uncertainty as to what their ability really is, as outcomes of agents' effort depend on both ability levels and their distributions.

3.3 The First-Period Contracting Problem

Agent $\hat{a}$'s, $\hat{a} \in \{a, b\}$, effort choice decision for the first period can be affected by his job market prospects for the second period. Thus, the first-period principal's problem can be stated as follows.

Problem 2: Choose an initial-period pay schedule $S_1^i(\hat{a})$ to maximize expected first-period shareholder profit:[6]

$$E[Y_1^i(\hat{a}) - S_1^i(\hat{a})] \text{, subject to}$$

(1) $Y_1^i(\hat{a}) = e_0^{\hat{a}} f(K^i) + \theta^{\hat{a}} g(K^i) + h(K^i)\varepsilon_1^i$,

(2) $e_1^{\hat{a}} \in \arg\max_{\hat{e}} E\left[-\exp\left\{-r\left(S_1^i(\hat{a}) - c(\hat{e}) + W_1^{\hat{a}}\right)\right\}\right]$,

s.t. $\quad Y_1^i(\hat{a}) = \hat{e} f(K^i) + \theta^{\hat{a}} g(K^i) + h(K^i)\varepsilon_1^i$,

(3) $E\left[-\exp\left\{-r\left(S_1^i(j) - c(e_0^{\hat{a}}) + W_1^{\hat{a}}\right)\right\}\right] \geq -\exp\left(-rW_0^{\hat{a}}\right)$.

The main difference between Problems 1 and 2 is that in Problem 1, the agent's first-period wealth consists of not only $S_1^i(\hat{a}) - c(e_0^{\hat{a}})$, direct compensation from the firm net of effort cost, but also $W_1^{\hat{a}}$, the certainty equivalent wealth the agent can expect from the second period contracting. Young agents have career concerns that impact on their choice of their first managerial position.

---

[6] Note that at time 0, shareholders' expected profit of the large firm for both periods 1, and 2 is $E[Y_1^L(b) - S_1^L(b) + \pi_1^L(b, W_1^{bL})\chi_A + \pi_1^L(a, W_1^{aL})(1-\chi_A)]$, where $\chi_A(\omega) = 1$ for $\omega \in \left\{\hat{\omega} \in \Omega \mid E[\theta^a - \theta^b \mid Y_1](\hat{\omega}) \leq A(K^S, K^L)\right\}$, and $\chi_A(\omega) = 0$, otherwise. Here, Ω is a complete description of all uncertainties in this paper. However, since it can be shown that the second-period profit $E[\pi_1^L(b, W_1^{bL})\chi_A + \pi_1^L(a, W_1^{aL})(1-\chi_A)]$ in equilibrium is independent of the agent's time-0 effort e_0^b, shareholders' optimal decision for the agent's time-0 effort can be computed ignoring the expected second-period profit. Thus, as far as optimal effort decisions are concerned, shareholders are only concerned with maximizing the expected profit from the first period. That is, optimal effort levels maximize $E[Y_1^L(b) - S_1^L(b)]$ subject to appropriate constraints, as stated in Problem 2.

Without loss of generality, we again assume optimal contracts are linear such that $S_1^i(\hat{a}) = \alpha^i + \beta^i Y_1^i(\hat{a})$ for agent $\hat{a}$ working for firm i.

Proposition 4: *Let firm* $k \in \{\text{S,L}\}$ *hires agent* $\hat{a} \in \{a,b\}$ *at time zero. Then fixed and incentive parameters* $(\alpha^{\hat{a}k}, \beta^{\hat{a}k})$ *for the optimal contract are given as follows:*

$$\begin{aligned}\alpha^{\hat{a}k} &= W_0^{\hat{a}} - \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}) + c(\hat{e}^{\hat{a}}) - g(K^S)m_{\theta^{\hat{a}}} - \beta^{\hat{a}k}(\hat{e}^{\hat{a}} f(K^k) + m_{\theta^{\hat{a}}} g(K^k)) \\ &\quad + \frac{r}{2}(\beta^{\hat{a}k} + g(K^S)p^{\hat{a}k})^2 \left(\sigma_{\theta^{\hat{a}}}^2 g^2(K^k) + \sigma^2 h^2(K^k)\right).\end{aligned} \tag{9}$$

$$\begin{aligned}\beta^{\hat{a}k} &= \frac{c'(e)}{f^k} - g(K^S)p^{\hat{a}k} \\ &= \frac{1}{1 + r\kappa\left(K^k\right)^{-2\gamma_f}\left(\sigma_{\theta^{\hat{a}}}^2 (K^k)^{2\gamma_g} + \sigma^2 (K^k)^{2\gamma_h}\right)} - \frac{\sigma_{\theta^{\hat{a}}}^2 (K^k K^S)^{\gamma_g}}{\sigma_{\theta^{\hat{a}}}^2 (K^k)^{2\gamma_g} + \sigma^2 (K^k)^{2\gamma_h}}\end{aligned} \tag{10}$$

Thus, the expected compensation to the agent is

$$E[S^k(Y_1^k)] = W_0^{\hat{a}} - \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}) - g(K^S)m_{\theta^{\hat{a}}} + \Psi(K^k, \sigma_{\theta^{\hat{a}}}) \tag{11}$$

and the expected profit of firm *k* hiring agent $\hat{a}$ is

$$E[Y_1^k - S^k] = m_{\theta^{\hat{a}}}\left(g(K^k) + g(K^S)\right) - W_0^{\hat{a}} + \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}) + \Psi\left(K^k, \sigma_{\theta^{\hat{a}}}\right), \tag{12}$$

where initial certainty equivalent wealth:

$$W_0^{\hat{a}} = 2m_{\theta^{\hat{a}}} g(K^S) + \Psi(K^S, \sigma_{\theta^{\hat{a}}}) + \Phi(K^S, K^S, \sigma_{\theta^{\hat{a}}}),$$

and $$\Psi\left(K^k, \sigma_{\theta^{\hat{a}}}\right) = \frac{1}{2}\frac{1}{\kappa(K^k)^{-2\gamma_f} + r\kappa^2\left(K^k\right)^{-4\gamma_f}\left(\sigma_{\theta^{\hat{a}}}^2 (K^k)^{2\gamma_g} + \sigma^2 (K^k)^{2\gamma_h}\right)}.$$

Recall that in the second period, there is no future career concern problems and the contract sensitivity is the marginal cost of effort per marginal expected output, $\frac{c'(e)}{f^k}$.

However, Proposition 4 also implies that, in the first-period contracting, the sensitivity is adjusted for the agent's career concern by product of the scale term for ability in the small firm and the updating talent regression slope coefficient, i.e., $g(K^S)p^k$. That is, in the first period, the contract sensitivity does not have to be equal to the marginal cost of effort per marginal expected output, because the agent has already built-in (implicit) incentives to work even without an explicit incentive contract. This kind of adjustment is well-known. See Gibbons and Murphy (1992).

Proposition 4 also implies that the expected pay differential between large and small firms is made up of the following two components:

$$E[S_1^L]-E[S_1^S]=\underbrace{W_0^b-W_0^a-E\left[W_1^b-W_1^a\right]}_{\substack{\text{current period reservation}\\ \text{wealth differential}}}+\underbrace{\Psi(K^L,\sigma_{\theta^b})-\Psi(K^S,\sigma_{\theta^a})}_{\text{effort production differential}}. \tag{13}$$

Note that the current period certainty equivalent reservation wealth is $W_0^{\hat{a}}-E\left[W_1^{\hat{a}}\right]$, that is, the certainty equivalent wealth for the agent's lifetime career (over the two periods) minus the expected future certainty equivalent wealth. Representing the difference in negotiating power between two agents in the labor market, the current period reservation certainty equivalent differential has the following structure:

$$\begin{aligned} W_0^b-W_0^a-E\left[W_1^b-W_1^a\right]&=\underbrace{(m_{\theta^b}-m_{\theta^a})g(K^S)}_{\text{ability differential}}+\underbrace{\Psi(K^S,\sigma_{\theta^b})-\Psi(K^S,\sigma_{\theta^a})}_{\substack{\text{effort production differential}\\ \text{to the small firm}}}\\ &\quad+\underbrace{\Phi(K^S,K^S,\sigma_{\theta^b})-\Phi(K^L,K^S,\sigma_{\theta^b})}_{\substack{\text{compensation for disadvantages}\\ \text{in furure job market}}} \end{aligned} \tag{14}$$

This structure tells us that sources of negotiating power lie in expected ability, the volatility of ability, and disadvantages/advantages of working for the large firm in the future job market.

To sum up, the sources of the difference in pay size between executives of large and small firms are (1) the ability production differential had each agent worked for the small firm (reference firm), (2) the effort production differential had each agent worker for the small firm, (3) discount for advantages the large-firm executive may experience in future executive labor markets, and (4) the actual effort production differential between the large and small firm. By contrast, in Edmans, Gabaix and Landier (2007) pay differentials are entirely determined by talent/ability differentials. In this paper, the talent differential is just one of many sources of the difference in pay size. In particular, if $\gamma_g < \gamma_h$, disadvantages suggested in the third source can occur, as the volatility of updated expectation of executive ability level after the first period will be higher for the large-firm executive, because dollar-return from production is more volatile for the large firm than it is for the small firm. That is, the large-firm profit outcome in the initial period provides a weaker signal as to agent ability than for the small-firm agent due to the volatility difference. However, if $\gamma_g > \gamma_h$, which is what we observe empirically, then working for the large firm can help send a less noisy signal about his ability to the future job market.

The second differential can increase the pay for the agent working for the large firm if the volatility of the agent's ability is lower than that of the other agent working for the small firm (reference firm), simply because the low volatility can improve effort incentives. However, note that this differential in fact has nothing to do actual improvement of effort incentives, but it is simply added as a consequence of labor market competition in which the small firm bids for the agent with low volatility in ability. On the other hand, the fourth differential is compensation for actual improvement of effort incentives.

Now, we examine effects of firm size on managerial salaries:

Proposition 5: *At time 0, agents a and b are hired by firms S and L, respectively, if and only if*

$$\Delta\pi_0 := (m_{\theta^b} - m_{\theta^a})\left(g(K^L) - g(K^S)\right) + \int_{K^S}^{K^L}\int_{\sigma_{\theta^a}}^{\sigma_{\theta^b}} \Phi_{K^k\sigma_\theta}(K^k, K^S, \sigma_\theta)d\sigma_\theta dK^k$$
$$+\int_{K^S}^{K^L}\int_{\sigma_{\theta^a}}^{\sigma_{\theta^b}} \Psi_{K\sigma_\theta}\left(K, \sigma_\theta\right)d\sigma_\theta dK \geq 0.$$

Remark: The necessary and sufficient condition for Proposition can be alternatively stated as

$$\begin{aligned}\Delta\pi_0 = {} & \underbrace{(m_{\theta^b} - m_{\theta^a})\left(g(K^L) - g(K^S)\right)}_{\text{talent contribution differential}} \\ & + \underbrace{\Phi(K^S, K^S, \sigma_{\theta^a}) - \Phi(K^S, K^S, \sigma_{\theta^b}) + \Psi\left(K^S, \sigma_{\theta^a}\right) - \Psi\left(K^S, \sigma_{\theta^b}\right)}_{\text{reservation CEO wealth differential}} \\ & + \underbrace{\Phi(K^L, K^S, \sigma_{\theta^b}) - \Phi(K^L, K^S, \sigma_{\theta^a})}_{\text{advantage of working for the large firm in future job market}} \\ & + \underbrace{\Psi\left(K^L, \sigma_{\theta^b}\right) - \Psi\left(K^L, \sigma_{\theta^a}\right)}_{\text{current period effort production differential}} \\ & \geq 0.\end{aligned}$$

Note that this condition can be satisfied if $m_{\theta^b} \geq m_{\theta^a}$, $(\sigma_{\theta^b} - \sigma_{\theta^a})\Psi_{K\sigma_\theta} \geq 0$ and $(\sigma_{\theta^b} - \sigma_{\theta^a})\Phi_{K^k\sigma_\theta} \geq 0$, and also that $sign(\Phi_{K^k\sigma_\theta}) = sign(\gamma_g - \gamma_h)$; and if $\gamma_g \leq 2\gamma_f$ and $2\gamma_h - 2\gamma_f - \gamma_g \leq 0$, then $\Psi_{K^k\sigma_\theta} \leq 0$.

Corollary 1: *If $m_{\theta^b} \geq m_{\theta^a}$, and $\sigma_{\theta^b} = \sigma_{\theta^a}$, then, at time 0, agents a and b are hired by firms S and L, respectively.*

Proposition 5 provides the necessary and sufficient condition under which agent *b* is hired by the large firm. When $\Delta\pi_0 > 0$, the large firm prefers agent *b* to *a*. In the following example, we use empirical data reported in Tables 1 and 4 to see if current CEOs hired by large firms may be justified based on the average ability of such managers.

Example 1: Suppose $K^S = 1{,}023$, $K^L = 25{,}132$, $r = 0.05$, $\kappa = 0.01$, $m_{\theta^a} = 1.2655$, $m_{\theta^b} = 1.2918$, $\gamma_f = 0.3812$, $\gamma_g = 0.9848$, $\gamma_h = 0.8131$, $\sigma = 0.7519$, $\sigma_{\theta^a} = 0.8234$, and $\sigma_{\theta^b} = 0.4280$. Then $\Delta\pi_0 = 1{,}418.04$, $E[S_1^S] = 8{,}834.07$, and $E[S_1^L] = 9{,}773.54$.

In this example, $\Delta\pi_0 > 0$ which implies that current CEOs of large firms might have been hired because the CEOs of large firms were expected to earn higher net profits to the large firms than CEOs of small firms were. This example is consistent with the popular intuition that CEOs of large firms have on average greater expected talent levels with lower talent volatilities.

However, Proposition 5 and Corollary 1 also allude to the possibility that large firms may not necessarily choose more talented agents, unless the volatilities are the same. Here, we provide a numerical example in which the large firm hires the manager with lower expected talent.

Example 2: Suppose $K^S = 1{,}023$, $K^L = 25{,}132$, $r = 0.05$, $\kappa = 0.01$, $m_{\theta^a} = 1.2918$, $m_{\theta^b} = 1.2655$, $\gamma_f = 0.3812$, $\gamma_g = 0.9848$, $\gamma_h = 0.8131$, $\sigma = 0.7519$, $\sigma_{\theta^a} = 0.8234$, and $\sigma_{\theta^b} = 0.4280$. Then $\Delta\pi_0 = 333.21$, $E[S_1^S] = 8{,}858.29$, and $E[S_1^L] = 9{,}749.32$.

In Example 2, agent *b* is hired by the large firm at time 0, although his expected talent level is lower than that of the other agent. Note however that agent *b*'s talent volatility is lower than that of the other agent, which helps improve work incentives. In this case, agent effort contribution can affect the outcome more than the expected talent differential can, and thus the large firm is more concerned with improving incentives than the talent differential, and consequently hires agent *b*.

The next proposition provides some sufficient conditions under which the agent working for the large firm is expected to be more highly paid than the other agent working for the small firm.

Proposition 6: *Suppose that agents a and b are hired by the small and large firms, respectively. If* $m_{\theta^b} \geq m_{\theta^a}$, $\sigma_{\theta^b} \leq \sigma_{\theta^a}$, $\gamma_g \leq \gamma_h$, *and* $\max[\gamma_g, \gamma_h] < 2\gamma_f$, *and then the first-period expected pay of the large firm manager is higher than that of the small firm.*

Evaluating the inequalities included in Proposition 5 utilizing the estimated elasticity values reported in Table 4 below for all firm sizes and measures, neither inequality is satisfied as $\max[\gamma_g, \gamma_h] > 2\gamma_f$ and $\gamma_g > \gamma_h$. Hence in the first period of the manager's career, we cannot guarantee that the larger firm manager will be paid more in equilibrium than the smaller firm manager.

4. Empirical Implementation

4.1 Model Specification

For empirical estimation purposes, we use equation (10) in Proposition 4 to express the stochastic production function (1) in terms of (at least theoretically) observables as follows:

$$Y_t^i(\hat{a}) \equiv \hat{Y}_t^{i\hat{a}} = \frac{1}{\kappa}\left(\beta_{t-1}^{i\hat{a}} + \left(K_{t-1}^S\right)^{\gamma_g} p_{t-1}^{\hat{a}i}\right) \times \left(K_{t-1}^{i\hat{a}}\right)^{2\gamma_f} + \hat{\theta}^{\hat{a}}(K_{t-1}^{i\hat{a}})^{\gamma_g} + \sigma \times (K_{t-1}^{i\hat{a}})^{\gamma_h} \times \varepsilon_t^i , \quad (15)$$

where $p_{t-1}^{\hat{a}i} = \dfrac{\sigma_{\theta^{\hat{a}}}^2 (\hat{K}_{t-1}^i)^{\gamma_g}}{\sigma_{\theta^{\hat{a}}}^2 (\hat{K}_{t-1}^i)^{2\gamma_g} + \sigma^2 (\hat{K}_{t-1}^i)^{2\gamma_h}}$, and $\hat{\theta}_t^{\hat{a}} \equiv E[\theta_t^{\hat{a}} \mid Y_{t-1}^k]$ is the expected conditional talent. Note that unobservable effort has been substituted out of the equation and replaced by the incentive contract following the lead of Baker and Hall (2004).

The comprehensive end-of-period wealth measure, $\hat{Y}_t^{i\hat{a}}$, is generated by the stochastic CEO production process with the CEO combining his effort imputable from his opening incentive contract and his talent with the available capital he has to work with given by the opening total value of firm assets, K_{t-1}. Such a comprehensive wealth approach to measuring CEO performance is both recommended and utilized by Baker and Hall (2004) and Gabaix and Landier (2008) on the grounds that the actions of the CEO this period have implications for shareholder and debtholder wealth well into the future, not in just the increment to claimant wealth in the current period.[7]

Observed output $\hat{Y}_t^{i\hat{a}}$ is calculated as firm claimant (shareholder plus debt holder) wealth at the end of period *t*. We separately analyze two sets of wealth measures: the first based on market values and the second on accounting values. We begin with the market value method. This is made up of three components. The first component is the total value of the firm's assets at the end of period *t* funded by both equity and debt, K_t. This is computed as the book value of total assets (item 6 from Compustat) plus the market value of total equity (item 199*(items 25+40)) less the book value of equity (item 60) less deferred taxes (item 74). To this is added the second component, namely the net cash flow to equity holders. This is made up cash distributed as dividends (item 127) plus the net value of shares repurchased (item 115 less item 108). Finally, the third component is added, namely the cash distributed to debt-holders. It takes the form of interest paid (item 15) plus long-term net debt reduction (item 114 less item 111) less the change (increase) in current debt (item 301). The accounting value method is very similar to the market value method except that now

[7] Note that these authors use essentially the total value of assets at the period end, K_t, as their wealth measure without taking into account net cash distributions to claimants.

the total value of assets is simply the book value given by Compustat Item 6. The second and third components are the same. In the absence of new net equity or debt issues, earnings generated by the manager are either retained and thus included in the closing value of assets or are paid out to claimants and thus still add to end of period wealth.

With respect to our market measure and in keeping with the regression analysis of Gabaix and Landier (2008, Table I), we use the opening market value of total assets, as described above as a component of the end of period wealth measure except lagged one period, as the best size proxy for the capital stock measure that is most associated with CEO pay, rather than income or sales that Gabaix and Landier (2008) show are inferior in their ability to explain CEO pay. Our alternate accounting measure is simply the opening *book* value of total assets.

To provide the estimated sensitivity for each year of the executive's career, $\hat{\beta}_{t-1}^{i\hat{a}}$, we use the opening sum of the executive's shareholding, restricted stock and the share-equivalent of the executive's option holdings relative to total shareholdings estimated from the Black-Scholes Delta formula (modified to include dividends).[8] The Black-Scholes values of option holdings are not provided in ExecuComp. They were computed using two different methods. Using the first method an inventory of option holdings was constructed for each CEO based on the ExecuComp data for newly issued options with a given strike price and expiry date. All share prices, shares on issue and stock split data was obtained from ExecuComp for consistency purposes.[9] A

[8] For the first year that the CEO appears in the database the sensitivity for that year is employed in lieu of the opening value since the opening value is not available.

[9] For many stocks inconsistencies arose between ExecuComp and CRSP which made it necessary to use the one data source for this purpose.

four-year escrow period was assumed with one-quarter of the options coming out of escrow each year and is described more fully in Garvey and Swan (2002). Only options most in the money were exercised according to data supplied by ExecuComp.

In addition to this inventory method, a simpler method described by Core and Guay (2002) was also computed and the two sets of results compared. It was found that the two sets of results were quite comparable and the more comprehensive inventory method was chosen in preference. In converting option holdings into share equivalents, attention was paid to the dilutionary effects of option issuance on shares outstanding. ExecuComp has data explicitly on this up until 1994 and was estimated for subsequent years. Perhaps the most significant component of estimated incentive values is shares held privately by the CEO. These are sourced from ExecuComp, either as a percentage of shares outstanding or as shares held. Great care was taken to ensure consistency in these computations using just ExecuComp and Compustat data as share numbers outstanding from CRSP were not always consistent and to remove cases where there were obvious transcription errors in either the share or option data or missing observations.

The manager's effort level implied from the opening incentive contract value, $\beta_{t-1}^{\hat{a}}$, together with his individual (unobservable as such) conditional talent factor, $\hat{\theta}_t^{\hat{a}}$, and stochastic volatility factor, is applied to the opening value of total assets under management given by K_{t-1} to generate the specified end of fiscal year wealth belonging to all claimants.

The CEO tenure with a particular firm is assumed to have a minimum length of one completed financial year and continue until the CEO resigns, retires or dies. Hence an observation on a CEO year's productivity performance is only included for completed years. We adopt as our unit of account each CEO-year but also compute the number of individual years of tenure with the *i*th firm. The length of an individual tenure, $n^{i\hat{a}}$, varies and is captured as an explanatory variable in the individual pay and CEO income regressions. The superscript $i\,\hat{a}$ refer to the value for each annual observation of the performance of the *i*th firm and the subscript *t*-1 to the beginning of fiscal year opening value.

We provide direct estimates of the parameters of the non-linear expression (15) in Table 4 below. A problem with the equation is that neither the individual talent factors, $\theta_t^{\hat{a}}$, themselves, nor the conditional expectations, $\hat{\theta}_t^{\hat{a}}$, are directly observable. To overcome this problem we first estimate equation (15) by focusing on the second-period problem (ignoring the first-period problem with a career concern). It is estimated as two separate components. The first component of equation (15) is estimated using non-linear least squares as:

$$\hat{Y}_t^{i\hat{a}} = \frac{\hat{\beta}_{t-1}^{i\hat{a}}}{\kappa}\left(\hat{K}_{t-1}^{i\hat{a}}\right)^{2\gamma_f} + \bar{\hat{\theta}}_t^{\hat{a}}(\hat{K}_{t-1}^{i\hat{a}})^{\gamma_g} + Controls + \xi_t\,, \tag{16}$$

where regression coefficient, $\bar{\hat{\theta}}_t^{\hat{a}}$, is the estimated mean conditional talent factor taken over all CEO-years in the sample. Note that all the other terms in equation (15) drop out as the random term ε_t^i is zero in expectation and the career concern (slope) term p^k from equation (4) is set to zero. Controls consist of both two digit Industry Dummies and the length of experience with the firm prior to the CEO appointment if an internal appointment, and ξ_t is the *iid* error term. Year dummies were deliberately excluded to investigate the model's capacity to explain real CEO pay rises over the sample period. However, only one of the industry control dummies was statistically significant in the non-linear least squares estimates but nonetheless generated large coefficients that added to noise and prevented conversion of the non-linear estimation. Hence the values of all industry dummies were set to zero. Note that the parameters of the dollar volatility term in equation (15), σ and γ_h, need to be estimated separately as the term ε_t^i in dollar volatility $\sigma\left(K_{t-1}^{ij}\right)^{\gamma_h}\varepsilon_t^i$, is a standard normal random variable with mean zero. These are estimated via equation (18) below.

Our sample consists of 18,835 career years of CEOs *not* from a regulated industry or an industry with an unusual capital structure. The two-digit codes excluded are 22 (utilities), 52 (finance and insurance), 55 (management of companies and enterprises) or exceeding 90 (public administration). Included CEO years have appeared in S&Ps ExecuComp over the period 1992-2006 with no missing observations and a minimum tenure of a full financial year. ExecuComp includes firms over these periods that have appeared within the top 1500 S&P firms.

All dollar amounts including the value of assets, the firm's total market and accounting income and the CEOs total pay are converted to constant dollars of 2006 using the CPI.

The second (accounting) measure is identical to the first except that the book value of total assets replaces the market value of total assets. Distributions in the form of dividend and interest payments and new cash injections remain as before. Both performance measures are deflated by the estimated average pay-performance sensitivity and then the natural logarithm taken to obtain the dependent variable in regression equation (A4) in the Appendix that is used to obtain starting values for the Non-Linear Least Squares estimation.

The sample of CEO yearly observations is ranked by size of opening total assets expressed in 2006 dollar values based on the CPI, $K_{t-1}^{i\hat{a}}$, and is split into two equal halves by number of observations, representing large and small firms separately. The sample utilizing market values is also split into CEO years with a positive contribution ($K_t^{i\hat{a}}$ + Net Cash $\text{Dist}_t - K_{t-1}^{i\hat{a}} > 0$) and with a negative contribution.

All productivity and associated data for CEO years based on market performance values are summarized in Table 1 and for accounting values, in Table 2. For space reasons the accounting estimates for the large and small samples separately are not presented. The mean terminal (end of period) wealth level is \$14,124m for the observations on CEO performance for the entire sample in Table 1 and is greater than the mean opening value of assets, \$13,077. This wealth measure nearly doubles for the sample of large firms to \$27,070m and for small firms, only \$1,179m. Hence size is highly skewed. Unsurprisingly, the mean $\hat{\beta}_{t-1}^{i\hat{a}}$ sensitivity coefficient at 0.0381 for small firms is about double that for large firms, 0.0173. Total mean annual pay from ExecuComp for large companies at \$7.8m in 2006 prices is many times higher than for small companies at \$2.3m. Inclusive of income from shares owned by CEOs, mean total CEO income is \$33m for large companies and \$7.9m for small. Note that only 63% of CEO years display positive market performance, generating end of year wealth well in excess of the opening asset value and it is the reverse for the negative performers.

The partial correlation coefficients (in levels) for the entire market-based sample are provided in Table 3. Perhaps the most striking feature of these correlations is the unsurprising 98% correlation between opening and end of year wealth (inclusive of distributions and net of new debt and equity issues). This illustrates the ubiquity of size and largely explains the exceptionally good model fit in Table 4 below. Striking also is the 66% correlation between size and dollar volatility.

<< INSERT TABLES 1 to 3 ABOUT HERE>>

The Non-Linear Least Squares CEO productivity regression results in real terms are summarized in Table 4 for the productivity measures based on market and book values and those displaying either positive or negative income, all based on the entire sample, and the large and small sub-samples that have been estimated separately using market productivity only. Starting values for the coefficients were obtained via the estimation of equations (A4) and (A6) in the Appendix. For the positive and negative performer breakdowns the same set of coefficients based on the entire sample have been employed. For the full sample, the results in column 1 for the market-based measure based on explicit incentives only (equation (16)), indicate that the estimated shadow price of effort represented by Kappa at 0.35 is far lower than for large firms, when estimated separately for the large-firm sample, with a value of 2.5. While this in itself is not surprising as one would expect large firms to be harder and more costly to manage, the magnitude of the difference quite large and should indicate a considerably lower shadow price of effort in smaller companies if we allow shadow prices and technology to differ between the two classes or organization. As far as we are aware, this is the first time that this coefficient has been estimated as Baker and Hall (2004) assume a value of one. Unfortunately, our Kappa estimate for small firms, also estimated separately, is very large and clearly anomalous.

<<INSERT TABLE 4 ABOUT HERE>>

For the entire sample using market performance measures, CEO effort productivity increases by 38% for each doubling of firm size (total assets under management) while ability productivity, which can be either positive or negative, increases by a much higher 98%. Thus there are approximately constant returns to scale in talent and this is true for all samples. Large firms have a higher effort-scale sensitivity of 47% when this sample is estimated separately but all of the scale elasticity estimates based either on market or accounting values are in agreement.

The estimated mean CEO conditional ability level, $\bar{\hat{\theta}}_t^{\hat{a}}$, for the market measure and full sample is 1.29, which is slightly greater than the mean of the predicted values of 1.28 found by treating the estimating equation (16) as an identity. The non-linear nature of the estimating equation ensures that the estimated and simulated means found by setting the regression residuals to zero will differ at least slightly. The mean estimated ability level for large firms is surprisingly low at 1.21, rising enormously to 2.74 for small firms. This higher mean talent level for the small-firm sample arises because of differences in production function coefficients and is thus not a like-with-like comparison. While are findings are surprising given the theoretical predictions of Rosen (1982) that the largest firms would necessarily hire the most able managers, we believe ours to be the first estimates for CEOs that do not rely on extreme value theory that automatically assigns the most talented managers to the largest firms. All the estimated coefficients in Table 4 for all firms and large firms are significant at the 1% level irrespective of the use of market or accounting measures.

The slopes of the actual and predicted terminal wealth levels were estimated. They indicate that predicted values are relatively unbiased with close to a 45 degree slope. The R2 for the entire sample is high at 99%, falling to only 53% for small firms.

Computed from the residuals of the estimating equation, conditional talent prediction estimates are constructed for every CEO year and summarized in Table 4. The means and standard deviations are reported for the large and small sub-samples separately along with the entire sample. The pairs of “All Firms”, “Positive Wealth Gain” and

Negative Wealth Gain" columns all utilize coefficients estimated for the entire sample whereas the remainder is confined to either the large or small firm sample. The predicted conditional mean of talent using the all firm market measure in the first column is 1.28 with a standard deviation of ability of 0.66. When computed separately for the large-firm sample it is 1.27 with a standard deviation of 0.43, rising to 1.29 for the small firm sample with a much higher standard deviation of talent of 0.82, or almost double. Hence the risk-adjusted (Sharpe) ratio is 2.95 for large-firm CEOs and 1.57 for small-firm CEOs, that is, almost double.

Hence, contrary to the Gabaix and Landier (2008) estimates that found negligible differences in ability levels from the CEO in their median company, number 250, in size and number one, we find a remarkable diversity in CEO talent as measured by *ex post* performance. These findings are consistent with our model in which the CEOs own *ex ante* ability may be unknown even to himself and where in the marketplace for CEOs it is possible that more capable managers are priced out of the market within the group of large companies. Within the context of our model, the tighter distribution from which CEOs of large companies are drawn is consistent with the far greater predictability of performance for large-company CEOs and their far higher pay. Our results also do not appear to be supportive of Hermalin's (2005) model in which it is advantageous to recruit CEOs of less precisely known talent as pay does not need to be raised to compensate for CEO risk aversion, unlike in our model.

In column 2 of Table 4 we provide estimates inclusive of period 2 career concerns relevant for agents that are not yet at the end of their careers. The slope term $p^{\hat{a}i}$ cannot be computed for each CEO as we do not know the distribution of talent prior to estimation. However, the average value can be estimated as follows: Assuming that CEOs who are rehired in smaller firms from their current firm move to an average small firm size of $\hat{K}_{t-1}^{S}$ = \$1,023m in equation (15), then it yields a slightly modified version of equation (16) above given by:

$$\hat{Y}_t^{i\hat{a}} = \frac{1}{\kappa}\left(\beta_{t-1}^{i\hat{a}} + \left(K_{t-1}^{S}\right)^{\gamma_g} \overline{p}\right) \times \left(K_{t-1}^{i\hat{a}}\right)^{2\gamma_f} + \overline{\hat{\theta}}^{\hat{a}} (K_{t-1}^{i\hat{a}})^{\gamma_g} + Controls + \xi_t, \quad (17)$$

that crudely takes into account career concerns by having a uniform talent updating slope term $p_{t-1}^{\hat{a}i} \equiv \overline{p}$. This term, according to theory, should be CEO-specific and observable by the board of the hiring firm but is not observable by the econometrician. The average term we estimate converges to $\overline{p}$ = 0.001625. Consequently, the estimated impact of career concerns, once the size of the firm relevant for movers is taken into account, is in excess of 1 and is thus seems very high relative to the incentive parameter. With the inclusion of career concerns in the estimated equation it has the effect of raising the estimated Kappa coefficient by over fifteen fold which indicates a higher shadow price of effort and thus a smaller role for explicit incentives as was anticipated and also greater scale economies in effort (higher Gamma_f). Apart from these changes, the other alterations are quite small.

The partial correlation matrix provided by Table 3 above shows that predicted conditional talent is positively and quite strongly correlated with total CEO income, end-of-period market wealth, total pay, and career length, but is negatively correlated with CEO equity-based incentives (Beta), and particularly CEO age. Younger CEOs appear to be more talented and will naturally also be more concerned about their future career. There is also a small negative correlation with years of experience with the firm prior to becoming a CEO. This indicates that externally recruited CEOs tend to have higher talent and that firm-specific knowledge appears to be a liability. Within the entire market-based sample there is a slight positive correlation with size (total assets) consistent with the higher mean talent level for large firm managers.

To obtain the third element in the production function, observations on average dollar volatility of the firm during each CEO year are used to estimate the elasticity of the stochastic production function with respect to volatility:

$$ln\left(CEO \text{ year dollar volatility}_t\right) = ln\left(\sigma\right) + \gamma_h ln\left(K_{t-1}^{i\hat{a}}\right) + \varepsilon_t^i , \tag{18}$$

utilizing the original 18,835 observations and the split samples of large and small firms that form the basis of the equation (18) estimates. These regression results are summarized in Table 4 above, along with the other coefficients of the stochastic production function. The results indicate that productivity is extremely sensitive to share price volatility with the scale elasticity ranging between 81 percent for the sample of large firms based on market values and as low as 66 percent based on accounting values for non-performing firms. The summary sigma constant measure ranges from a low of 0.58 for large firms to 3.8 for positive performing firms based on accounting values.

The next question to be addressed is how total CEO pay responds to both increases in total assets under management and to conditional talent differences. While it is well-established that CEO pay is higher in larger companies, we are not aware of studies showing the responsiveness of pay to differences in talent levels. To address these questions the log of ability, size, volatility and other variables are regressed on the log of a comprehensive measure of fiscal year CEO total (flow) pay sourced from ExecuComp in constant 2006 dollars:

$$\begin{aligned} ln\left(\text{Ttl py}_t\right) &= ln\left(\text{Fix py}\right) + \rho_\theta ln\left(\hat{\theta}_t^{\hat{a}}\right) + \rho_\beta ln\left(\beta^{i\hat{a}}\right) + \rho_K ln\left(K_{t-1}^{i\hat{a}}\right) + \rho_\sigma ln\left(\text{Dol Vol}_t\right) \\ &+ \rho_{yrs\,off} ln\left(Car\ lth_{it}\right) + \rho_{\text{pre}\,CEO}\, ln\left(\text{Exp}\right) + \rho_{CEO\,Dual} Dum_{Dual\,t} + Res\ Dum_t + Cont_t + \varepsilon_t \end{aligned} , \tag{19}$$

where *ln* is the natural logarithm, ρ_θ is the elasticity with respect to ability, ρ_K is the elasticity of pay with respect to the opening value of total assets under management, ρ_σ the elasticity with respect to the risk borne by the manager (dollar volatility), $\rho_{years\ in\ office}$ is the elasticity of the length of the CEO's tenure with the *i*th firm, $\rho_{\text{pre CEO experience}}$ is the elasticity with respect to years with the firm prior to appointment as CEO, and $\rho_{\text{CEO duality}}$ the impact of CEO-Chair duality. Experience with the firm prior to appointment was included when ExecuComp records such information so as to examine the role of firm-specific experience and the existence of internal CEO selection tournaments in setting CEO pay. Otherwise, it is assumed that the CEO was hired either externally or with little firm-specific knowledge prior to assuming the role.

Since the predicted conditional talent levels include negative values and thus prevent the estimation of elasticity measures, the estimates were normalized with a mean of zero and standard deviation of unity. The distribution was then shifted to the right to ensure that all conditional talent estimates are positive prior to taking logs. A comprehensive measure of total pay from ExecuComp is used. Pay consists of salary plus bonus plus long-term incentive plan plus the value of new options and restricted stock allocated. Additional controls consist of two-digit industry dummies (not shown). Year dummies were deliberately excluded so as to be able to examine the capacity of the modeling to predict rising real pay levels over the sample period. The results are summarized in Table 5.

<< INSERT TABLE 5 ABOUT HERE>>

The impact of the estimated talent for each CEO year in elasticity form on pay is shown in the second row of the Table. These impact estimates range from 55% for all firms to 123% for firms with positive income utilizing the accounting productivity measure and 46% for large firms based on market productivity. This indicates that CEOs employed by firms with positive performance capture about 69% of their exceptional talent in the form of higher pay based on the market measure. By contrast, for negative performers the relationship is far lower at 16% indicating a smaller but still positive relationship between pay and talent. The elasticity of pay with respect to the incentive share, $\hat{\beta}_{t-1}^{i\hat{a}}$, is negative across the board. This indicates that CEOs are penalized by the board in terms of flow incentives when they possess stock incentives, either shares or the share-equivalents of option holdings. Hence they are seen as substitutes by the board. The very high statistical significance of our talent measure indicates that we have been successful in estimating talent from the CEO production function quite independently of pay, and indeed, CEOs are rewarded for talent.

For all firms based on the market measure the elasticity with respect to total assets is 28% and lower at 23% for large firms and thus fairly consistent with the literature but is on the low side. This is to be expected because, unlike the traditional literature, in our regressions managerial talent is held constant. Accounting measures produce slightly lower estimates of around 20% and for firms with positive market performance the rate is 30%. The risk borne by the CEO is captured by the inclusion of the stock dollar volatility term, as indicated by the inclusion of risk in the pay schedule, equation (5) above. In the market-based regressions it has typically an elasticity of 19% but is higher at 27% for all firms based on accounting measures. Since size is highly correlated with risk, the combined effect of size and risk is very high at 47% for each doubling of firm size. Years in office is significantly rewarded with a relatively small positive elasticity of around 8% for additional years in office.

When the influence of career concerns on pay levels is taken into account in column 2, the differences are quite small.

The elasticity estimates for years of experience with the firm prior to CEO appointment indicate that external appointees are paid more and that this experience is not rewarded. In fact, it is penalized with an elasticity of around 5%. The fact that pay falls with internal seniority casts doubt on the efficacy of internal tournaments for promotion to CEO (see, for example, Lazear and Rosen (1981)). CEOs who accept the dual role of board chair typically receive about 11.5% higher pay. There is negligible difference between positive and negative performers. Hence, this finding provides no support for rent-seeking arguments along the lines of Bebchuk and Fried (2004), or the claim that CEOs are systematically overpaid. Finally, CEOs who die in office are paid at a much lower rate but Table 6 below reveals that managers who die in office have exceptionally high income even after controlling for talent. This is probably because older managers are more likely to own shares and, consequently, to receive less direct pay given evidence of substitutability.

In Table 6 the same model as in Table 5 is used to explain total CEO income inclusive of share, and share equivalent of option holdings, ownership. The estimated pay sensitivity $\beta_{t-1}^{\hat{a}}$, CEO income share times the change in the firm's market value, is added to the ExecuComp flow pay estimate used in Table 5 to obtain estimates of the CEOs total income inclusive of incentives. Since firms experiencing negative shareholder income can result in overall negative CEO income, this turns out to be the case for several thousand CEO-years. For consistency, the same log specification as in equation (19) was utilized to explain CEO income, requiring that negative observations be dropped.

Comparing Table 6 with Table 5, it is apparent that CEO talent plays an even more important role in rewarding CEOs using a comprehensive income measure relative to the simple pay measure. For the entire sample using the market measure the sensitivity of income to talent is now much higher at 226%, rising to 242% for positive performers. This high income talent sensitivity is consistent with the very high correlation between talent and CEO income shown in Table 3 above. These findings suggest that CEOs are very aware of their own talent when deciding how many shares to own. Thus while the $\beta_{t-1}^{\hat{a}}$ sensitivity measure tends to be small for CEOs of large firms, it especially rewards talented managers. Once again, the sensitivity of non-performers income to talent is quite low. The sensitivity of income to $\beta_{t-1}^{\hat{a}}$ is, naturally, positive with a typical elasticity of around 20%, rising to 27% for positive performers. However, the lower level of penalties for negative performers may be due in part to the truncation of negative CEO income at zero. The asset under management elasticity is low at 17% based on market productivity. Perhaps the most surprising finding is the high exposure of income to risk. For the overall sample it is 35%, rising as high as 42% for small firms with the accounting measure. This indicates once again that CEOs require much higher expected pay for bearing firm-specific risk due to high diversification costs.

<< INSERT TABLE 6 ABOUT HERE>>

Table 7 reports results indicating the ability of the CEO panel pay model summarized in Table 5 to explain the growth in real pay levels over the sample period. The actual mean pay, number of annual observations for market-based and accounting-based predictions and the corresponding estimates for the five categories of estimates are set out for each of eight sample years. For the overall category, the actual pay increased by 117% from 1993 to 2005 in real terms. The base year is 1993 instead of 1992 due to the small number of observations in the commencement year. Moreover, 2005 is chosen instead of 2006 because the very high average pay in 2006 seems anomalous. Predicted pay increased by 46% whereas actual pay increased by 117%. Since there are no year or time dummies involved, the predicted rise is entirely due to explicable economic factors. Table 8 shows that these factors are firm size and risk (dollar volatility). For the 500 largest companies, firm size has increased by 115% and dollar volatility by 178%.

<< INSERT TABLES 7 AND 8 ABOUT HERE>>

5. Conclusions

Our modeling shows that when it is sufficiently productive, the large firm expectedly pays higher salary than the small firm. (See Proposition 5). In a managerial assignment world with CEO talent common knowledge, it can be socially optimal that large firms (or firms with better production technologies) hire managers with high abilities. Thus, it is conventionally argued that there should be a positive relationship between pay and firm size because large firms hire high ability managers who deserve high pay. However, in an agency world, a large firm (with better production technology) may not always be willing to hire a high (expected) ability manager, partly because most ability rent belongs to the agent in labor market competition and partly because salaries are affected by both the agent expected ability and its volatility. We argue that even when a large firm hires a low-ability manager, the expected pay for the low ability manager can be higher than that for a high-ability manager who is hired by a small firm, if the large firm's productivity and the firm size are sufficiently higher and larger than those of the small firm. We find that, indeed, CEOs in large firms are paid a lot more than in small firms but on average have actually slightly lower conditional talent. More importantly, CEOs in large firms have much higher talent risk-adjusted ability as ability dispersion is lower than for small firms (at about half).

We also find that unlike Jensen and Murphy (1990) or Baker and Hall (2004), one may not claim a negative relationship between the sensitivity and the firm size as problematic without looking at relative productivities across firms. When we check these estimated productivities we find, indeed, that the sensitivity relationship with firm size is optimally negative in equilibrium. Schaefer (1998) found that the pay-performance sensitivity is inversely proportional to the square root of firm size.

We analyze managerial career paths which can also affect the contract sensitivity. Since we find that managerial ability contribution increases faster than the total market productivity volatility as the firm size increases, we expect that a manager who previously worked for a large firm will be given a contract with a higher sensitivity than a manager who previously worked for a small firm. We find significant statistical support for this hypothesis.

Finally, we present a number of new and surprising empirical results that indicate there is some degree of alignment between CEO productivity with respect to scale and CEO pay. A noteworthy aspect of our findings is that while talented high-performing CEOs are financially rewarded, poorly performing CEOs do not seem to be severely penalized.

References

Aggarwal, Rajesh K., and Andrew A. Samwick, 1999, "The other side of the trade-off: The impact of risk on executive compensation," *Journal of Political Economy* 107, 65–105.

Baker, G. and B. J. Hall, 2004, "CEO Incentives and Firm Size," *Journal of Labor Economics*, 22 (4), 767-798.

Barro, Jason, and Robert Barro, 1990, "Pay, Performance and Turnover of Bank CEOs," *Journal of Labor Economics* 8, 448-481.

Bayard, Kimberly and Kenneth R. Troske, 1999, "Examining the Employer-Size Wage premium in the Manufacturing, Retail Trade, and Service Industries Using Employer-Employee Matched Data," *American Economic Review* 89, 99-103.

Bebchuk, Lucian A., and Jesse M. Fried, 2004, *Pay without Performance: The Unfulfilled Promise of Executive Compensation*, Cambridge, MA: Harvard University Press.

Bebchuk, Lucian Arye and Yaniv Grinstein, 2005. "Firm Expansion and CEO Pay," Olin Centre for Law, Economics and Business, Harvard University Discussion paper No. 533, updated 2007.

Becker, Bo, 2006, "Wealth and Executive Compensation," *Journal of Finance* 61 (1), 379-397.

Conyon, M., 1997, "Corporate Governance and Executive Compensation," *International Journal of Industrial Organization* 15 (4), 493-510.

Conyon, Martin J. and K. J. Murphy, 2000, "The Prince and the Pauper? CEO Pay in the United States and United Kingdom," *Economic Journal* 110, 640-671.

Core, John and Wayne Guay, 2002, Estimating the Value of Employee Stock Option Portfolios and Their Sensitivity to Price and Volume, *Journal of Accounting Research* 40 (3), 613-630.

Cosh, Andrew, 1975, "The Remuneration of Chief Executives in the United Kingdom," *Economic Journal* 85, 75-94.

Cosh, A. and A. Hughes, 1997, "Executive Remuneration, Executive Dismissal and Institutional Shareholdings," *International Journal of Industrial Organization* 15 (4), 469-492.

Edmans, Alex, Xavier Gabaix and Augustin Landier, 2007, "A Calibratable Model of Optimal CEO Incentives in Market Equilibrium," Working Paper, MIT Sloan School of Management.

Fama, Eugene., 1980, "Agency problems and the theory of the firm," *Journal of Political Economy* 88 (2), 288-307.

Frydman, Carola and Raven E. Saks, 2007, "Executive Compensation: A New View from a Long-term Perspective, 1936-2005," Working Paper, Federal Reserve Board, Washington DC.

Gabaix, Xavier and Augustin Landier, 2008, "Why Has Ceo Pay Increased So Much," *Quarterly Journal of Economics* 123(1), 2008, 49-100.

Garvey, Gerald T., and Peter L. Swan, 2002, "What Can Market Microstructure Contribute to Explaining Incentive Pay?: Liquidity and the Use of Stock-Based Compensation," Working Paper, Claremont Graduate University and University of New South Wales.

Giannetti, Mariassunta. 2007. Serial CEO Incentives and the Structure of Managerial Contracts. ECGI Finance Working Paper No. 183/207, Stockholm School of Economics, CEPR and ECGI.

Gibbons, Robert and Kevin J. Murphy, 1992, "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 100, 468-505.

Hermalin, Benjamin E., 2005, "Trends in Corporate Governance," *Journal of Finance* 60 (5), 2351-2384.

Holmstrom, B., 1999, "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies* 101, 835-860.

Holmstrom, B. and P. Milgrom, 1987, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica* 55, 303-328.

Idson, Todd L., and Walter Y. Oi, 1999, "Workers are More Productive in Large Firms," *American Economic Review* 89 (2), 104-108.

Jensen, M. and K. J. Murphy, 1990, "Performance pay and top-management incentives," *Journal of Political Economy* 98, 225-264.

Joskow, Paul, Nancy Rose, and Andrea Shepard, 1993, Regulatory Constraints on CEO Compensation," *Brookings papers on Economic Activity: Microeconomics* MCMXCIII, 1-72.

Kaplan, S., 1997, "Top Executive Incentives in Germany, Japan and the US: A Comparison," Working Paper, University of Chicago.

Kato, T., 1997, "Chief Executive Compensation and Corporate Groups in Japan: New Evidence from Micro Data," *International Journal of Industrial Organization* 15 (4), 493-510.

Kato, T. and K. Kubo (2006). "CEO compensation and firm performance in Japan: Evidence from new panel data on individual CEO pay," *Journal of the Japanese and International Economies*; 20(1): 1

Kostiuk, P. F., 1989, "Firm size and executive compensation," *The Journal of Human Resources* 25 (1), 90-105.

Lazear, E. and S. Rosen, 1981, "Rank-order tournaments as optimal labor contracts," *Journal of Political Economy* 89, 841–64.

Lewellen, W. G. and B. Huntsman, 1970, "Managerial pay and corporate performance," *American Economic Review* 60 (4), 710-720.

Lewis, H. G., 1986, *Union relative wage effects*. Chicago: University of Chicago Press.

McKnight, Phillip J. and Cyril Tomkins, 1999, "Top Executive Pay in the United Kingdom: A Corporate Governance Dilemma," *International Journal of the Economics of Business* 6, 223-243.

Merhebi, R., K. Pattenden, P. L. Swan and X. Zhou, 2006, "Australian Chief Executive Officer Remuneration: Pay and Performance," *Accounting and Finance* 46, 481-497**.**

Rajgopal, Shivaram, Terry Shevlin and Valentina Zamora, 2006, CEOs' Outside Employment Opportunities and the Lack of Relative Performance Evaluation in Compensation Contracts," *Journal of Finance* 61 (4), 1813-1844.

Roberts, D. R., 1956, "A general theory of executive compensation based on statistically tested propositions," *Quarterly Journal of Economics* 70 (3), 270-294.

Rose, N., and A. Shepard, 1997, "Firm Diversification and CEO Compensation: Managerial Ability or Executive Entrenchment?" *RAND Journal of Economics* 28, 489-514.

Rosen, S., 1982, "Authority, control and the distribution of earnings", *The Bell Journal of Economics* 13 (2), 311-323.

Rosen, S., 1992, "Contracts and the Market for Executives," in Lars Werin and Hans Wijkander, eds., *Contract Economics*, Cambridge, MA; Oxford: Blackwell.

Schaefer, S., 1998, "The Dependence of Pay-Performance Sensitivity on the Size of the Firm," *Review of Economics and Statistics* 80, 436-443.

Schättler, H. and J. Sung, 1993, "The First-Order Approach to Continuous-Time Principal Agent Problem with Exponential Utility," *Journal of Economic Theory* 61, 331-371.

Zhou, X., 1999, "CEO Pay, Firm Size, and Corporate Performance," *Canadian Journal of Economics*, 33, 213-251.

APPENDIX

Proofs

Proof of Proposition 1:Note that the sign of the performance sensitivity with respect to firm size:

$$sign\left(\frac{\partial}{\partial K^i}\left(\frac{\kappa\hat{e}}{(K^i)^{\gamma_f}}\right)\right) = -sign\left(\frac{(\gamma_g-\gamma_f)(K^i)^{2(\gamma_g-\gamma_f)-1}\sigma_\theta^2}{\sigma_\theta^2(K^k)^{2(\gamma_g-\gamma_h)}+\sigma^2} + (\gamma_h-\gamma_f)(K^i)^{2(\gamma_h-\gamma_f)-1}\right)$$

$$= -sign\left(\frac{(\gamma_g-\gamma_f)(K^i)^{2(\gamma_g-\gamma_h)}\sigma_\theta^2}{\sigma_\theta^2(K^k)^{2(\gamma_g-\gamma_h)}+\sigma^2} + (\gamma_h-\gamma_f)\right)$$

$$= -sign\left((\gamma_g+\gamma_h-2\gamma_f)-(\gamma_g-\gamma_f)\frac{\sigma^2}{\sigma_\theta^2(K^k)^{2(\gamma_g-\gamma_h)}+\sigma^2}\right).$$

This quantity is $<\ 0$ if $\gamma_h-\gamma_f\ \geq\ \gamma_f-\gamma_g>0$, and $>\ 0$ if $0<\gamma_g-\gamma_f\ \leq\ \gamma_f-\gamma_h$.

The rest of the statement of the proposition is obvious. □

Proof of Proposition 2: Let

$$\beta(K^k,K^i,\sigma_{\theta^{\hat{a}}}) = \frac{1}{1+r\kappa\sigma^2\left(\frac{(K^i)^{2(\gamma_g-\gamma_f)}\sigma_{\theta^{\hat{a}}}^2}{\sigma_{\theta^{\hat{a}}}^2(K^k)^{2(\gamma_g-\gamma_h)}+\sigma^2}+(K^i)^{2(\gamma_h-\gamma_f)}\right)}$$

Then, we have

$$\beta(K^L,K^i,\sigma_{\theta^b})-\beta(K^S,K^i,\sigma_{\theta^a})$$
$$=\int_{K^S}^{K^L}\frac{\partial}{\partial K^k}\beta(K^k,K^i,\sigma_{\theta^b})dK^k+\int_{\sigma_{\theta^a}}^{\sigma_{\theta^b}}\frac{\partial}{\partial\sigma_\theta}\beta(K^S,K^i,\sigma_\theta)d\sigma_\theta.$$

Since $sign\left(\partial\beta/\partial K^k\right)=sign(\gamma_g-\gamma_h)$, and $\frac{\partial\beta}{\partial\sigma_\theta}<0$, if $\gamma_g-\gamma_h\geq 0$ and $\sigma_{\theta^b}<\sigma_{\theta^a}$, then $\beta(K^L,K^i,\sigma_{\theta^b})-\beta(K^S,K^i,\sigma_{\theta^a})>0$; and if $\gamma_g-\gamma_h\leq 0$ and $\sigma_{\theta^b}>\sigma_{\theta^a}$, then $\beta(K^L,K^i,\sigma_{\theta^b})-\beta(K^S,K^i,\sigma_{\theta^a})<0$. □

Proof of Proposition 3: Suppose that $\hat{a}$ and $\hat{b}$ are hired for the second period by the small and large firms, respectively. Then, by condition (i), we have $\pi^S(\hat{a}, W_1^{\hat{a}S}) \geq \pi^S(\hat{b}, W_1^{\hat{b}L})$. Thus,

$$g(K^S)E[\theta^{\hat{a}} \mid Y_1] + \Phi(K^{k_{\hat{a}}}, K^S, \sigma_{\theta^{\hat{a}}}) - W_1^{\hat{a}S} \geq g(K^S)E[\theta^{\hat{b}} \mid Y_1] + \Phi(K^{k_{\hat{b}}}, K^S, \sigma_{\theta^{\hat{b}}}) - W_1^{\hat{b}L},$$
$$g(K^S)E[\theta^{\hat{a}} - \theta^{\hat{b}} \mid Y_1] + \Phi(K^{k_{\hat{a}}}, K^S, \sigma_{\theta^{\hat{a}}}) - \Phi(K^{k_{\hat{b}}}, K^S, \sigma_{\theta^{\hat{b}}}) \geq W_1^{\hat{a}S} - W_1^{\hat{b}L}.$$

We also have:

$$\pi^L(\hat{b}, W_1^{\hat{b}L}) \geq \pi^L(\hat{a}, W_1^{\hat{a}S}),$$
$$g(K^L)E[\theta^{\hat{b}} \mid Y_1] + \Phi(K^{k_{\hat{b}}}, K^L, \sigma_{\theta^{\hat{b}}}) - W_1^{\hat{b}L} \geq g(K^L)E[\theta^{\hat{a}} \mid Y_1] + \Phi(K^{k_{\hat{a}}}, K^L, \sigma_{\theta^{\hat{a}}}) - W_1^{\hat{a}S},$$
$$W_1^{\hat{a}S} - W_1^{\hat{b}L} \geq g(K^L)E[\theta^{\hat{a}} - \theta^{\hat{b}} \mid Y_1] + \Phi(K^{k_{\hat{a}}}, K^L, \sigma_{\theta^{\hat{a}}}) - \Phi(K^{k_{\hat{b}}}, K^L, \sigma_{\theta^{\hat{b}}}),$$

Combining the above inequalities, we have:

$$g(K^S)E[\theta^{\hat{a}} - \theta^{\hat{b}} \mid Y_1] + \Phi(K^{k_{\hat{a}}}, K^S, \sigma_{\theta^{\hat{a}}}) - \Phi(K^{k_{\hat{b}}}, K^S, \sigma_{\theta^{\hat{b}}}) \geq W_1^{\hat{a}S} - W_1^{\hat{b}L}$$
$$\geq g(K^L)E[\theta^{\hat{a}} - \theta^{\hat{b}} \mid Y_1] + \Phi(K^{k_{\hat{a}}}, K^L, \sigma_{\theta^{\hat{a}}}) - \Phi(K^{k_{\hat{b}}}, K^L, \sigma_{\theta^{\hat{b}}}),$$

that is,

$$\left(g(K^L) - g(K^S)\right)E[\theta^{\hat{a}} - \theta^{\hat{b}} \mid Y_1]$$
$$\leq \Phi(K^{k_{\hat{a}}}, K^S, \sigma_{\theta^{\hat{a}}}) - \Phi(K^{k_{\hat{b}}}, K^S, \sigma_{\theta^{\hat{b}}}) - \Phi(K^{k_{\hat{a}}}, K^L, \sigma_{\theta^{\hat{a}}}) + \Phi(K^{k_{\hat{b}}}, K^L, \sigma_{\theta^{\hat{b}}}).$$

If $(\hat{a}, \hat{b}) = (a, b)$, then the above inequality implies

$E[\theta^a - \theta^b \mid Y_1] \leq A(K^S, K^L; \sigma_{\theta^a}, \sigma_{\theta^b})$.

If $(\hat{a}, \hat{b}) = (b, a)$, then the same inequality implies

$E[\theta^a - \theta^b \mid Y_1] \geq A(K^S, K^L; \sigma_{\theta^a}, \sigma_{\theta^b})$.

On the other hand, by the zero profit condition for the small firm in the definition of equilibrium in the executive labor market, reservation certainty equivalent wealth levels $(W_1^{\hat{a}S}, W_1^{\hat{a}L})$ are determined as follows:

$\pi^S(\hat{a}, W_1^{\hat{a}S}) = 0,$ and $\pi^S(\hat{b}, W_1^{\hat{b}L}) = 0.$ That is,

$$\pi^S(\hat{a}, W_1^{\hat{a}S}) = g(K^S)E[\theta^{\hat{a}} \mid Y_1^{k_{\hat{a}}}(\hat{a})] + \Phi(K^{k_{\hat{a}}}, K^S, \sigma_{\theta^{\hat{a}}}) - W_1^{\hat{a}S} = 0,$$
$$\pi^S(\hat{b}, W_1^{\hat{b}L}) = g(K^S)E[\theta^{\hat{b}} \mid Y_1^{k_{\hat{b}}}(\hat{b})] + \Phi(K^{k_{\hat{b}}}, K^S, \sigma_{\theta^{\hat{b}}}) - W_1^{\hat{b}L} = 0.$$

Therefore, the assertion of the proposition follows. □

Proof of Proposition 4: Suppose that at time 0, agent $\hat{a}$ is hired by firm *k*. Then by Proposition 3, we know that agent $\hat{a}$ will move to firm *i* (=*S*,*L*) for the second period with a certainty equivalent wealth of $W_1^{\hat{a}} (= g(K^S) E[\theta^{\hat{a}} \mid Y_1^k(\hat{a})] + \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}))$. Note that $W_1^{\hat{a}}$ is unaffected by the agent's choice of a firm to join for the second period. Then, given contract $S_1^k(\hat{a}) = \alpha^{\hat{a}k} + \beta^{\hat{a}k} Y_1^k(\hat{a})$, the agent's expected utility at time 0 is:

$$
\begin{aligned}
& E\left[-\exp\left\{-r\left(S_1^k(\hat{a}) - c(\hat{e}) + W_1^{\hat{a}k}\right)\right\}\right] \\
&= E\left[-\exp\left\{-r\left(\alpha^{\hat{a}k} + \beta^{\hat{a}k} Y_1^k(\hat{a}) - c(\hat{e}) + g(K^S) E[\theta^{\hat{a}} \mid Y_1^k(\hat{a})] + \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}})\right)\right\}\right] \\
&= E\left[-\exp\left\{-r\left(\begin{array}{l} \alpha^{\hat{a}k} + \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}) - c(\hat{e}) \\ + g(K^S)\left(m_{\theta^{\hat{a}}} - p^k \overline{e} f(K^k) - p^k m_{\theta^{\hat{a}}} g(K^k)\right) \\ + (\beta^{\hat{a}k} + g(K^S) p^k) Y_1^k(\hat{a}) \end{array}\right)\right\}\right] \\
&= -\exp\left\{-r\left(\begin{array}{l} \alpha^{\hat{a}k} + \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}) - c(\hat{e}) \\ + g(K^S)\left(m_{\theta^{\hat{a}}} - p^k \overline{e} f(K^k) - p^k m_{\theta^{\hat{a}}} g(K^k)\right) \\ + (\beta^{\hat{a}k} + g(K^S) p^k)(\hat{e} f(K^k) + m_{\theta^{\hat{a}}} g(K^k)) \\ - \frac{r}{2}(\beta^{\hat{a}k} + g(K^S) p^k)^2 \left(\sigma_{\theta^{\hat{a}}}^2 g^2(K^k) + \sigma^2 h^2(K^k)\right) \end{array}\right)\right\} \qquad \text{(A2)} \\
&= -\exp(-r W_0^{\hat{a}})
\end{aligned}
$$

By the FOC with respect to effort level *e*, we have first-period pay sensitivity of:

$$\beta^{\hat{a}k} = \frac{c'(e)}{f^k} - g(K^S) p^k.$$

Note that in equilibrium, $\hat{e} = \overline{e}$, and thus the definition of $W_0^{\hat{a}}$ in equation (A2) implies equation (9).

Thus, first-period expected pay is:

$$E[S^k(Y_1^k)] = W_0^{\hat{a}} - \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}) + c(\hat{e}) - g(K^S) m_{\theta^{\hat{a}}} + \frac{r}{2}\left(\sigma_{\theta^{\hat{a}}}^2 g^2(K^k) + \sigma^2 h^2(K^k)\right)\left(\frac{c'(e)}{f^k}\right)^2,$$

and the expected profit to firm *i* for the first period is:

$$
\begin{aligned}
E[Y_1^k - S_1^k] = {} & e f(K^k) + m_{\theta^{\hat{a}}}\left(g(K^k) + g(K^S)\right) - W_0^{\hat{a}} + \Phi(K^k, K^S, \sigma_{\theta^{\hat{a}}}) \\
& - c(\hat{e}) - \frac{r}{2}\left(\sigma_{\theta^{\hat{a}}}^2 g^2(K^k) + \sigma^2 h^2(K^k)\right)\left(\frac{c'(e)}{f^k}\right)^2.
\end{aligned} \qquad \text{(A3)}
$$

Then the FOC with respect to effort e for firm i to maximize expected profit implies:

$$\frac{\kappa e}{\left(K^k\right)^{\gamma_f}} = \frac{1}{1 + r\kappa\left(K^k\right)^{-2\gamma_f}\left(\sigma^2_{\theta^{\hat{a}}}\left(K^k\right)^{2\gamma_g} + \sigma^2\left(K^k\right)^{2\gamma_h}\right)}.$$

Thus, the sensitivity of the contract at time 0 becomes as stated in (10), and the expected compensation as in (11).

On the other hand, by Definition 1-(ii), the equilibrium certainty equivalent wealth of agent $\hat{a}$ is $W_0^{\hat{a}} = 2m_{\theta^{\hat{a}}} g(K^S) + \Phi(K^S, K^S, \sigma_{\theta^{\hat{a}}}) + \Psi(K^S, \sigma_{\theta^{\hat{a}}})$. Thus by substituting this certainty equivalent wealth and the FOC back into equation (A3), we have equation (12). □

Proof of Proposition 5: If agent $\hat{a}$ ($\in \{a,b\}$) were hired by the small firm, Proposition 4 implies the net profit to the firm would be as follows:

$$E[Y_1^S - S^S(\hat{a})] = m_{\theta^{\hat{a}}}\left(g(K^S) + g(K^S)\right) - W_0^{\hat{a}} + \Phi(K^S, K^S, \sigma_{\theta^{\hat{a}}}) + \Psi\left(K^S, \sigma_{\theta^{\hat{a}}}\right).$$

If agent $\hat{a}$ ($\in \{a,b\}$) were hired by the large firm, the profit to the firm would be as follows:

$$E[Y_1^L - S^L(\hat{a})] = m_{\theta^{\hat{a}}}\left(g(K^L) + g(K^S)\right) - W_0^{\hat{a}} + \Phi(K^L, K^S, \sigma_{\theta^{\hat{a}}}) + \Psi\left(K^L, \sigma_{\theta^{\hat{a}}}\right).$$

Since by Definition 1-(ii), the small firm is indifferent between the two agents in equilibrium, we have $E[Y_1^S - S^S(a)] = E[Y_1^S - S^S(b)] = 0$. Thus,

$$\begin{aligned} W_0^a - W_0^b &= 2(m_{\theta^a} - m_{\theta^b})g(K^S) + \Phi(K^S, K^S, \sigma_{\theta^a}) - \Phi(K^S, K^S, \sigma_{\theta^b}) \\ &\quad + \Psi\left(K^S, \sigma_{\theta^a}\right) - \Psi\left(K^S, \sigma_{\theta^b}\right). \end{aligned}$$

On the other hand, Definition 1-(i) implies the following condition should hold in equilibrium for agent b to be hired by the large firm.

$$\begin{aligned} &(m_{\theta^b} - m_{\theta^a})\left(g(K^L) + g(K^S)\right) + W_0^a - W_0^b \\ &+ \Phi(K^L, K^S, \sigma_{\theta^b}) - \Phi(K^L, K^S, \sigma_{\theta^a}) + \Psi\left(K^L, \sigma_{\theta^b}\right) - \Psi\left(K^L, \sigma_{\theta^a}\right) \geq 0. \end{aligned}$$

By substitution,

$$\begin{aligned}
&(m_{\theta^b}-m_{\theta^a})\left(g(K^L)-g(K^S)\right)+\Phi(K^S,K^S,\sigma_{\theta^a})-\Phi(K^S,K^S,\sigma_{\theta^b})\\
&+\Psi\left(K^S,\sigma_{\theta^a}\right)-\Psi\left(K^S,\sigma_{\theta^b}\right)+\Phi(K^L,K^S,\sigma_{\theta^b})\\
&-\Phi(K^L,K^S,\sigma_{\theta^a})+\Psi\left(K^L,\sigma_{\theta^b}\right)-\Psi\left(K^L,\sigma_{\theta^a}\right)\\
&=(m_{\theta^b}-m_{\theta^a})\left(g(K^L)-g(K^S)\right)+\int_{K^S}^{K^L}\int_{\sigma_{\theta^a}}^{\sigma_{\theta^b}}\Phi_{K^k\sigma_\theta}(K^k,K^S,\sigma_\theta)d\sigma_\theta dK^k\\
&+\int_{K^S}^{K^L}\int_{\sigma_{\theta^a}}^{\sigma_{\theta^b}}\Psi_{K\sigma_\theta}\left(K,\sigma_\theta\right)d\sigma_\theta dK\geq 0.
\end{aligned}$$

The above inequality holds under the stated hypotheses of the proposition □

Proof of Proposition 6: From equations (13) and (14), we have

$$\begin{aligned}
E[S_1^L]-E[S_1^S]&=(m_{\theta^b}-m_{\theta^a})g(K^S)+\Psi(K^S,\sigma_{\theta^b})-\Psi(K^S,\sigma_{\theta^a})\\
&+\Phi(K^S,K^S,\sigma_{\theta^b})-\Phi(K^L,K^S,\sigma_{\theta^b})+\Psi(K^L,\sigma_{\theta^b})-\Psi(K^S,\sigma_{\theta^a})\\
&=(m_{\theta^b}-m_{\theta^a})g(K^S)+2\int_{\sigma_{\theta^a}}^{\sigma_{\theta^b}}\Psi_{\sigma_\theta}(K^S,\sigma_\theta)d\sigma_\theta\\
&-\int_{K^S}^{K^L}\Phi_{K^k}(K^k,K^S,\sigma_{\theta^b})dK^k+\int_{K^S}^{K^L}\Psi_K(K,\sigma_{\theta^b})dK.
\end{aligned}$$

For the second statement, note that since $\gamma_g<2\gamma_f$, $\gamma_h<2\gamma_f$, and $\gamma_g\leq\gamma_h$, we have $\frac{\partial}{\partial K^k}\Phi(K^k,K^S,\sigma_\theta)<0$, and $\frac{\partial}{\partial K}\Psi(K,\sigma_\theta)>0$. Therefore, $E[S_1^L]-E[S_1^S]\geq 0.$ □

Starting Values for Non-Linear Estimation

In order to be able to estimate the production function in levels starting values of the parameters are required for non-linear estimation of the period 2 model excluding career concerns. Equation (15) is modified to remove the term p^k and then rearranged as:

$$\left\{\hat{Y}_t^{i\hat{a}}-\left[\bar{\hat{\theta}}_t^{\hat{a}}(\hat{K}_{t-1}^{i\hat{a}})^{\gamma_g}+\hat{\sigma}(\hat{K}_{t-1}^{i\hat{a}})^{\gamma_h}\varepsilon_t^i\right]\right\}=\frac{1}{\kappa}\left(\hat{\beta}_{t-1}^{i\hat{a}}+\left(\hat{K}_{t-1}^S\right)^{\gamma_g}\bar{p}\right)\left(\hat{K}_{t-1}^{i\hat{a}}\right)^{2\gamma_f}. \qquad \text{(A4)}$$

Since the econometrician cannot observe the CEO-specific talent updating terms in equation (15), we have $p_{t-1}^{\hat{a}i}\equiv\bar{p}$ in equation (A4) above as a coefficient to be estimated for period 1 in the model with this coefficient set to zero in period 2.

We obtain initial starting estimates of the effort and ability production elasticities in turn, beginning with the effort elasticity. These starting values are then used in the direct estimation of the non-linear production function. We take advantage of the fact that the ability of the $\hat{a}$ th agent, $\theta_t^{\hat{a}}$, is a drawing from a random distribution and thus may not be systematically related to capital stock size $\hat{K}_{t-1}^{i\hat{a}}$ and that ε_t^i is a standard-normal random variable with a zero mean. Thus as an approximation we can take the expression in square brackets on the LHS of equation (A4) to be zero in expectation. On taking logarithms we now obtain a simple estimable equation using ordinary least squares (OLS):

$$ln\left\{\frac{\hat{Y}_t^{ia}}{\left(\hat{\beta}_{t-1}^{i\hat{a}}+\left(\hat{K}_{t-1}^{S}\right)^{\gamma_g}\overline{p}\right)}\right\}=ln\left(\frac{1}{\kappa}\right)+2\gamma_f ln\left(\hat{K}_{t-1}^{i\hat{a}}\right), \tag{A5}$$

with the intercept estimate $\hat{\alpha}_0 = ln\left(\frac{1}{\kappa}\right)$, the (common) marginal cost of effort coefficient, $\hat{\kappa}=e^{-\hat{\alpha}_0}$, the estimated effort elasticity with respect to the production function, $\hat{\gamma}_f = \frac{1}{2}\hat{\alpha}_1$, where $\hat{\alpha}_1$ is the slope coefficient. These values are then used as starting values in the non-linear estimation of the regression equation based on the production function, equations (16) and (17) in the text. The use of the non-linear approach is true to the assumed additive nature of the specified error structure.

The starting values for the ability elasticity and mean ability level can now be estimated: Once again setting the error term ε_t^i to its expected value of zero and p^k to zero, we have by rearranging equation (15) in the text:

$$ln\left[\hat{Y}_t^i(\hat{a})-\frac{1}{\hat{\kappa}}\left(\hat{\beta}_{t-1}^{i\hat{a}}+\left(\hat{K}_{t-1}^{S}\right)^{\gamma_g}\overline{p}\right)\left(\hat{K}_{t-1}^{i\hat{a}}\right)^{2\hat{\gamma}_f}\right]=ln\left(\bar{\hat{\theta}}_t^{\hat{a}}\right)+\gamma_g ln(\hat{K}_{t-1}^{i\hat{a}}), \tag{A6}$$

where $\bar{\hat{\theta}}_t^{\hat{a}}$ denotes the mean level of ability for the sample.

Table 1: Summary Statistics of CEO Careers, 1992-2006, Based on Market Productivity/Wealth

Fiscal year values in constant 2006 dollars based on the CPI. Sources are S&P ExecuComp, S&P Compustat and CRSP. All CEOs excluding those in financial services and with tenure of at least one year in one firm are included. Market wealth/productivity consists of the total market value of assets (equity plus total debt) at fiscal year end plus the net value of all distributions to equity and debt holders during the year. Pay-performance sensitivity ($\hat{\beta}_{t-1}^{i\hat{a}}$) consists of the proportion of shares on issue held by the CEO at fiscal year open from ExecuComp inclusive of restricted stock plus the share equivalent (hedge ratio value) of his option holdings based on the Black Scholes formula. Total Pay consists of the value of salary plus bonus plus restricted stock grants plus the value of option grants as reported by ExecuComp and converted to 2006 prices. Total CEO income consists of Total Pay as before plus $\hat{\beta}_{t-1}^{i\hat{a}}$ times the change in the firm's equity market value. The dollar volatility of the firm's stock is computed from CRSP data as the product of the standard deviation of returns and the opening value of market capitalization for each financial year. The CEOs firm experience prior to being appointed CEO is computed from the date the CEO joined the firm until appointed CEO where this is recorded by ExecuComp. Where this information is not reported the CEO is assumed to have been externally recruited. The average age of the CEO over his tenure is computed for the smaller sample of CEOs for which ExecuComp supplies this information. The larger and smaller firm samples are obtained by equally dividing the entire sample of size-ranked CEO fiscal years. The sample of positive and negative performances are found by dividing the sample between CEO fiscal years in which the income to claimants on the firm (dividends plus capital gains plus distributions to debt holders) is positive and the remainder for which the income is negative.

Variable	No.	Mean	Median	Std Dev	Min	Max
Overall Sample						
Mkt Val Terminal Wlth ($M)	18,835	14,124	2,680	49,886	-131	1,172,257
Mkt Val Ttl Asts ($M)	18,835	13,077	2,470	46,059	8	1,078,253
Beta (PPS)	18,835	0.0278	0.0030	0.0675	0.0000	0.7370
Total Pay ($000)	18,662	5,059	2,471	19,395	0	2,268,428
Total Income ($000) incl. Shares	18,649	20,575	3,087	611,080	-32,528,084	47,694,393
Dol Volat ($M)	18,835	1,947	425	7,364	1	220,876
Career Length (Yrs)	18,835	6.7	6	3.5056	1	15
Yrs Exp (Pre-CEO)	18,835	6.6	2	8.7115	0	48
CEO Age (Yrs)	13,158	55.6	56	7.8	29	91
Sample of Large Firms						
Mkt Val Terminal Wlth ($M)	9,417	27,070	8,839	68,130	40	1,172,257
Mkt Val Ttl Asts ($M)	9,417	25,132	8,206	62,866	2,472	1,078,253
Beta (PPS)	9,417	0.0174	0.0016	0.0535	0.0000	0.5829
Total Pay ($000)	9,346	7,824	4,460	26,953	0	2,268,428
Total Income ($000) incl. Shares	9,344	33,149	5,209	858,054	-32,528,084	47,694,393
Dol Volat ($M)	9,417	3,647	1,198	10,130	11	220,876
Career Length (Yrs)	9,417	6.8	6	3.4552	1	15
Yrs Exp (Pre-CEO)	9,417	8.6	4	10.0425	0	43
CEO Age (Yrs)	6,334	56.1	56	7.3	29	85

Table 1: Continued.

Variable	No. Obs.	Mean	Median	Std Dev	Min	Max
Sample of Small Firms						
Mkt Val Terminal Wlth ($M)	9,418	1,179	967	963	-131	26,611
Mkt Val Ttl Asts ($M)	9,418	1,023	897	631	8	2,470
Beta (PPS)	9,418	0.0381	0.0062	0.0776	0.0000	0.7370
Total Pay ($000)	9,316	2,284	1,477	3,072	0	67,725
Total Income ($000) incl. Shares	9,305	7,949	1,800	93,708	-322,722	7,912,450
Dol Volat ($M)	9,418	248	180	286	1	16,507
Career Length (Yrs)	9,418	6.6	6	3.5513	1	15
Yrs Exp (Pre-CEO)	9,418	4.6	2	6.5485	0	48
CEO Age (Yrs)	6,824	55.1	55	8.3	29	91
Sample of Firms with Positive Market Performace						
Mkt Val Terminal Wlth ($M)	11,942	16,368	3,362	55,525	13	1,172,257
Mkt Val Ttl Asts ($M)	11,942	13,617	2,633	47,789	8	966,654
Beta (PPS)	11,942	0.0275	0.0030	0.0677	0.0000	0.7370
Total Pay ($000)	11,839	5,343	2,731	22,303	0	2,268,428
Total Income ($000) incl. Shares	11,835	48,424	5,341	668,894	-581,074	47,694,393
Dol Volat ($M)	11,942	1,624	405	5,168	1	127,765
Career Length (Yrs)	11,942	6.8	6	3.4988	1	15
Yrs Exp (Pre-CEO)	11,942	6.9	2	8.9239	0	48
CEO Age (Yrs)	8,563	55.6	56	7.6	30	91
Sample of Firms with Negative Market Performace						
Mkt Val Terminal Wlth ($M)	6,893	10,236	1,757	37,886	-131	1,007,422
Mkt Val Ttl Asts ($M)	6,893	12,142	2,191	42,882	9	1,078,253
Beta (PPS)	6,893	0.0282	0.0031	0.0671	0.0000	0.5780
Total Pay ($000)	6,823	4,566	2,046	12,862	0	805,983
Total Income ($000) incl. Shares	6,814	-27,794	796	491,170	-32,528,084	5,295,721
Dol Volat ($M)	6,893	2,508	464	10,071	2	220,876
Career Length (Yrs)	6,893	6.4	6	3.5012	1	15
Yrs Exp (Pre-CEO)	6,893	6.1	2	8.3035	0	48
CEO Age (Yrs)	4,595	55.4	55	8.3	29	87

Table 2: Summary Statistics of CEO Careers, 1992-2006, Based on Accounting Productivity

Fiscal year values in constant 2006 dollars based on the CPI. Sources are S&P ExecuComp, S&P Compustat and CRSP. All CEOs excluding those in financial services and with tenure of at least one year in one firm are included. Accounting wealth/productivity consists of the total book value of assets (equity plus total debt) at fiscal year end plus the net value of all distributions to equity and debt holders during the year. The sample of positive and negative performances are found by dividing the sample between CEO fiscal years in which the income to claimants on the firm (accounting income plus distributions to debt holders) is positive and the remainder for which the income is negative. The remaining variables are defined as in Table 1.

Variable	No. Obs.	Mean	Median	Std Dev	Min	Max
Overall Sample						
Book Val Terminal Wlth ($M)	18,835	9,351	1,576	39,130	-19	1,126,323
Book Val Ttl Asts ($M)	18,835	8,554	1,412	35,536	4	922,600
Beta (PPS)	18,835	0.0278	0.0030	0.0675	0.0000	0.7370
Total Pay ($000)	18,662	5,059	2,471	19,395	0	2,268,428
Total Income ($000) incl. Shares	18,649	20,575	3,087	611,080	-32,528,084	47,694,393
Dol Volat ($M)	18,835	1,947	425	7,364	1	220,876
Career Length (Yrs)	18,835	6.7	6	3.5056	1	15
Yrs Exp (Pre-CEO)	18,835	6.6	2	8.7115	0	48
CEO Age (Yrs)	13,158	55.6	56	7.8	29	91
Sample of Firms with Positive Accounting Performace						
Book Val Terminal Wlth ($M)	15,287	9,898	1,732	41,599	8	1,126,323
Book Val Ttl Asts ($M)	15,287	8,704	1,502	36,808	8	922,600
Beta (PPS)	15,287	0.0291	0.0031	0.0693	0.0000	0.7370
Total Pay ($000)	15,154	5,255	2,626	21,130	0	2,268,428
Total Income ($000) incl. Shares	15,147	25,826	3,584	669,654	-32,528,084	47,694,393
Dol Volat ($M)	15,287	2,028	456	7,593	3	220,876
Career Length (Yrs)	15,287	6.9	7	3.5048	1	15
Yrs Exp (Pre-CEO)	15,287	6.9	2	8.9205	0	48
CEO Age (Yrs)	10,818	55.7	56	7.8	29	91
Sample of Firms with Negative Accounting Performace						
Book Val Terminal Wlth ($M)	3,548	6,994	947	25,799	-19	715,355
Book Val Ttl Asts ($M)	3,548	7,907	1,074	29,428	4	800,179
Beta (PPS)	3,548	0.0220	0.0026	0.0584	0.0000	0.6271
Total Pay ($000)	3,508	4,209	1,910	8,458	0	151,597
Total Income ($000) incl. Shares	3,502	-2,138	1,474	219,884	-8,470,434	8,355,927
Dol Volat ($M)	3,548	1,597	305	6,273	1	136,208
Career Length (Yrs)	3,548	5.8	5	3.3762	1	15
Yrs Exp (Pre-CEO)	3,548	5.1	2	7.5709	0	40
CEO Age (Yrs)	2,340	54.8	55	8.0	29	90

Table 3: Partial Correlation Coefficient Matrix for Entire Sample Utilizing Company Market Productivity, 1992-2006, with the Inclusion of our Talent Estimates

	Talent	Wealth	Asset	Beta	Pay	Income	Volat	Career	Exper	Age
Pred Talent (Theta)	1									
Mkt Terminal Wealth	0.0394	1								
Mkt Val Total Assets	-0.0027	0.9812	1							
Beta (income share)	-0.0261	-0.0500	-0.0529	1						
Total Pay (ExecuComp)	0.0316	0.1550	0.1551	-0.0353	1					
Total Income incl. Shares	0.0911	0.0917	0.0340	0.0843	0.0305	1				
Dollar Volatility	-0.0215	0.6032	0.6628	-0.0336	0.1623	-0.0133	1			
CEO Career Length (Yrs)	0.0146	0.0298	0.0242	0.1604	0.0075	0.0182	0.0249	1		
Yrs Experience (Pre-CEO)	-0.0190	0.1596	0.1600	-0.0489	0.0075	-0.0047	0.1025	0.0606	1	
CEO Age	-0.0595	0.0195	0.0170	0.1282	-0.0018	-0.0098	-0.0116	0.0441	0.0825	1

Table 4: CEO Production Function - Non-Linear Regression Equation Estimates

Non-linear regression estimates of the CEO production function as given by equation (16) in the text. The dependent variable is the residual based on the financial year total dollar market wealth (total market value of assets at end of year plus total net distributions made up of dividends and interest payments less the net value of new equity and debt capital raisings) for each of the three market samples, All Firms, Large Firms, and Small Firms and total dollar accounting wealth (total book value of assets at end of year plus total distributions made up of dividends and interest payments) for each of the three accounting samples. The three coefficients, Kappa (κ), Gamma_f (γ_f), and Gamma_g (γ_g), are estimated separately for the entire market sample and entire accounting sample and the two subsamples of large and small stocks. The coefficient representing the impact of career concerns is estimated only for the entire sample using the market method. The coefficients for the entire market sample are applied to firm CEO years with positive incomes (dividends plus capital gains plus total payments to debt holders) and also to firm-years with negative incomes. The average Theta ($\bar{\theta}$) (CEO ability) factor is estimated for the entire sample and the large and small sub-samples. The mean and standard deviation, values of Theta (θ^{ij}) are also implied by treating the estimated production function as an identity with differing Theta values for each CEO year are reported for the large and small subsamples respectively.

Coefficient	**All Firms**			**Pos Wlth Gain**		**Neg Wlth Gain**		**Large**	**Small**
	Mkt.		**Accg.**	**Mkt.**	**Accg.**	**Mkt.**	**Accg.**	**Mkt.**	
	NA	**Career**	NA	NA	NA	NA	NA	NA	NA
Kappa (κ)	0.3548*	17.8000*	0.3926*	0.3548*	0.3926*	0.3548*	0.3926*	2.5028*	23.7534*
(t-value)	(4.67)	(6.11)	(5.20)	(4.67)	(5.20)	(4.67)	(5.20)	(3.76)	(2.87)
Shadow Price Career (p)	NA	1.625E-03*	NA	NA	NA	NA	NA	NA	NA
(t-value)	NA	(6.05)	NA	NA	NA	NA	NA	NA	NA
Gamma_f (γ_f)	0.3812*	0.5533*	0.3788*	0.3812*	0.3788*	0.3812*	0.3788*	0.4675*	0.6195*
(t-value)	(43.58)	(315)	(47.16)	(43.58)	(47.16)	(43.58)	(47.16)	(42.99)	(28.37)
Gamma_g (γ_g)	0.9848*	0.9606*	0.9970*	0.9848*	0.9970*	0.9848*	0.9970*	0.9895*	0.8797*
(t-value)	(8,920)	(689)	(12,079)	(8,920)	(12,079)	(8,920)	(12,079)	(6,329)	(1,181)
Est. Av. Ability (θ))	1.2896*	1.2911*	1.1307*	1.4012*	1.1673*	1.0498*	0.8884*	1.2143*	2.7383*
(t-value)	(725)	(534)	(970)	(850)	(1,228)	(474)	(278)	(513)	(201)
Slope: Pred Prod	1.0012*	0.9997*	1.0003*	0.9991*	0.9998*	1.0047*	0.9984*	1.0009*	1.0575*
(t-value)	(698)	(699)	(944)	(816)	(1,196)	(460)	(269)	(477)	(104)
RSq	0.9627	0.9629	0.9793	0.9824	0.9894	0.9685	0.9533	0.9602	0.5341
Pred Ability (θ)) Mean	1.2787	1.3293	1.1377	1.4944	1.1987	0.9049	0.8749	NA	NA
Pred Ability (θ)) Std Dev	0.6563	0.7683	0.5144	0.7270	0.5461	0.2027	0.1866	NA	NA
	Sample of Large Firms			NA	NA	NA	NA	NA	NA
Pred Ability (θ)) Mean	1.2655	1.3085	1.1321	NA	NA	NA	NA	NA	NA
Pred Ability (θ)) Std Dev	0.4280	0.5315	0.2662	NA	NA	NA	NA	NA	NA
	Sample of Small Firms			NA	NA	NA	NA	NA	NA
Pred Ability (θ)) Mean	1.2918	1.3501	1.1434	NA	NA	NA	NA	NA	NA
Pred Ability (θ)) Std Dev	0.8234	0.9473	0.6770	NA	NA	NA	NA	NA	NA
Dollar Volatility Regression									
Gamma_h (γ_h)	0.8131*	0.8131*	0.6562*	0.8131*	0.6562*	0.8131*	0.6562*	0.8403*	0.7726*
(t-value)	(255)	(255)	(150)	(255)	(150)	(255)	(150)	(118)	(93.41)
Sigma σ	0.7519*	0.7519*	3.8039*	0.7519*	3.8039*	0.7519*	3.8039*	0.5824*	0.9897
(t-value)	(10.99)	(10.99)	(40.31)	(10.99)	(40.31)	(10.99)	(40.31)	(8.14)	(0.19)
RSq	0.7754	0.7754	0.5456	0.7754	0.5456	0.7754	0.5456	0.5967	0.4809

*Significant at 1%; **Significant at 5%; ***Significant at 10%

Table 5: Determinants of CEO (Flow) Pay Levels (in Logarithms) Based on ExecuComp Pay Data, 1992-2006, in 2006 Prices

Equation (19) in the text is estimated for each of the five groups identified in Table 4 and using the predicted ability/talent estimates from Table 4 for each of the five groups.

Variable	All Firms			Pos Income		Neg Income		Large	Small
	Mkt		Acc	Mkt	Acc	Mkt	Acc	Mkt	Mkt
	NA	Career	NA	NA	NA	NA	NA	NA	NA
Intercept Log(Fix Pay)	3.3452*	3.2561*	3.1251*	3.4332*	3.1580*	3.0111*	3.5239*	3.4851*	3.5873*
(t-value)	(24.88)	(24.12)	(21.35)	(19.39)	(19.91)	(14.13)	(12.82)	(15.66)	(20.42)
Log Pred. Ability (θ)	0.5544*	0.5185*	0.5982*	0.6897*	1.2308*	0.1626*	-0.0786	0.4576*	0.4832*
(t-value)	(22.62)	(21.64)	(11.15)	(15.03)	(18.54)	(2.96)	(1.16)	(12.28)	(19.68)
Log Beta (PPS)	-0.0119*	-0.0137*	-0.0125*	-0.0031	-0.0060**	-0.0271*	-0.0270*	-0.0094**	-0.0130*
(t-value)	(4.66)	(5.37)	(4.85)	(0.99)	(2.10)	(6.38)	(4.55)	(2.03)	(4.66)
Log Total Assets	0.2756*	0.2763*	0.2083*	0.2919*	0.2431*	0.2956*	0.2495*	0.2263*	0.2618*
(t-value)	(23.64)	(23.67)	(25.91)	(19.04)	(25.97)	(14.07)	(13.92)	(11.36)	(15.27)
Log Dollar Volatility	0.1882*	0.1892*	0.2663*	0.1752*	0.2326*	0.1678*	0.2155*	0.2055*	0.1705*
(t-value)	(15.94)	(16.00)	(32.58)	(11.12)	(24.66)	(8.11)	(11.78)	(11.37)	(11.03)
Log Years in Office	0.0832*	0.0831*	0.0884*	0.0819*	0.0998*	0.0853*	0.0441***	0.0937*	0.0765*
(t-value)	(7.13)	(7.11)	(7.50)	(5.57)	(7.63)	(4.44)	(1.65)	(4.83)	(5.57)
Log Yrs Pre-CEO Exp	-0.0547*	-0.0548*	-0.0520*	-0.0560*	-0.0498*	-0.0492*	-0.0548*	-0.0642*	-0.0419*
(t-value)	(8.42)	(8.42)	(7.92)	(7.01)	(6.98)	(4.40)	(3.35)	(6.45)	(4.94)
Chair-CEO Duality	0.1137*	0.1122*	0.1131*	0.1078*	0.1143*	0.1228*	0.0762**	0.1547*	0.0757*
(t-value)	(7.48)	(7.38)	(7.36)	(5.72)	(6.78)	(4.80)	(2.14)	(5.98)	(4.34)
Departure-Resigned	-0.0129	-0.0132	-0.0293	-0.0180	-0.0390	-0.0005	0.0009	0.0159	-0.0443
(t-value)	(0.53)	(0.54)	(1.19)	(0.56)	(1.38)	(0.01)	(0.02)	(0.37)	(1.61)
Departure-Retired	-0.0670*	-0.0664*	-0.0681*	-0.0830*	-0.0781*	-0.0237	-0.0004	-0.0640**	-0.0858*
(t-value)	(3.52)	(3.49)	(3.54)	(3.57)	(3.77)	(0.72)	(0.01)	(2.34)	(3.21)
Departure-Deceased	-0.3104*	-0.3204*	-0.3059*	-0.3870*	-0.3139*	-0.1722	-0.1547	-0.2422**	-0.3764*
(t-value)	(4.61)	(4.75)	(4.49)	(4.83)	(4.35)	(1.39)	(0.80)	(1.97)	(5.11)
No Observations	16,490	16,490	16,490	10,540	13,584	5,950	2,906	8,284	8,206
RMSQ	0.8786	0.8797	0.8884	0.8666	0.8810	0.8928	0.8701	0.9884	0.7448
RSq	0.4129	0.4114	0.3997	0.4113	0.4051	0.4147	0.4267	0.2636	0.2605

*Significant at 1%; **Significant at 5%; ***Significant at 10%

Table 6: Determinants of CEO Income Levels (in logs), Based on Execucomp Pay plus Imputed Share Income, 1992-2006, in 2006 Prices

The same equation (19) is estimated as in Table 5 except that the dependent variable is now CEO income made up of total pay as given by ExecuComp plus the income from shares and option holdings equivalents based on the $\hat{\beta}_{t-1}^{i\hat{a}}$ income share of the annual change in equity market capitalization.

Variable	All Firms			Positive Income		Negative Income		Large	Small
	Mkt		Acc	Mkt	Acc	Mkt	Acc	Mkt	
	NA	Career	NA	NA	NA	NA	NA	NA	NA
Intercept (Fixed Pay)	5.1232*	4.7053*	5.6044*	5.1763*	5.8603*	3.4791*	5.6422*	7.0767*	3.7956*
(t-value)	(28.66)	(26.48)	(26.30)	(27.81)	(25.50)	(9.90)	(13.88)	(24.53)	(15.49)
Log Pred. Ability (θ)	2.2576*	2.2669*	1.0484*	2.4160*	1.9238*	0.2532**	-0.0971	1.7748*	2.2376*
(t-value)	(63.69)	(66.61)	(14.93)	(50.66)	(22.11)	(1.96)	(1.10)	(38.35)	(52.59)
Log Beta (PPS)	0.1982*	0.1917*	0.2201*	0.2649*	0.2381*	-0.0028	0.1173*	0.2333*	0.1666*
(t-value)	(63.25)	(61.60)	(62.63)	(81.54)	(63.43)	(0.44)	(13.28)	(45.89)	(42.09)
Log Total Assets	0.1678*	0.1691*	0.0314*	0.2059*	0.0709*	0.3396*	0.2391*	0.0998*	0.0584**
(t-value)	(11.55)	(11.76)	(2.79)	(13.02)	(5.55)	(11.20)	(8.79)	(4.62)	(2.37)
Log Dollar Volatility	0.3486*	0.3507*	0.4675*	0.3434*	0.4229*	0.1271*	0.2535*	0.3955*	0.3126*
(t-value)	(23.41)	(23.80)	(40.24)	(21.16)	(32.47)	(4.21)	(8.89)	(19.97)	(13.69)
Log Years in Office	0.0474*	0.0439*	0.0673*	0.0666*	0.0741*	0.0176	0.0676***	0.0592*	0.0325***
(t-value)	(3.29)	(3.08)	(4.15)	(4.43)	(4.24)	(0.64)	(1.72)	(2.83)	(1.65)
Log Yrs Pre-CEO Exp	-0.0572*	-0.0576*	-0.0460*	-0.0500*	-0.0350*	-0.0705*	-0.0899*	-0.0564*	-0.0484*
(t-value)	(7.17)	(7.30)	(5.12)	(6.11)	(3.69)	(4.47)	(3.75)	(5.31)	(3.94)
Chair-CEO Duality	0.1636*	0.1587*	0.2014*	0.1694*	0.1975*	0.0973*	0.0878***	0.0478***	0.2374*
(t-value)	(8.76)	(8.58)	(9.56)	(8.78)	(8.80)	(2.70)	(1.67)	(1.72)	(9.48)
Departure-Resigned	-0.1464*	-0.1439*	-0.2016*	-0.1200*	-0.2117*	-0.1024**	-0.1063	-0.0835***	-0.1596*
(t-value)	(4.84)	(4.80)	(5.91)	(3.63)	(5.62)	(1.96)	(1.44)	(1.79)	(4.06)
Departure-Retired	-0.2499*	-0.2459*	-0.2597*	-0.2222*	-0.2682*	-0.1580*	-0.1637**	-0.2271*	-0.2650*
(t-value)	(10.83)	(10.76)	(9.99)	(9.34)	(9.87)	(3.50)	(2.29)	(7.80)	(6.97)
Departure-Deceased	0.2685*	0.2423*	0.2394**	0.1677**	0.2156**	0.0841	0.5196***	0.4183*	0.0651
(t-value)	(3.15)	(2.87)	(2.49)	(2.03)	(2.19)	(0.40)	(1.70)	(3.26)	(0.58)
No Observations	14,167	14,167	14,167	10,386	11,931	3,781	2,236	7,304	6,863
RMSQ	1.0020	0.9917	1.1302	0.8810	1.0981	1.0073	1.1251	0.9979	0.9799
RSq	0.5645	0.5735	0.4460	0.6424	0.4765	0.3753	0.3553	0.5276	0.5535

*Significant at 1%; **Significant at 5%; ***Significant at 10%

Table 7: Actual and Predicted Mean CEO Total Pay Based on ExecuComp by Years in 2006 Prices, $000

The pay prediction model as set out in equation (19) and Table 5 is used to predict pay levels in the prices of 2006 by the eight years specified in the table below and for the five sample categories broken down by market and accounting productivity measures. Actual average pay based on ExecuComp and the sample sizes for the years involved are also presented.

Year	Act Mean Pay	No Obs	Mkt.	No Obs	Accg.
All Firms					
1992	3,207	317	4,659	317	4,617
1993	2,755	1014	3,646	1014	3,548
1995	2,910	1425	3,437	1425	3,273
2000	7,029	1413	6,096	1413	6,328
2002	5,704	1293	5,438	1293	6,115
2004	5,816	1211	5,274	1211	5,324
2005	5,972	1150	5,337	1150	5,438
2006	8,697	916	5,718	916	5,863
Large Firms					
1992	3,577	256	4,938	251	5,080
1993	3,454	529	4,405	541	4,524
1995	4,433	601	4,659	606	4,531
2000	10,170	717	7,921	695	8,046
2002	8,383	681	7,039	682	7,708
2004	8,273	673	6,586	686	6,716
2005	8,197	671	6,521	676	6,739
2006	12,297	578	6,748	579	7,081
Small Firms					
1992	1,797	61	1,386	66	1,666
1993	1,955	485	1,612	473	1,585
1995	1,783	824	1,490	819	1,470
2000	3,988	696	1,893	718	2,617
2002	2,714	612	1,665	611	2,115
2004	2,606	538	1,712	525	1,829
2005	2,798	479	1,733	474	1,874
2006	2,510	338	1,751	337	1,923
Positive Income					
1992	3,281	241	4,715	271	4,710
1993	2,807	740	3,988	862	3,872
1995	3,018	1,076	3,973	1,204	3,639
2000	7,367	791	6,676	1,100	7,226
2002	6,050	528	4,982	918	6,592
2004	5,996	884	5,667	997	5,924
2005	6,409	693	5,813	881	5,839
2006	9,451	634	6,352	777	6,462
Negative Income					
1992	2,765	76	4,630	46	4,942
1993	2,459	274	2,961	152	2,826
1995	2,323	349	2,146	221	2,428
2000	5,841	622	5,154	313	4,518
2002	4,856	765	5,255	375	4,789
2004	4,980	327	4,557	214	3,700
2005	4,540	457	4,614	269	4,573
2006	4,481	282	4,356	139	3,955

Table 8: Increase in Firm Size and Dollar Volatility, 1992-2006, in $2006 Prices

Observations on the total market value of assets and risk (dollar volatility) are presented for two years, 1992 and 2006, for the top 100 firms and top 500.

Sample	Year	Mean	Median	Std	Min	Max
		Total Market Value of Assets				
Top 100	1993	61,278	42,094	57,681	21,457	344,716
	2005	145,612	77,645	187,291	37,211	1,078,253
	% Change	137.6%	84.5%	224.7%	73.4%	212.8%
Top 500	1993	18,807	8,349	33,649	2,820	344,716
	2005	40,492	14,901	98,938	4,732	1,078,253
	% Change	115.3%	78.5%	194.0%	67.8%	212.8%
		Dollar Volatility				
Top 100	1993	4,082	3,027	3,941	232	23,561
	2005	11,904	8,606	12,294	330	76,798
	% Change	191.6%	184.3%	212.0%	42.1%	226.0%
Top 500	1993	1,435	767	2,263	90	23,561
	2005	3,972	1,989	6,895	79	76,798
	% Change	176.8%	159.2%	204.7%	-12.7%	226.0%