# Essays on Executive Compensation

Shingo Takahashi

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics.

Chapel Hill
2006

Approved by
Advisor: Thomas Mroz
Reader: David Blau
Reader: Chuanshu Ji
Reader: Matthias Kahl
Reader: Wilbert van der Klaauw

ABSTRACT

Shingo Takahashi: Essays on Executive Compensation

(Under the direction of Thomas Mroz)

Chapter 1 provides empirical evidence of the effect of stock options and total compensation on the job turnover of corporate Chief Executive Officers (CEOs). Our estimates indicate that both the amount and the composition of the compensation package are significant determinants of turnover probability. Holding the total amount of compensation constant, an increase in the proportion of stock options in the total compensation from its median level (0.48) to the 75th percentile level (0.67), would result in a decrease in annual turnover probability from 16 percent to 13.5 percent. On the other hand, holding the proportion of stock options constant, if the total compensation increases from the median level ($2.5 million) to the 75 percentile level ($5 million), turnover probability would decrease to 14 percent. In Chapter 2 we develop a model to describe the relationship between incentive and tenure in a principal-agent setting. One of the standard results of principal agent theories is that pay-for-performance sensitivity increases with the agent's tenure, but this has been rejected by prior empirical studies in CEO compensation literature. In our model, uncertainty dictates if the principal chooses input-based compensation or output-based compensation, where input-based compensation is less incentive intensive. We show that the principal is more likely to choose input-based compensation later in the agent's tenure. This demonstrates that pay-for-performance sensitivity decreases with the agent's tenure -result consistent with the prior empirical findings in the CEO compensation literature. Chapter 3 reexamines the relationship between pay-for-performance sensitivity and tenure using CEO compensation data. Our estimates indicate that there is a strong and positive relationship between pay-for-performance sensitivity and CEO tenure. For CEOs with tenure less than or equal to six years, an improvement in firm performance from the median level to the 75th percentile level would only lead to a 0.06 percent increase in total compensation. For CEOs with tenure of seven years or more, the same improvement in firm performance would lead to an 8 percent increase in total compensation. Our new findings strongly support standard principal-agent theories, but do not support our model in Chapter 2.

TABLE OF CONTENTS

LIST OF TABLES

**Chapter I**

**The Structure of Compensation and CEO Job Turnover**

# 1 Introduction

CEO compensation has been explained commonly by the principal agent theory, where the shareholders of the firm are the principal and the CEO is the agent. Past empirical studies have mainly focused on the effect of firm performance on either compensation or on turnover probability. However, none of the prior literature investigated the relationship between compensation *and* turnover probability.

When investigating the compensation vs. turnover relationship, we consider not only the *amount* but also the *form* of the compensation package. Since 1990, we have seen an unprecedented increase in the use of stock options as part of CEO compensation. A possible effect of the increased use of stock options on turnover probability has been implied by some researchers (Anderson, Banker and Ravindran, 2000), but we know almost nothing about the actual statistical impact. This is the motivation of this study: to document the relationship between compensation and turnover behavior of CEOs. The form of compensation is represented by the *proportion* of stock options in the total compensation package.

Investigating the effect of the form of compensation on CEO turnover is also interesting on theoretical grounds since stock options can be seen as deferred compensation; Standard option pricing theory implies that option holders would wait to exercise until the strike date in order to maximize their profit. Moreover, stock options granted to CEOs usually have vesting period of about three years and un-vested options are usually forfeited if the CEOs leave their firms. One implication of deferred compensation is that it reduces turnover probability since it provides an incentive for CEOs to remain at the firm longer.

The contributions of my study are fourfold. (1) This study provides new empirical evidence about the link between the form of CEO compensation and CEO turnover probability, (2) documents the effect of total compensation on CEO turnover probability, (3) proposes a reasonable choice of instrumental variables to cope with the endogeneity problem which arises in estimating such relationships and (4) estimates the relationships between turnover and total amount and the form of compensation using a joint system of equations that incorporates unobserved heterogeneity, which deals with endogeneity issues most efficiently.

We find that both the amount and the form of total compensation have significant effects on CEO turnover probability. Turnover probability for our median CEO is 16%. An increase in the *proportion* of options in total compensation from the median level to the 75 percentile level (an increase in the proportion from 0.47 to 0.68) would decrease the turnover probability from 16% to 13.5%, while an increase in the *amount* of total compensation from the median level to the 75 percentile level, which is an increase from $2 million to $4 million, would decrease turnover probability from 16% to 14.0%. Thus, if the probability of separation were constant over time, the expected number of years that CEOs would hold office would increase by 19 percent (6.2 years to 7.9 years) when the proportion of options increases from the median level to the 75 percentile level, while it would increases by 14% (6.2 years to 7.4 years) when total compensation increases from the median level to the 75 percentile level, holding the proportion of deferred compensation constant.

Prior empirical work regarding the incentive effect of stock options customarily investigated how total compensation is tied to firm performance where options are simply added to total compensation (for example, Hall and Liebman, 2000). However, this approach ignores the particular characteristics of options as deferred compensation. Our estimated negative relationship between the form of compensation and CEO turnover probability provides fresh evidence that stock options are used as deferred compensation to provide incentives for CEOs to strive longer, binding CEOs to their firms.

Additionally, our study finds that, although changes in firm performance have a relatively small effect on the turnover probability, the turnover related pay-for-performance sensitivity is greater than the comparable figure estimated by Jensen and Murphy (1990, A). We find that the median CEOs with 4 years of experience would lose 56.9 cents for each $1000 lost by shareholders. This is larger than a comparable Jensen and Murphy's figure by a factor of 7.(Jensen and Murphy estimated that CEOs would lose 8.6 cents for each $1000 lost by shareholders.) This difference mainly stems from the fact that our estimated turnover probability is much higher than that of Jensen and Murphy's.

Jensen and Murphy's estimated turnover probability is 4.6 percent for CEOs whose age is 53 years old while our comparable estimate is 16 percent. Jensen and Murphy's small estimate may be partly because they did not exclude the left censored observations which would have biased their estimate downwards. It could also be that their estimate was contaminated with a similar problem with age variables that we found in ExecuComp data set. [1]

Finally, our study finds little support for managerial entrenchment hypothesis through interlocking directorship. The presence of interlocking directorship did not appear to have a negative effect on turnover probability for CEOs whose tenure is less than 10 years after controlling for biases due to left censoring (or the interrupted spells biases). We only find a negative effect of interlocking directorship when the sample contains left censored observations. However, since we do not find a negative effect when we eliminate the left censored observations, we are not able to tell if such negative effect is due to a true interlocking directorship effect.

## 2 Theories and prior empirical studies

This section summarizes several existing theories and the prior empirical studies that describe CEO compensation. First I outline the principal agent theory, second managerial entrenchment and skimming theories, and third the matching theory.

[1] Age is missing for more than 70% of the observations. Now, divide the sample into two sub-sample, one for non-missing age, and the other for missing age. Then, the average exit rate for the first sub-sample is about 0.019 while the average exit rate is 0.13. Therefore, restricting the sample to the observations for non-missing age variable produces unrealistically low turnover rate.

## 2.1 Principal agent theory and prior empirical work

CEO compensation has been commonly explained by principal agent theories, where the shareholders of the firm are the principal and the CEO is the agent. Since ownership is separated from management in public corporations, asymmetries of information occur between the CEO and the shareholders. This leads the principal to form a contract with the CEO as a way of directing the CEO's actions towards their interests. Prior theoretical work has studied many types of contracts. The greater part of the literature on CEO compensation has been devoted to testing the pay-for-performance contract. However, another type of contract, the "deferred compensation contract" seems particularly informative about some aspects of CEO compensation, especially the stock option component of CEO pay. In this section, I briefly outline the theory and empirical work on "deferred compensation" followed by a summary of the theory and empirical work concerning pay-for-performance contracts.

**Deferred compensation :** If a CEO remains with a firm for an extended period of time, it does not imply that the firm will pay the CEO his/her marginal product every period of time. For example, suppose the CEO exerts effort, $e$ , which takes only two values , 0 or 1. $e$=0 is interpreted as shirking and $e = 1$ interpreted as the CEO exerting effort. Assume the CEO incurs the cost of making effort equal to c/2 only when he or she exerts an effort. If the CEO shirks, he or she will be caught shirking with probability equal to p. The maximum penalty associated with being caught shirking is zero compensation. Then, the firm should pay at least $w^* = c/(2p)$ in order to induce CEO to exert effort (Prendergast, 1999, P44). This wage entails a rent equal to $\frac{(1-p)c}{2p}$. Now suppose that the CEO works for two periods. Let $\delta$ be the discount factor. Then it is not necessary for the firm to pay the CEO $w^*$ for both periods. In fact, it can be shown the contract such that the firm pays the CEO $w^* = \frac{(1-p)c}{2p}$ in the second period and pays the CEO $w_1 > \frac{2}{2p} - \frac{\delta(1-p)c}{2p}$ in the first period would induce effort in both periods. This means that rent associated with the first year is set strictly less than the rent associated with second period. Intuitively, this offer induces effort in both periods since, if the CEO is caught shirking in the first period, the CEO is fired, therefore loses not only the rent associated with first period wage, but also the future rent which is set greater than the first year rent.

By providing greater rent in the second period, the firm effectively keeps the CEO exerting an effort in both periods, thus providing a longer term incentive. This feature is particulary attractive to firms since firms may want to motivate their CEOs in order to align the CEOs' interests with the firms' long run profits (Eaton and Rosen, 1983). For example, suppose that the firm has old equipment which should be replaced. Without a long run incentive, the CEOs may keep the old equipment in order to make a large profit one year, and take that record as an advertisement to get a job elsewhere (Eaton and Rosen, 1983).

Notice that stock options can be seen as deferred compensation. A standard option pricing theory predicts that it is better for the option holders to wait to exercise the stock option until the strike date in order to maximize their profit. Although some evidence of early exercises have been reported (Hemmer, Matsunaga, Shevlin, 1996), giving stock options certainly creates an incentive for the CEO to remain longer. This is combined with the fact that stock options granted to CEOs usually have vesting periods of about three years. Since un-vested options are forfeited if CEOs leave their positions, this creates additional incentive for CEOs to strive longer.

One implication is that the greater use of deferred compensation binds CEOs to their firms. However, there is no prior literature investigating the relationship between the use of stock options and the turnover probability. The closest study would be Eaton and Rosen (1983) that considers stock options as a deferred compensation to bind CEOs to the firm. They used the proportion of stock options in the total compensation to represent the form of compensation. However, their empirical study focuses on what determines the form of compensation, not the relationship between the form of compensation and turnover probability. This omission becomes one of our motivations to document the relationship between the form of compensation and turnover probability.

**Pay-for-performance contract :** The pay-for-performance contract ties the CEOs' performance with compensation and it is the most standard result from principal-agent theory. The most straightforward empirical test for the principal agent theory is to estimate the pay for performance sensitivity. Jensen and Murphy (1990, A) estimated the following equation, $\Delta(Salary + Bonus)_t = a + b\Delta(Shareholders'wealth)_t$ where b is interpreted as the pay for performance sensitivity and $\Delta$ denotes the first difference. They find that a $1000 increase in shareholders' wealth will increase CEO annual compensation only by 2.2 cents.

We can understand the CEO turnover-performance sensitivity in the context of pay for performance sensitivity if we view turnover as zero compensation. Jensen and Murphy's estimates (1990, A) suggest that the annual turnover probability of a CEO whose age is 53 would increase from 4.6% to 5.7% if the firm's net market return deteriorates from 0% to -50%. They show that, due to such an increase in turnover probability, the CEO would lose 8.6 cents for each $1000 lost by shareholders. They concluded that such pay-for-performance sensitivity is too weak to be consistent with the principal agent theory.

Hall and Liebman (1998) argue that the small pay for performance sensitivity in Jensen and Murphy (1990, A) is due to the fact that Jensen and Murphy's compensation measure does not incorporate stock option holdings. Instead, Hall and Liebman define the CEOs' wealth as the summation of salary, bonus and stock holdings. Their estimate suggests that, if the firm's performance improves from the 10 percentile level to the median level, the CEO's wealth would increase by as much as $7.6 million.

## 2.2 Managerial entrenchment and skimming

The managerial entrenchment hypothesis asserts that the separation of ownership and management gives CEOs effective control of the compensation determination process. Bertrand and Mullainathan (2001) show that under the managerial entrenchment hypothesis, CEO pay level is constrained by the unwillingness of CEOs to draw shareholder's attention (p, 902). Under the managerial entrenchment hypothesis, performance due to *observable* luck may be rewarded. On the other hand, under a simple principal agent hypothesis, rewarding the CEO for observable luck will not give him/her any incentive, and therefore paying for luck cannot be the optimal contract for the principal. Bertrand and Mullainathan (2001) use several measures for observable luck, such as an increase in oil prices, and show that CEO compensation is sensitive to lucky dollars as much as general dollars. Further, they find that such tendency is weaker for less entrenched CEOs (indicated by the presence of block-shareholders). They conclude that CEO compensation is better explained by the managerial entrenchment hypothesis.

The studies of managerial entrenchment summarized above mainly focus on the effect on compensation. It is also reasonable to assume that CEOs also have a motivation to entrench themselves from the threat of dismissal. My study also investigates the relationship between managerial entrenchment and turnover probability by investigating the effect of interlocking directorship to see if entrenchment through an interlocking directorship would reduce CEO turnover. [2] We use a dummy variable for interlocking directorship as a proxy for managerial entrenchment.

## 2.3 Matching theory

Although the matching theory is not often used to describe CEO compensation, it is one of the few theories that directly relates compensation and turnover probability.

[2]Interlocking directorship can be best described by the following situation. Suppose "CEO A" previously served as a member of a board for "CEO B" who is in a different firm. Oftentimes, we can find a situation where "CEO B" later serves as a member of the board for "CEO A". This can create close personal connections between CEO A and the member of the board (CEO B).

Jovanovic (1979) is one of the most prominent papers that contributed to the matching theory. One distinguishing feature of the matching theory is that uncertainty exits both for CEOs and shareholders. The productivity of the CEO depends on the match between the CEO and the firm: Exact information about the match is not perfectly observed by both parties.

A simple form of the matching model is the following. In the beginning of the period, the manager and the firm jointly draw a match parameter $\theta$, assumed $N(0, \sigma_0^2)$. The output of the firm at period $t$ is given by $y_t = \theta + u_t$ where $u_t$ is a noise term, assumed $N(0, \sigma_u^2)$ and independent of $\theta$. The match between the manager and the firm can only be inferred by observing the output which includes the noise term. By observing the output repeatedly over time, however, the assessment about the match becomes more precise. For each period, the wage offer to the CEO is assumed to be its conditional expected output level, $E(y_t|y_1, \ldots, y_{t-1})$. After the offer is made to the CEO, he or she decides whether or not to take this offer. If he/she does not take the offer, he/she quits and finds a new match.

The implications of matching theory are, (i) the probability of subsequent turnover is negatively correlated with the current wage rate, (ii) turnover probability increases with the riskiness of the environment $\sigma_u^2$, (iii) wage rises with tenure and (iv) turnover probability is negatively correlated with tenure.

## 2.4 Summary of the implications

First, the literature on deferred compensation suggests that the greater use of stock options in total compensation would reduce turnover probability. Pay-for-performance contract literature predicts that turnover probability is negatively related to firm performance. Third, the managerial entrenchment hypotheses implies that an increase in the extent of managerial entrenchment reduces the turnover probability. Lastly, the matching theory predicts that turnover probability decreases with the amount of compensation, increases with the riskiness of the environment and decreases with tenure. This discussion suggests that my empirical work should incorporate variables such as the returns to shareholders to proxy for the firm's performance as well as the measures to control for these possible influences. My choice of variables is described in more detail in Section 5.

# 3 Trend in CEO compensation

Our primary data sources are ExecuComp and Compustat published by Standard and Poors. ExecuComp covers detailed information about the five most highly paid executives in each company within the S&P 500, S&P Midcap 400, and S&P SmallCap 600 firms. My study focuses on compensation only of CEOs. The sample covers the years 1993 to 2003. All the compensation figures in my study are deflated by the Consumer Price Index, with year 2003 as the base year. Most of the important variables are from ExecComp, except the data on R&D spending and dividend payout which are from

Compustat.

We define total compensation as the sum of salary, bonus and the ex ante value of stock options. The value of stock options is calculated using the Black Sholes stock price formula. These three elements are chosen since they constitute the largest part of CEO compensation. They make up of nearly 90% of the whole CEO compensation package for all the sample years.

The yearly average of total compensation is seen in Table 1. The increase from 1993 to 2000 is rather dramatic. The average total compensation in 1993 is \$1.93 million. This figure nearly quadrupled to \$6.66 million in year 2000. However, average total compensation declined after year 2000. In fact, this is the year in which the information technology industry stock "bubble" burst. Table 2 shows the yearly averages for the three components separately. As can be seen, the value of stock options increases substantially to reach a peak in year 2000 and then declines sharply. Both salary and bonus show modest but steady increases through the whole sample period.

It is evident from Table 2 that the composition of compensation packages has changed significantly; the importance of stock option grants has increased dramatically. To see this more clearly, I calculate the "option mix" defined as $\frac{stock\ option}{total\ compensation}$. Table 3 shows yearly averages of the option mix. Option mix increases from about 0.3 in 1993 to about 0.5 in year 2001. This figure declines after 2001, however, with the mix of option in year 2004 still at 0.42.

# 4 Data description

Due to the 1992 requirement by the SEC, ExecuComp data contain a detailed breakdown of CEO compensation, including salary, bonuses, restricted stock grants and the Black-Sholes value of stock options. This makes ExecuComp a more attractive data set than alternative sources such as the Executive Compensation Survey published by Forbes; Forbes data does not contain stock option information.

## 4.1 Sample criteria

We treat each CEO-firm combination as a unique CEO. We also require that each individual became a CEO on or after year 1993, the year our sample period began, so that that there are no left censored observations in our sample. This is the requirement that distinguishes our sample from most of the prior studies about CEO turnover. To our knowledge, most of the studies about CEO turnover do not explicitly address the problem associated with left censoring, presumably including CEOs whose tenure started before their sample period. As is well known, such inclusion of left censored observation causes biases in the estimates. A more detailed description of our sample criteria can be found in Appendix A.

After eliminating observations that do not match our sample criteria, we obtain an unbalanced panel data set that contain 5350 CEO-years of observation including 1221 corporations and 1625 CEOs from 1993 to 2003.

## 4.2 Descriptive statistics

Table 4 shows the median amounts of salary, bonus, stock option and total compensation for the sample that passed our criteria. The median values are computed by CEO tenure. CEO tenure is the length of time a CEO holds office within a specific firm. All three components of total compensation show an increasing trend. The median level of salary is $0.5 million in the first year of tenure, increasing to $0.855 million at the 10th year of tenure. Bonus appears to be the smallest component. Median CEOs receive about $0.3 million in the first year, which increases to $0.8 million 10 years later. Stock option is unarguably the largest component of total compensation. The median CEO receives about $0.93 million worth of stock options in the first year of their tenure, and this increases to $1.56 million at the 10th year of tenure.

The median level of total compensation shows a steady upward trend in the first 10 years of CEO tenure. The median CEO receives total compensation of about $2 million dollars in the first year, increasing as much as 90% in the next 10 years. The option mix does not show an obvious trend, as can be seen from Table 4. The median level of option mix in the first year is 0.51, hovering between 0.45 and 0.5 through our sample period. The option mix trend shows that CEOs receive a considerable portion of total compensation in the form of stock options for the first 10 years of their tenure. This means that the amount that CEOs would forfeit when they are dismissed is fairly high. Needless to say, the benefits of stock options are realized only if the stock price increases. Therefore, a CEO compensation package with such a high option mix bears significant risk.

# 5 Models

Our primary objective is to document the relationship between turnover probability and the amount of compensation, and the relationship between turnover probability and the "form" of compensation. The form of compensation is represented by the proportion of stock options in total compensation. In this section we propose three different models to estimate such a relationship.

Model 1 is a simple Panel data logit discrete hazard model with the dependent variable equal to zero if the CEO does not leave after the end of the financial year, and equal to one if the CEO leaves after the end of the financial year. Model 2 and Model 3 deal with the possible endogeneity in total compensation and option mix by estimating a joint system of equations. Model 2 estimates the system of equations using a two stage method. The problem with endogeneity arises because of the presence of unobserved heterogeneity which causes correlations among the error terms in the system of equations. In Model 3, we explicitly incorporate a time invariant unobserved heterogeneity term in the system and estimate the coefficients using a maximum likelihood method.

## 5.1 Model 1: Basic Model-Panel Data logit discrete hazard model

Model 1 is a single equation panel data logit discrete hazard model written as,

$$y_{it} = \beta_0 + \beta_1 log(Total\ compensation)_{it}$$

$$+ \quad \beta_2(Option\ mix)_{it} + Z_{it}^{'}\beta + \mu_{it}$$

$$such\ that\ if$$

$$y_{it} \geq 0\ then\ leave\ the\ firm\ at\ the\ end\ of\ year\ t$$

$$y_{it} < 0\ then\ stay\ in\ the\ firm\ for\ the\ next\ year.$$

$i$ indexes each CEO and $t$ indexes the year. Total compensation is defined as the sum of annual salary, bonuses and stock options. Option mix is computed as $\frac{Options}{Total\ compensation}$. $\mu_{it}$ is an error term that is assumed standard logistic. In order to compute robust standard errors for the estimates of coefficients, I allow possible correlation among error terms within the same CEO. Nonetheless, error terms are assumed to be independent across different CEOs. $Z_{it}$ is the vector of variables that directly affects the turnover probability, but not correlated with the error term $\mu_{it}$.

There are some reasons to believe that total compensation and option mix are endogenous variables. Therefore, applying a simple logit model may result in biased estimates. This leads to considering a joint system of equations. Model 2 deals with the endogeneity by a two stage method. Model 3 deals with endogeneity by incorporating a time invariant unobserved heterogeneity term in the system.

## 5.2 Model 2: Two stage instrumental variable estimation to deal with endogeneity

Compensation may be set according to the firm's performance and may be determined endogenously. As for the form of compensation, many researchers report that firms may determines the form of compensation based on"investment opportunity sets," such as the market to book asset ratio. (i.g., Gaver and Garver,1993). Details of such discussions along with possible factors that affect total compensation and option mix are presented in section 6.

To deal with endogeneity, we consider the following system of equations.

$$\begin{aligned} Turnover\ equation &: y_{it} = \beta_0 + \beta_1 log(Total\ compensation)_{it} \\ &\quad +\beta_2 (Option\ mix)_{it} + Z_{it}^{'}\beta + \mu_{it} \quad (1) \\ Compensation\ equation &: log(Total\ compensation)_{it} = \alpha_0 + X_{it}^{'}\alpha + \varepsilon_{it}^{comp} \quad (2) \\ Option\ mix\ equation &: (Option\ mix)_{it} = \gamma_0 + X_{it}^{'}\gamma + \varepsilon_{it}^{mix} \quad (3) \end{aligned}$$

$Z_{it}$ in equation (1) is the vector of variables that directly affect turnover probability. $X_{it}$ is the vector of variables that affect the compensation and option mix. $\varepsilon_{it}^{comp}$ is assumed normal with the mean equal to zero and variance equal to $\sigma^{comp}$. $\varepsilon_{it}^{mix}$ is also assumed normal with mean zero and variance $\sigma^{mix}$. We continue to assume that $\mu_{it}$ follows standard logistic. However, we do not assume that those errors terms, $\mu_{it}$, $\varepsilon_{it}^{comp}$ and $\varepsilon_{it}^{mix}$, are independent.

Estimating the turnover equation is our main goal. The simplest way to estimate the equation is by a two-stage-method. First, we estimate compensation and option mix equations using Ordinary Least Square. Then we replace the compensation and option mix in the turnover equation with their predicted values. I refer to the total compensation and option mix equations as *the first stage equations*.

Appropriateness of the two stage method depends on the correlation between $\mu$ and $\varepsilon^{Comp}$, and a correlation between $\mu$ and $\varepsilon^{mix}$. If there is no correlation, then estimating the turnover equation with simple logit regression without the two stage method is more appropriate. This may be unlikely, however, due to the presence of unobservable variables that affect both the turnover equation and the first stage equations. For example, there may be an unobservable variable that characterizes the riskiness of the environment in which the firm operates. This variable is likely to increase total compensation (Garen, 1994), and at the same time, increase turnover probability, causing a positive correlation between $\mu$ and $\varepsilon^{comp}$. If we do not take care of endogeneity, such a positive correlation between error terms causes upward bias in the estimated coefficient for total compensation. The two stage method allows for both time invariant heterogeneity, and time varying heterogeneity that causes contemporaneous errors to be correlated.

To our knowledge, there has been no prior research examining the relationship between turnover probability and compensation. Therefore, there are no agreed upon instruments. Moreover, our second stage regression (turnover equation), involves a limited dependent variable which requires logit regression. The finite sample performance of the two stage method of this kind is studied by Bollen, Guilky and Mroz (1995). They provide researchers with simple ways to test the validity of instruments, along with other practical guidance for effective application of such two stage methods. The following summarizes their suggestions.

**Irrelevance test for excluded variable :** As for identification, the X variables in the first stage equations should contain at least one variable that is excluded from Z variables in the turnover equation. Relevance of the excluded variables should be tested. This is done by testing the null hypothesis that the coefficients for the excluded variables in the first stage equations are jointly equal to zero. If we fail to reject the null hypothesis, using two stage method would simply add noises in the second stage and thus, it is more appropriate to estimate a single turnover equation without the two stage method.

**Test for over-identification :** Over-identifying restrictions should be tested, i.e., the excluded variables should influence the turnover equation *only through* the first stage equations. A simple test that is the following: (i) replace total compensation and option mix in the turnover equation with their predicted values and (ii) put the "excluded" variables (instruments) in the turnover equation, then estimate the turnover equation using logit regression. If our exclusion restrictions are valid, the coefficients for instruments will be jointly close to zero, in which case, we say that we fail to reject the over-identification test.[3]

**Exogeneity test :** Whether these suspected endogenous explanatory variables are indeed endogenous should be tested since taking care of endogeneity when it is actually exogenous is costly in terms of precision. The simplest test is to put the predicted errors of the first stage equations in the turnover equation (without replacing total compensation and option mix with their predicted values) and estimate the turnover equation using logit regression. If the variables are actually exogenous, the coefficients of the predicted errors will be close to zero. If the coefficients are jointly *NOT* close to

[3]Note that we cannot put all the excluded variables in the turnover equation since this causes perfect multi-collinearity: If we have k excluded variables, we only put in k-2 variables. Therefore, if the model is exactly identified, we cannot test the validity of the instruments.

zero, we reject the exogeneity hypothesis.

We summarize the choice of explanatory variables along with the rationale for the choice of exclusion restrictions in section 6. It should be noted here that we include a number of lagged variables in the explanatory variable which makes it difficult to use CEOs who leave at the end of the first year since some lagged variables are not attainable for such CEOs. Therefore, the two stage method uses only individuals who stayed at least two years in the firm. This could cause a potential selection bias. Model 3 addresses one way to correct for such selection bias as well as dealing with endogeneity.

### 5.3 Model 3: Heterogeneity model

The problem associated with endogeneity arises from the correlation among error terms. The previous section dealt with this problem using a two stage method. Yet, another method is to incorporate an unobserved explanatory variable. Let $\chi_i$ be the unobserved explanatory variable for individual i. This term summarizes all the time invariant unobserved heterogeneity not captured by the observed explanatory variables. Therefore, our heterogeneity model is written as

$$\begin{aligned} Turnover\ equation \ &: \ y_{it} = \beta_0 + \beta_1 log(Total\ compensation)_{it} \\ &\quad +\beta_2 (Option\ mix)_{it} + Z'_{it}\beta + \rho_1 \chi_i + \mu_{it} \qquad (4) \\ Compensation\ equation \ &: \ log(Total\ compensation)_{it} = \alpha_0 + X'_{it}\alpha \\ &\quad +\rho_2 \chi_i + \varepsilon_{it}^{comp} \qquad (5) \\ Option\ mix\ equation \ &: \ (Option\ mix)_{it} = \gamma_0 + X'_{it}\gamma + \rho_3 \chi_i + \varepsilon_{it}^{mix} \qquad (6) \end{aligned}$$

where $\rho_1$, $\rho_2$ and $\rho_3$ are the factor loads. $\chi_i$ is assumed standard normal. $\varepsilon_{it}^{comp}$ and $\varepsilon_{it}^{mix}$ are assumed normal with mean zero, and variances $\sigma^{comp}$ and $\sigma^{mix}$ respectively. We continue to assume that $\mu_{it}$ are standard logistic. However, unlike the two stage method, we assume that those error terms are independent. We also assume that $\chi_i$ is independent of the error terms. Therefore, we have, $\varepsilon_{it}^{comp} \perp \varepsilon_{it}^{mix} \perp \mu_{it} \perp \chi_i$. In other words, the correlation in residuals are captured by $\chi_i$ when $\rho_1$, $\rho_2$ and $\rho_3$ are not all equal to zero. This method differs from Model 2 in that it only deals with time invariant heterogeneity and does not allow for time varying heterogeneity that causes contemporaneous errors to be correlated.

As stated in the previous section, Model 2 (the two stage method) uses only individuals who survived for at least two years, and hence drops individuals who left at the end of the first year. This causes a potential selection bias. To deal with such a selection problem, we incorporate a selection equation which is described as,

$$\begin{aligned} Selection\ equation : I_{it_i} &= \theta_0 + W'_{it_i}\theta + \rho_4\chi_i + \mu_{it_i}^{initial} \\ such\ that\ if & \\ I_{it_i} &\geq 0\ then\ leave\ the\ firm\ at\ the\ end\ of\ first\ year \\ I_{it_i} &< 0\ then\ stay\ in\ the\ firm\ in\ the\ next\ period \end{aligned} \quad (7)$$

where $t_i$ is the year in which the individual became a CEO. $I_{it_i}$ is the latent variable such that if it is greater than 0, the CEO exits in the first year, and if it is smaller than zero, the CEO stays in the firm. $W_{it_i}$ is a set of exogenous variables that directly affect the initial year turnover. The $\rho_4\chi_i$ term controls for the possibility of self-selection bias in year 2 and later. Therefore, our heterogeneity model becomes a system of four equations (4), (5), (6) and (7). Again, we assume that $\mu_{it}^{init}$ are independent from all other error terms.

If all the equations were linear, our model would be a standard multi-equation random effect model. However, turnover equation and selection equation are non-linear. Therefore, we estimate the system in the following way. Let $T_i$ be the year in which individual i exit the firm. If individual did not exit the firm during the sample period, $T_i = 2003$, and this individual is said to be right censored. Again $t_i$ is the year in which an individual became a CEO. Let $D_{it}^{exit}$ be the dummy variable which equals one if individual $i$ exits at year t and zero otherwise. Let $D_{it_i}^{init}$ be the dummy variable which equals one if individual i exits at the end of the first year of his or her tenure.

Since all the error terms, $\mu_i$, $\varepsilon_i^{comp}$, $\varepsilon_i^{mix}$ and $\mu_{it_i}^{init}$ are independent conditional on $\chi_i$, individual i's likelihood contribution is written as,

$$\begin{aligned} L_i(\Phi|\chi_i) &= \prod_{t\,=\,t_i+1}^{Ti} \{[1 - logit(\widetilde{Z}_{it}'\widetilde{\beta} + \rho_1\chi_i)]^{D_{it}^{exit}}[logit(\widetilde{Z}_{it}'\widetilde{\beta} + \rho_1\chi_i)]^{1-D_{it}^{exit}} \\ &\times \ \phi(log(Total\ compensation)_{it} - \widetilde{Z}_{it}'\widetilde{\alpha} - \rho_2\chi_i, \sigma^{comp}) \\ &\times \ \phi((Option\ mix)_{it} - \widetilde{Z}_{it}'\widetilde{\gamma} - \rho_3\chi_i, \sigma^{mix})\} \\ &\times \ [1 - logit(\widetilde{W}_{it_i}'\widetilde{\theta} + \rho_4\chi_i)]^{D_{it_i}^{init}}[logit(\widetilde{W}_{it_i}'\widetilde{\theta} + \rho_4\chi_i)]^{1-D_{it_i}^{init}} \qquad (8) \end{aligned}$$

*where*

$$\begin{aligned} logit(v) &= \frac{e^v}{1+e^v} \\ \phi(v,\sigma) &= \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{v^2}{2\sigma^2}) \end{aligned}$$

The term, $\widetilde{Z'_{it}}\widetilde{\beta}$, represents the observable part of equation (4). Other terms with a tilde have the same meaning. $\Phi$ is the union of all the coefficients to be estimated. To obtain the unconditional likelihood, we integrate out $\chi_i$. Unconditional likelihood contribution of individual i is give by,

$$L_i(\Phi) \quad = \quad \int_{-\infty}^{\infty} L_i(\Phi|v)\frac{1}{\sqrt{2\pi}}exp(-\frac{v^2}{2})dv \qquad (9)$$

Unfortunately, we do not have a closed form for this. Therefore, we approximate $L_i(\Phi)$ using the Gauss-Hermite approximation with 10 mass points.

$$L_i(\Phi) \quad \approx \quad \widetilde{L_i}(\Phi) = \sum_{k=1}^{10} w_k L_i(\Phi|v_k) \qquad (10)$$

where the weights $w_k$ and the support point $v_k$ are chosen using 10 points Gauss-Hermite formula.

Let N be the number of individuals in the sample. We maximize the following likelihood function over $\Phi$ to obtain the estimated coefficients.

$$L(\Phi) = \prod_{i=1}^{N} \widetilde{L_i}(\Phi) \qquad (11)$$

# 6 Choice of explanatory variables

This section briefly outlines our choice of explanatory variables along with our choice of excluded variables. We only describe selected variables. Table 5 is the list of our choice of explanatory variables. Detailed definitions for those variables can be found in Appendix A. First, we describe the choice of Z variables - the variables that are included in the turnover equation. Second, we describe the variables that affect total compensation and option mix. Third, we provide the rationale for the choice of our excluded variables. The excluded variables may be referred to as instruments. We use the terms instruments and excluded variables interchangeably.

### 6.1 Choice of Z variables: variables that directly affect turnover

Existing theories summarized in section 2 suggest that factors such as firm performance, managerial entrenchment, CEO tenure and riskiness of the environment in which the firm operates directly affect CEO turnover probability. In empirical work, it is necessary to find variables to proxy for those factors.

We use the following three variables to proxy for firm performance: (i) returns to shareholders, (ii) natural log of market to book asset ratio, and (iii) yearly percentage change in sales. The use of the market to book asset ratio requires some explanation. The book value of assets can be viewed as assets already in place, while a positive difference between market value and book value can be viewed as assets which will be in place in the future. A greater market to book asset ratio is a proxy for growth opportunity. Therefore, the market to book asset ratio can be seen as one indicator of firm performance.

Following Garen (1994) and Aggarwal and Samwick (1999) I use the R&D ratio (computed as $\frac{R\&D}{Book\ value\ of\ assets}$) and stock price volatility to proxy for the riskiness of the environment in which the firm operates. In addition, I use a dummy variable for observations with no R&D expenditure [4]. Many firms have no such expenditure, and

[4]ExecuComp reports a number of R&D observations as missing. My study assumed that whenever R&D figure is missing, $R\&D = 0$ . I believe that such assumption is reasonable and would not bias my

there may be a systematic difference between firms that make R&D expenditures and firms that do not.

We do not use an age variable. We found that the age variable in ExecuComp (p_age_2) is problematic in two ways. First, age is missing for more than 70% of the observations. Second, when we divide the sample into two sub-samples, one for observations whose age variable is not missing, and the other for observations whose age variable is missing, the average turnover rate for the sub-sample whose age variable is not missing is 0.019, whereas the average turnover rate for the sub-sample whose age variable is missing is 0.13. Therefore, incorporating age variable produces extremely small and unrealistic turnover rate. In fact, our preliminary work restricted the sample to only those with an observable age measure, leading to an unrealistically low median turnover rate of 0.026.

## 6.2 Variables that affect total compensation and option mix:

Some existing theories suggest that total compensation is influenced by factors such as firm performance, managerial entrenchment and CEO tenure. The inclusion of the industry average total compensation stems from the idea that when the board decides the compensation level, it may consider how its peer firms are paying their CEOs. It is also important to consider the timing at which compensation is determined. Typically, salary and stock option grants are determined at the *beginning* of the financial year, whereas bonuses are determined *during* the financial year. Therefore, it is likely that

estimates as many missing observation are from such industries as apparel or food industry, which are not usually considered R&D intensive industry salary and stock options depend more on the variables of the previous period. For this reason, I incorporate lags of those variables.

As for the determinants of option mix, much research suggests that a firm's "growth opportunities" may affect the composition of its CEO's compensation. Such research argues that as growth opportunities increase, the observability of managerial action decreases, increasing information asymmetry between management and shareholders. Using market based incentive plans can reduce the agency cost associated with information asymmetry. Therefore, we expect a higher proportion of stock options in total compensation for firms with abundant growth opportunities.

Common variables to proxy for growth opportunities are market to book asset ratio (Myers, 1977), R&D to book value of asset ratio (Gaver and Gaver, 1993) and percentage change in sales (Anderson, Banker and Ravindran, 2000). Dividend yield is used as the inverse indicator of growth opportunities (Anderson et al, 2000). The use of dividend yield is because growth firms tend to have greater amounts of investment expenditure and hence, lower dividend payout (Anderson et al, 2000).

## 6.3 Choice of excluded variables

Whether a variable can be excluded from the turnover equation is a matter of degree. Ideally, we want to find variables that affect the turnover equation *only through* the total compensation and option mix. In reality, such variables may be difficult to find. However, some variables may affect compensation and option mix very strongly while affecting $y$ in the turnover equation very weakly. Such a variable may be used as an excluded variables. Our choice of excluded variables are listed in Table 5. Notice that X variables are the union of all the excluded variables and the variables that directly affect turnover.

First, note that many of the lagged variables are used as excluded variables. Consider for example, exclusion of the lag of the interlocking directorship dummy. It is reasonable to assume that previous years' interlocking directorship has a weaker impact on determining whether the CEO is to be retained next year. One can imagine an extreme situation where, there was a member of the board with interlocking directorship in the *previous year*, but that person is out *this year*. If such is the case, it is difficult for this person to have influence over the CEO retention decision or CEO compensation during this year.

Other lagged variables are in the excluded category. This is due to my assumption that the current variables summarize most of the relevant information about the past. In particular, we assume that the board of directors mostly uses current information for their CEO retention decision. Therefore, I used lagged variables as excluded variables. Dividend yield is used as an excluded variable. This is due to the fact that, by construction, the returns to shareholders contain dividend information, and therefore, the dividend information is already included in the turnover equation. I validate the choice of excluded variables via the statistical test described in section 5.2.

Finally, the choice of the variable for the selection equation ($W$ variables) is presented in Table 5. Table 6 is the summary statistics of our main explanatory variables.

# 7 Main empirical results

This section summarizes our empirical results. First, we presents the estimates of selected variables for the three models. Second, we detail the effects of the amount and the composition of the CEO compensation on turnover probability. Third, we present the estimated effect of interlocking directorship on the turnover probability. Finally, using the estimated turnover probability, we compute the pay-for-performance sensitivity due to the threat of dismissal.

## 7.1 Estimation results across the models

Table 8 presents the estimated coefficients for selected variables for Model 1, the single turnover equation model, and Model 3, the heterogeneity model. We do not report Model 2 since we fail to reject Mode 1 in favor of Model 2. We failed to reject the exogeneity of total compensation and option mix, as can be seen in Table 10, even if our instruments are valid: The excluded variables appear to be significant determinants of both total compensation and option mix, as seen in Table 9. The excluded variables also appear to be valid instruments as is shown in Table 10. However, both total compensation and option mix do not appear to be endogenous, as can be seen from Table 10. Our instruments include variables such as the lagged value of R&D ratio, dividend yield, and industry average of option mix, which are reported in Table 5.

Model 1 is nested with Model 3 given the restriction that four factor loads are zero. The $\chi^2$ test of the restriction, however, rejected the hypothesis. Thus, we focus on Model

3 as being the most relevant. Notice, however, that the significant effect of heterogeneity only comes from the covariance between total compensation, option mix and the initial selection equations as can be seen from Table 8. There appears to be no correlation with the error terms in the turnover equation. In fact, the estimates for the turnover equation for Model 3 are never substantially different from Model 1. The estimated coefficients for total compensation is -0.176 for Model 1 while it is -0.246 for Model 3. Therefore, the estimated coefficient for total compensation increases by 40% in magnitude when we choose the heterogeneity model.[5] The estimated coefficient for option mix is -0.937 for Model 1 while it is -1.032 for Model 3; the coefficient increases by 10% in magnitude when we choose the heterogeneity model.

## 7.2 The effect of the amount and the structure of compensation on CEO job turnover

In this section, we report the relationship between estimated turnover probability and total compensation, and the relationship between the turnover probability and option mix. To interpret the magnitude of the estimated coefficients in an informative way, we first define our "median" CEO at his or her fourth year of tenure. This is reported in the third column of Table 11. These are the median values for our explanatory variables for CEOs at their fourth year of tenure. Our median CEO has a total compensation of about \$2.52 million, with an option mix equal to 0.476. The second and fourth columns

[5] Since the variance of composite error term $\rho_1\chi_i + \mu_{it} = \rho_1^2 + \pi^2/3$, more appropriately, we multiply the coefficients for Model 3 by $\sqrt{\frac{\pi^2/3}{\pi^2/3+\rho_1^2}} = 0.997$ to compare coefficients across the model. As can be seen, this does not make a big difference.

of Table 11 report 25 percentile and 75 percentile values for those variables for the CEOs at their fourth year of tenure.

Having defined the median CEOs, we compute the change in turnover probability when a particular variable changes from the median level to the 75 percentile or to the 25 percentile, holding all other variables constant. Table 12 reports the changes in turnover probability based on Model 3, the heterogeneity model. Upon computing the logit of leaving, we have integrated out the heterogeneity term.

Our median CEO has a turnover probability of 0.16. An increase in total compensation from the medial level to the 75 percentile level would reduce turnover probability to 0.14, while a decrease in total compensation to the 25 percentile level would increase turnover probability to 0.184. The negative relationship between total compensation and turnover probability is consistent with the matching theory.

A change in option mix also appears to have a sizable effect on turnover probability: an increase in option mix from the median level to its 75 percentile level would reduce turnover probability from 0.16 to 0.135, while, if the option mix decreases from the median level to the 25 percentile level, turnover probability increases from 0.16 to 0.19. Thus, there is a considerable negative relationship between CEO turnover probability and option mix. Such strong and negative relationship confirms our supposition that stock options are used as a deferred compensation to provide incentive for CEOs to thrive longer.

To see the magnitude of the change in turnover probability more clearly, we estimate the expected remaining years CEO would serve the firm if the turnover probability were constant over time. Table 13 summarizes the results. The expected remaining years for the median CEOs are 6.2 years. If total compensation increases from the median level to the 75 percentile level, the expected remaining years would increase to 7.1 years, an increase by 15 percent. On the other hand, if the option mix increase from the median level to the 75 percentile level, expected remaining years increases from 6.2 years to 7.4 years. This is an increase of a similar magnitude. Both total compensation and option mix appear to be significant determinants of the turnover probability.

Prior empirical work regarding the incentive effect of stock options customarily investigated how total compensation is tied to firm performance (for example, Hall and Liebman, 2000). In doing so, they simply included options in the total compensation, treating them as cash. However, this approach ignores the particular characteristics of options as deferred compensation. Our estimated negative relationship between the form of compensation and CEO turnover probability provides fresh evidence that stock options are used as deferred compensation to bind CEOs to the firm, and that this feature of stock options provides longer term incentives to CEOs.

## 7.3 Managerial entrenchment through interlocking directorship

We found little evidence of the managerial entrenchment through interlocking directorship. First, we did not find evidence of such managerial entrenchment for our main results. The estimated coefficient for interlocking directorship dummy for our heterogeneity model is 0.371, and this is statistically insignificant (see Table 8).

Such a result may indicate that the importance of the board being filled with the CEO's supporters may be minimal. For example, an interlocking directorship can occur simply because two firms are involved in a business relationship (Hallock, 1997). In such cases, the board of director with interlocking directorship would have little interest in keeping CEOs with bad performance. This case may be especially relevant for our sample since our sample contains CEOs whose tenure is ten years or less. It would be difficult for inexperienced CEOs to have discretion in choosing new board members. Note that we incorporate left-censored observations, the estimated coefficient for interlocking directorship becomes -0.291, with t-statistics equal to 2.32. This negative relationship may be spurious, resulting from biases due to left censoring.

## 7.4 Incentives generated by the threat of dismissal

Table 12 also reports the sensitivity of CEO turnover probability to the change in firm performance measured by the returns to shareholders. A change in firm performance appears to have a smaller impact on turnover probability compared to total compensation and option mix: when the returns to shareholders decrease from the median level to the 25 percentile level, the turnover probability increases from 0.161 to 0.172, and when the returns to shareholders improves from the median level to the 75 percentile level, the turnover probability would decrease only to 0.150. Such results translate to a decrease in the expected remaining years from 6.21 to 5.81 years when firm performance deteriorates from the median level to the 25 percentile level, and an increase in the expected remaining years from 6.21 to 6.67 when firm performance improves from the median level to the 75 percentile level. For either case, the changes in the expected remaining years are small, ranging between 6% to 7%.

Such a weak relationship between CEO performance and turnover probability may suggest that CEOs are not disciplined by a threat of dismissal, also confirmed by Jensen and Murphy (1990, A). Jensen and Murphy computed turnover related pay-for-performance sensitivity and showed that CEOs would lose only 8.3 cents for each $1000 lost by shareholders. However, there is a notable difference between Jensen and Murphy's results and our results; that is, our estimate for turnover probability is much higher than Jensen and Murphy's. Their estimate for turnover probability for a 53 year old CEO is 0.046 while our comparable estimate is 0.16. [6] Our estimate is larger than Jensen and Murphy's by more than a factor of 3.

Jensen and Murphy's small estimates may be partly because they did not exclude the left censored observations which would have biased their estimate downwards. It could also be that their estimate was contaminated with a similar problem with age variables, which is already discussed in section 6.1. Due to the problem with the age variable in ExecuComp, inclusion of an age variable produces unrealistically low turnover rate. We suspect that Jensen and Murphy might have underestimated the turnover related pay-for-performance sensitivity. Therefore, we reexamine the turnover related pay-for-performance sensitivity based on our estimated turnover probability.

[6] Since the median age for CEOs with 4 years of experience is 50 years old, we believe that the our estimated turnover probability is comparable to Jensen and Murphy's figure.

First we compute the upper bound of wealth losses when our median CEO is dismissed at the fourth year of tenure. We assume that the CEO's wage profile would follow the median values presented in the Table 4 and that the CEO would have no alternative employment opportunities for five years.[7] This means that the CEO's wealth losses when he or she is dismissed at the fourth year of tenure is the present discount value (PDV) of the lost total compensation from the fifth year of tenure to the tenth year of tenure plus present discounted value of forfeited stock options. We assume that all the unvested options are forfeited and that all the stock options vest in three years. Therefore, the wealth losses are written as

$$\begin{aligned} & Wealth\ Losses \\ = \ & PDV\ of\ cumulative\ income\ from\ 5th\ tenure\ until\ 10th\ tenure \\ + \ & PDV\ of\ stock\ options\ granted\ from\ 2rd\ tenure\ to\ 4th\ tenure \end{aligned} \tag{12}$$

Using the discount value of 3%, the upper bound of wealth losses if the median CEO is dismissed are \$20.65 million, with PDV of cumulative income equal to \$17.64 million and PDV of forfeited stock options equal to \$3.01 million.

Table 15 summarizes the expected wealth losses due to the threat of dismissal at different levels of firm performance. Expected value of wealth losses at the fourth year of tenure with a given turnover probability are computed as $(Wealth\ losses\ if\ the\ CEO\ ) \times (turnover\ \ probability)$. When the returns to shareholders are at the median level, the

[7]Our assumptions differ from Jensen and Murphy's. They assumed that the CEO would have no alternative job from age 53 until 65, and that they would receive \$1 million until age of 65 if they were not dismissed.

expected wealth losses are \$3.32 million. When the firm performance is at the 25 percentile level, the expected wealth losses are \$3.55 million. Therefore, a deterioration in firm performance from the median level to 25 percentile level would cause an increase in the expected CEO wealth losses by \$0.23 million.

Now, we compute the pay for performance sensitivity due to the threat of dismissal. Since we could not tell if a CEO is fired or quit, the terminology, dismissal, is used as a shorthand of turnover. The median of the firm value for our sample (when CEOs are at the third year of tenure) is about \$1.74 billion. Therefore, when the returns to shareholders deteriorate from the median level to the 75 percentile level, shareholders would lose about \$426.56 million.[8] This means that CEO would lose 56.9 cents for each \$1000 lost by shareholders. This estimate is considerably higher than the comparable figure estimated by Jensen and Murphy (1990, A): Jensen and Murphy's estimate is that the CEO would loose only 8.6 cents for each \$1000 lost by share: Our estimate of turnover related pay-for-performance sensitivity is greater than Jensen and Murphy's estimate by factor of 7.

Again, this difference may be from the fact that Jensen and Murphy did not exclude the left censored observations, or it may be because of a similar problem with the age variable which we found in ExecuComp data set. While we do not attempt to test

[8] Because of the definition of returns to shareholders, we can compute the change in shareholders' wealth, which is defined as $(frim\ value)_t - (frim\ value)_{t-1} + (dividend\ payout)_t$, as $(Return\ to\ shareholders) * (Firmvalue)_{t-1}$. The median level of the returns to shareholders is 0.028, making the change in shareholders' wealth \$102 million. 25 percentile level of the return is -0.189, making the change in shareholders' wealth -\$427.56 million. Therefore, a worsening of the return from the median level to the 25 percentile level would have lost shareholders' wealth by \$426.56 million.

whether our estimated pay-for-performance sensitivity is consistent with the principal agent theory, it may be the case that previous literature grossly underestimated the turnover related pay for performance sensitivity.

# 8 Sensitivity of estimates to the inclusion of left censored observations

In order to avoid biases arising from left censoring , our main estimation uses the sample that contains only the CEOs who became CEOs after our sample period began. Much of the prior research regarding CEO turnover, however, does not address such an issue. In this section, we present the sensitivity in the estimated coefficients when we included the left censored observations.

Inclusion of left censored observations are problematic. For example, suppose we have two types of CEOs, one with a high propensity for turnover and the other with a low propensity of turnover. Further, suppose that they started their tenure prior to our sample period, say 1988. Then, CEOs who have a high propensity of turnover may not survive until the beginning our sample period, 1993, and only the CEOs with a low propensity of turnover survive until our sample period. This means that if we include left censored observations, we artificially increase the number of individuals who have a lower propensity of turnover, causing bias in our turnover probability estimates. In this section, we illustrate the sensitivity in the estimates when we include the left censored observations.

Table 12 shows the estimated coefficient for Model 3 for the sample that excludes left censored observations, and for the sample that includes left censored observation. The coefficient for total compensation for the sample without left censored observation is -0.246 while it is -0.111 for the sample with left censored observations. Inclusion of left censored observations lowers the estimate in magnitude by more than 50%. The coefficient for option mix is -1.03 for the sample without left censored observations while it is -0.8 for the sample with left censored observations. The coefficients for returns to shareholders decreases almost by 50% in magnitude. Therefore, inclusion of left censoring severely underestimate the sensitivity of the logit of leaving to the change in total compensation, option mix and returns to shareholders.

The coefficient for interlocking directorship changes the sign when we include the left censored observations. The coefficient for interlocking directorship is 0.37 for the sample without left censored observation, while it is -0.29 for the sample with left censored observations. However, we are not able to conclude whether this negative relationship is due to the true effect of an interlocking directorship.

In sum, we found that inclusion of left censored observation underestimates the effect of total compensation, option mix and returns to shareholders. In addition, the sign of coefficients for interlocking directorship becomes negative, though this negative effect of interlocking directorship may be spurious. The prior studies regarding CEO turnover mostly investigated the relationship between the firm performance and CEO turnover (i.e., Warner, Watts and Wruck, 1988). However, since such studies have ignored the issues about left censoring, they might have grossly underestimated the effect of firm performance on CEO turnover.

# 9 Conclusion

This paper contributed to the CEO compensation literature by providing a link between the form of CEO compensation and CEO turnover probability, and the link between the amount of the compensation and turnover probability. The form of compensation was represented by the proportion of stock options as part of CEO compensation. We estimated these relationships by using a joint equations model that incorporates a time invariant unobserved explanatory variable. We found that both the amount and the form of total compensation have significant impact on CEO turnover probability.

We found that an increase in the *proportion* of options in total compensation from the median level to the 75 percentile level would decrease the turnover probability from 16% to 13.5%, while an increase in the *amount* of total compensation from the median level to the 75 percentile level would decrease turnover probability from 16% to 14%. This would mean that, if the probability of separation were constant over time, an increase in the proportion of option mix from the median level to the 75 percentile level would increase the expected remaining lifetime of CEO from 6.2 years to 7.4 years (a 19% increase), and the increase in total compensation from the median level to the 75 percentile level would also lead to a change of a similar magnitude. The significantly negative effect of option mix is consistent with the view that stock options are used as a deferred compensation that provides longer incentive to CEOs.

Our study finds little support for the managerial entrenchment hypothesis through interlocking directorship and consistent evidence for matching hypothesis: The presence of interlocking directorship did not appear to have negative effect on turnover probability for CEOs whose tenure is less than 10 after controlling for the potentially severe biases due to left censoring.

Although changes in firm performance has a relatively small effect on the turnover probability, the turnover related pay-for-performance sensitivity is greater than the comparable figure estimated by Jensen and Murphy (1990, A). Our estimated turnover probability implies that the median CEOs with 4 years of experience would lose 56.9 cents for each $1000 lost by shareholders. This is greater than a comparable figure estimated by Jensen and Murphy's by factor of 7: They estimated that CEOs would lose only 8.6 cents for each $1000 lost by shareholders. Finally, we found that estimated coefficients are very sensitive to the inclusion of left censored observations. Inclusion of left censored observations severely underestimates the effect of total compensation, option mix and the returns to shareholders. This data issue may have lead other researchers to underestimate the turnover related pay-for-performance sensitivity.

Table 1: Yearly average of total compensation

| Year | # of obs | Average total comp |
|---|---|---|
| 1993 | 1330 | 1.93 |
| 1994 | 1495 | 2.26 |
| 1995 | 1572 | 2.21 |
| 1996 | 1639 | 2.89 |
| 1997 | 1699 | 3.59 |
| 1998 | 1627 | 4.09 |
| 1999 | 1653 | 5.36 |
| 2000 | 1656 | 6.66 |
| 2001 | 1590 | 5.91 |
| 2002 | 1575 | 4.21 |
| 2003 | 1494 | 3.46 |

Compensation figures are in million dollars

Table 2: Yearly average of salary, bonuses and options

| Year | # of obs | Salary | Bonus | Option |
|---|---|---|---|---|
| 1993 | 1330 | 0.63 | 0.50 | 0.79 |
| 1994 | 1495 | 0.63 | 0.55 | 1.09 |
| 1995 | 1572 | 0.63 | 0.56 | 1.03 |
| 1996 | 1639 | 0.63 | 0.67 | 1.59 |
| 1997 | 1699 | 0.62 | 0.66 | 2.31 |
| 1998 | 1627 | 0.62 | 0.65 | 2.81 |
| 1999 | 1653 | 0.63 | 0.75 | 3.98 |
| 2000 | 1656 | 0.64 | 0.78 | 5.24 |
| 2001 | 1590 | 0.65 | 0.67 | 4.60 |
| 2002 | 1575 | 0.66 | 0.70 | 2.84 |
| 2003 | 1494 | 0.67 | 0.83 | 1.95 |

Compensation figures are in million dollars

Table 3: Yearly average of Option mix

| Year | # of obs | Average option mix |
|---|---|---|
| 1994 | 1495 | 0.29 |
| 1995 | 1572 | 0.28 |
| 1996 | 1639 | 0.33 |
| 1997 | 1699 | 0.37 |
| 1998 | 1627 | 0.44 |
| 1999 | 1653 | 0.47 |
| 2000 | 1656 | 0.47 |
| 2001 | 1590 | 0.51 |
| 2002 | 1575 | 0.48 |
| 2003 | 1494 | 0.42 |

Table 4: Median compensation by tenure

| CEO tenure | # of obs | Salary | Bonus | Options | Total | Option Mix |
|---|---|---|---|---|---|---|
| 1 | 1519 | 0.504 | 0.302 | 0.930 | 1.991 | 0.515 |
| 2 | 1210 | 0.602 | 0.390 | 0.842 | 2.000 | 0.453 |
| 3 | 928 | 0.641 | 0.457 | 0.930 | 2.186 | 0.446 |
| 4 | 640 | 0.688 | 0.452 | 1.092 | 2.525 | 0.476 |
| 5 | 432 | 0.713 | 0.494 | 1.137 | 2.598 | 0.495 |
| 6 | 269 | 0.753 | 0.543 | 1.351 | 2.950 | 0.486 |
| 7 | 168 | 0.797 | 0.689 | 1.365 | 3.141 | 0.470 |
| 8 | 104 | 0.849 | 0.519 | 1.916 | 3.593 | 0.574 |
| 9 | 52 | 0.911 | 0.736 | 1.552 | 3.518 | 0.494 |
| 10 | 20 | 0.855 | 0.829 | 1.565 | 3.860 | 0.429 |

Compensation figures are in million dollars

Table 5: Choice of independent variables and exclusion restrictions

| Variables that directly affect turnover (Z variables) | Excluded variables |
|---|---|
| $(Returns\ to\ shareholders)_t$ | $(Inter\ locking\ directorship\ dummy)_{t-1}$ |
| $log(Market\ to\ book\ asset\ ratio)_{t-1}$ | $(\%\ change\ in\ sales)_{t-1}$ |
| $\Delta log(market\ to\ book\ asset\ ratio)_t$ | $log(dividend\ yeild + 1)_t$ |
| $(\%\ change\ in\ sales)_t$ | $log(dividend\ yeild + 1)_{t-1}$ |
| $(Interlocking\ directirship\ dummy)_t$ | $(\frac{R\&D}{assets})_{t-1}$ |
| tenure | $Dummy\{R\&D = 0\}_{t-1}$ |
| $(tenure)^2$ | $(industry\ average\ option\ share)_t$ |
| $log(sales)_t$ | $(industry\ average\ option\ share)_{t-1}$ |
| $(\frac{R\&D}{assets})_t$ | $(industry\ average\ total\ comp)_t$ |
| $Dummy\{R\&D = 0\}_t$ | $(industry\ average\ total\ comp)_{t-1}$ |
| log(stock price volatility) | |
| Year dummies | |
| industry dummies | |

Table 6: Choice of variables for initial turnover equation

| Variables |
|---|
| $log(Market\ to\ book\ asset\ ratio)_t$ |
| $\Delta log(market\ to\ book\ asset\ ratio)_t$ |
| $(\%\ change\ in\ sales)_t$ |
| $(Interlocking\ directirship\ dummy)_t$ |
| $log(sales)_t$ |
| $(\frac{R\&D}{assets})_t$ |
| $Dummy\{R\&D = 0\}_t$ |
| log(stock price volatility) |
| $log(dividend\ yeild + 1)_t$ |
| Year dummies |
| industry dummies |
| $(industry\ average\ total\ comp)_t$ |
| $(industry\ average\ option\ share)_t$ |
| Outside succession dummy |
| S&P 500 dummy |
| S&P midcap dummy |
| earnings per share dummy |
| number of board meeting during the year |

Table 7: Data summary statistics (1993 - 2003)

| Variable | # of obs | Mean | St div | Min | Max |
|---|---|---|---|---|---|
| $The returns to shareholders$ | 3831 | 0.34 | 9.45 | -1.0 | 582.9 |
| Market to book asset ratio | 5350 | 2.07 | 2.09 | 0.36 | 78.6 |
| $\%sales change$ | 5350 | 10.8 | 42.08 | -96.7 | 1533.1 |
| Interlocking directorship dummy | 5350 | 0.04 | 0.20 | 0 | 1 |
| Tenure | 5350 | 2.96 | 1.96 | 1 | 11 |
| $\frac{R\&D}{Assets}$ | 5350 | 0.03 | 0.07 | 0 | 1.18 |

Table 8: Estimation results by models

| | Model 1 | Model 3 |
|---|---|---|
| log(Total compensation) | -0.176 | -0.246 |
| | (0.084) | (0.140) |
| Option share | -0.937 | -1.032 |
| | (0.299) | (0.323) |
| Return to Shareholders | -0.298 | -0.300 |
| | (0.176) | (0.122) |
| Interlocking directorship dummy | 0.349 | 0.371 |
| | (0.317) | (0.273) |
| $\rho_1$ (turnover equation) | – | 0.148 |
| | | (0.153) |
| $\rho_2$ (Total comp equation) | – | 0.749 |
| | | (0.024) |
| $\rho_3$ (Option mix equation) | – | 0.172 |
| | | (0.006) |
| $\rho_4$ (Selection equation) | – | -1.214 |
| | | (0.461) |

a. Inside the brackets are robust standard errors.
b. BGM results show that total compensation and option mix are exogenous, we do not use first stage equations in Model 2. Therefore Mode 1 and Model 2 are identical

Table 9: Irrelevance test for excluded variables

| Equations | Test stat | $Adjusted R^2$ | Decision |
|---|---|---|---|
| log(total compensation) | 4.56 | 0.449 | Reject irrelevance |
| | (1.83) | | |
| Option share | 5.12 | 0.219 | Reject irrelevance |
| | (1.83) | | |

Table 10: Tests for over-identification and exogeneity

| | Joint test for Total compensation and Option Mix | |
|---|---|---|
| Equations | Test stat (5% Crit val) | Results |
| Over-identification | 3.75 (15.51) | Fail to reject |
| Exogenous test | 2.48 (5.99) | Fail to Reject |

Table 11: The median, the 25 percentile and the 75 percentile values for CEOs at their 4th year of tenure

| Variables | 25th percentile | Median | 75th percentile |
|---|---|---|---|
| Total compensation (in million) | 1.31 | 2.52 | 5.00 |
| Option share | .279 | .476 | .672 |
| Returns to shareholders | -.189 | .059 | .340 |
| Log(Market to book asset ratio) | 0.203 | .484 | .484 |
| $\Delta log(Market\ to\ book\ asset\ ratio)$ | -.158 | -.006 | .121 |
| $\%changeinsales$ | -1.45 | 6.46 | 15.73 |
| Interlocking directorship dummy | – | 0 | 1 |
| tenure | | 4 | |
| tenuresq | | 16 | |
| Sales (in million) | 0.63 | 1.78 | 5.13 |
| R & D ratio | 0 | .0031 | .0341 |
| R & D zero dummy | 1 | 0 | 0 |
| log(Volatility) | -1.28 | -1.00 | -.670 |

Table 12: Sensitivity of CEO turnover probability due to changes in selected variables

| | Model 3: Heterogeneity Model | | |
|---|---|---|---|
| Variables | 25th percentile | 50th percentile | 75th percentile |
| Total compensation | 0.184 | 0.161 | 0.140 |
| Option mix | 0.191 | 0.161 | 0.135 |
| returns to Shareholders | 0.172 | 0.161 | 0.150 |

Table 13: Remaining expected lifetime from 5th year of tenure

| | Model 3 (Preferred Model) | | |
|---|---|---|---|
| Variables | 25th percentile | 50th percentile | 75th percentile |
| Total Compensation | 5.43 years | 6.21 years | 7.14 years |
| Option share | 5.24 years | 6.21 years | 7.41 years |
| $\Delta(shareholders'wealth)$ | 5.81 years | 6.21 years | 6.67 years |

Table 14: Irrelevance test for excluded variables

| CEOs with 25 percentile performance | | CEOs with 25 percentile performance | | | |
|---|---|---|---|---|---|
| Turnover probability | Expected wealth losses | Turnover probability | Expected Wealth losses | Difference in expected wealth losses | Pay-performance sensitivity from the median to the 25th percentile performance |
| 0.172 | $3.55 million | 0.161 | $3.32 million | $0.23 million | 53.9 cents per $1000 |

Table 15: Sensitivity of estimates to the inclusion of left censored observations

| | A sample without left censored obs | A sample with left censored obs |
|---|---|---|
| Variables | Model 3 | Model 3 |
| Total compensation | -0.246 | -0.111 |
| | (0.140) | (0.068) |
| Option mix | -1.032 | -0.803 |
| | (0.323) | (0.170) |
| Returns to shareholders | -0.300 | -0.170 |
| | (0.122) | (0.058) |
| Interlocking directorship | 0.371 | -0.291 |
| | (0.237) | (0.126) |
| $\rho_1$ (turnover equation) | 0.148 | -0.009 |
| | (0.153) | (0.066) |
| $\rho_2$ (Total comp equation) | 0.749 | 0.64 |
| | (0.024) | (0.006) |
| $\rho_3$ (Option mix equation) | 0.172 | 0.161 |
| | (0.006) | (0.003) |
| $\rho_4$ (Selection equation) | -1.214 | – |
| | (0.461) | – |

## Chapter II

## An alternative theory of the relationship between tenure and incentive in a principal-agent setting.

# 1 Introduction

One of the standard results of principal agent theories is that pay-for-performance sensitivity increases with tenure as CEO, but this has not been supported by prior empirical studies. Murphy (1986) uses CEO compensation data between 1974 and 1984 and finds that pay-for-performance sensitivity is smaller for CEOs whose tenure is greater than 4.6 years than for CEOs with shorter tenure. This result means that pay-for-performance sensitivity decreases with CEO tenure, a result which contradicts the standard principal-agent theories. Gibbons and Murphy (1992) also use CEO compensation data between 1971 and 1989 and similarly find that pay-for-performance sensitivity decreases as CEO tenure increases, contradicting standard principal-agent theories.

Such results leave us unsure if CEO compensation is consistent with standard-principal agent theories and raise a question about whether standard principal-agent theories like Gibbons and Murphy's (1992) accurately describe CEO compensation. Since prior literature consistently finds a *negative* relationship between incentive and tenure, we find it worthwhile to develop a model that hypothesizes a negative relationship between incentive and tenure.

Our model is an extension of Prendergast's model (2002), which shows that there is a *positive* relationship between incentive and uncertainty. In his model, uncertainty dictates if the principal chooses "input-based compensation" or "output-based compensation". For input-based compensation, the agent is rewarded by his or her input, i.e., if he or she is keeping busy. On the other hand, output-based compensation ties compensation to output, i.e., the firm's value. Therefore, output based compensation is an incentive intensive compensation. Prendergast shows that under greater uncertainty, the principal is more likely to choose output-based compensation, demonstrating the positive relationship between incentive and uncertainty.

Our study's reasoning is simple. Uncertainty regarding the outcome of the business stems not only from the business environment but also from uncertainty regarding the CEO's ability. The CEO's ability becomes gradually known to the principal over time, thus decreasing the principal's uncertainty of the business outcome. Then, assuming Prendergast result, we see that the principal is more likely to choose input-based compensation for CEO's with longer tenure because of the reduced uncertainty. This suggests a negative relationship between incentive and tenure.

## 2 Prior literature

Murphy (1986) characterizes the incentive contract for CEOs in a multi-period principal-agent setting. In his model, the agent works for a fixed number of periods. His model implies that pay increases are spread evenly throughout the remaining years, i.e., good performance is rewarded as a shift in the earning-tenure profile. This in turn implies that pay-for-performance sensitivity increases with tenure because for later years, the reward is spread among fewer years. Thus the same increase in output causes greater increase in annual compensation.[9] The econometric model is $\Delta ln(Salary + Bonus)_{it} = \alpha + \beta(Years\ as\ CEO) + \delta(Rate\ of\ return\ on\ common\ equity)_{it} + \epsilon_{it}$. Murphy uses CEO compensation data between 1974 and 1984. He divides the data into three sub-samples; the first sub-sample is for CEOs with tenure of 4.6 and less, the second sub-sample is for CEOs with tenure between 4.6 and 9.9 and the third sub-sample is for CEOs with tenures greater than 9.9. Murphy finds that $\delta$ is smaller for those with longer tenure. Thus the estimated results show that pay-for-performance actually decreases with tenure, which is inconsistent with the standard results of principal-agent theories. [10]

Gibbons and Murphy (1992) develop a model that characterizes the incentive contract when the agent has reputational career concern. In their model, good current performance positively affects the CEO's assessed ability, as assessed by the shareholders. The assessed ability is interpreted as the "reputation" of the CEO. A good reputation causes an increase in his or her future compensation, thus creating incentives for CEOs to exert effort even without an explicit incentive contract. However, the reputational concerns disappear as the CEO reaches his or her retirement. To supplement such a loss

[9]Therefore, the implication that pay-for-performance increases with tenure depends on the finiteness of a CEO's lifetime.

[10]They conclude that those results are supported by learning model instead.

in incentive, firms may intensify pay-for-performance sensitivity as the CEO approaches his or her retirement. Gibbons and Murphy's model has two implications: (i) Holding CEO tenure constant, pay-for-performance sensitivity increases as CEOs reach their retirement. (ii) Holding years to retirement constant, pay-for-performance sensitivity increases as CEO tenure increases. The second implication is relevant for our study.

The reason for the second implication is as follows. First, because of the agent's risk aversion, the standard trade-off between incentive and uncertainty holds, i.e., pay-for-performance sensitivity decreases with the uncertainty of business performance. Suppose that the firm's output is given by $y_t = \eta + a_t + \epsilon_t$ where $\eta$ is the unobservable ability of a CEO (unobservable both by CEOs and the shareholders), $a_t$ is the effort and $\epsilon_t$ is the usual independent random shocks. Notice that there are two sources of uncertainty to both shareholders and the CEO, $\eta$ and $\epsilon_t$. Although $\eta$ is not directly observable (both by the shareholders and by the CEO), it can be inferred by repeatedly observing the firm's output over time. Because the uncertainty about the ability of the CEO decreases over time, so does the overall uncertainty of the business outcome. By the inverse relationship between incentive and uncertainty, Gibbons and Murphy show that pay-for-performance sensitivity should decrease with tenure.

Their econometric model to test the pay-for-performance incentive relationship is the following. $\Delta(Salary + Bonuses)_{it} = \beta_1(low\ tenure\ dummy) \times \Delta ln(Firm\ Value)_{it} + \sum_{n=1972}^{1988} \beta_n(n^{th}\ year\ dummy_{it}) \times \Delta ln(Firm\ value)_{it} + \beta'(other\ variables)_{it}$. The theory predicts that $\beta_1$ is negative, reflecting the negative relationship between pay-for-performance sensitivity and CEO tenure. They used CEO compensation data between 1971 and 1989 to estimate this model. They find a positive but statistically insignificant coefficient. The implication of their theory that incentive and uncertainty have a positive relationship is thus not strongly supported by the data. Nonetheless, they finds evidence that pay-for-performance sensitivity increases as the agent reaches his or her retirement, although the magnitude of the increase in pay-for-performance sensitivity is surprisingly small. For the CEOs with cash earning of \$562,000, 10% change in shareholders wealth corresponds to \$7,300 of CEO 's wealth change for CEOs more than three years from retirement, but \$9,500 for CEOs fewer than thee years to retirement.

Note again that the basic reason why pay-for-performance increases with tenure in Gibbons and Murphy (1999) is because of the standard inverse relationship between incentive and risk. Contrary to such standard results, Prendergast (2002) shows that under some conditions pay-for-performance sensitivity is non-decreasing with uncertainty. In his model, uncertainty dictates if the principal chooses "input-based compensation" or "output-based compensation". For input-based compensation, the agent is rewarded by his or her input, i.e., if he or she is keeping busy. On the other hand, output-based compensation ties the compensation to the output, i.e., the firm's value. It is assumed that making output-based compensation incurs a fixed cost to the firm so firms want to avoid output-based compensation if necessary. Assuming that the agent observes the business environment better than the principal, if the uncertainty of the business environment is significant, the principal delegates the decision making power to the agent. However, to keep the agent from abusing the delegated power, the principal uses output-based compensation. In other words, under greater uncertainty, the benefit of delegating decision making power to the agent exceeds the cost of making output-based compensation. If the business environment is less uncertain, the principal keeps the decision making power to himself or herself and uses input-based compensation. Since output-based compensation ties compensation to output, we are likely to observe greater pay-for-performance sensitivity when uncertainty is large.

Tadelis (2002) shows that incentive from a reputational concern is ageless: The optimal incentive contract is the same for both younger and older workers. In Tadelis' model, reputation is a tradeable good. Suppose that the agent works for two periods, then retire. At the beginning of the period, the agent buys a "name" by working for a company. Upon retirement, he or she sells the name. In the first period, the reputation affects the second period income which creates an incentive. In the second period, the reputation determines the price at which the agent can sell the name, creating continued concerned for reputation. The author shows that the incentive in the first period and the second period are quantitatively equivalent. Tadelis is mainly concerned with the relationship between incentive and the years to retirement, thus his task is not to show the relationship between incentive and tenure. However, if all agents are assumed to begin their tenure as CEOs at the same ages, his model suggests that pay-for-performance sensitivity is constant with regard to tenure.

# 3 Basic intuition of our model

The basic result of standard principal-agent theories like Gibbons and Murphy (1992), that a positive relationship exists between incentive and tenure, is derived from the standard inverse relationship between incentive and uncertainty. Since the principal's uncertainty about the agent's ability decreases over time, the inverse relationship between incentive and uncertainty leads to the conclusion that the incentive intensity of compensation is higher for an agent with longer tenure.

Our alternative theory shows otherwise that the incentive intensity of the compensation package increases with tenure. We borrow the framework from Prendergast (2002) which shows that incentive intensity of compensation *decreases* with uncertainty. Suppose that the principal hires an agent to perform one of n tasks. The principal makes one of the following two types of contracts; Input based-contracts or output-based contracts. For input based contracts, the principal chooses a task that the agent should perform and pay according to the agent's input (how busy the agent keep himself of herself). For output-based contracts, the principal delegates the task-choosing decision to the agent, but to keep the agent from abusing the delegated power, the principal ties compensation to the output: Delegation of decision making power always comes with an output-based contract. Since output-based contracts tie compensation to output, it is an incentive-intensive compensation.

Prendergast makes two assumptions: (i) The agent knows better than the principal about how the business should be conducted. (ii) Making an output-based contract is more costly than making an input-based contract for the principal. Under such conditions, the principal faces the following trade-offs: First, the principal will want to delegate the decision to the agent (since the agent knows the business environment better), but making output-based contract is expensive. Second, it is cheaper to make input-based contracts but the principal faces the risk of choosing the wrong task. Note that the more uncertain the environment, the more likely it is that the principal makes the wrong decision. Thus, under large uncertainty, the benefit of delegating the decision to the agent is more likely to exceed the cost of output-based contracts. This means that the principal is likely to choose output-based compensation when uncertainty is large. On the other hand, if the uncertainty about the business outcome is small, the principal is likely to choose the correct task. Thus, cost of output-based compensation is likely to outweigh the benefit of delegating the decision making power to the agent. This means that when uncertainty is small, the principal is more likely to choose input based compensation. Since output-based compensation is the incentive intensive compensation, a positive relationship exist between incentive and uncertainty.

We extend Prendergast' model (2002) to a multi-period setting to show that incentive decreases with tenure. Uncertainty regarding the outcome of the business stems not only from the business environment but also from uncertainty regarding the CEO's ability. The CEO's ability is gradually revealed over time, decreasing the overall uncertainty of the business outcome. Using Prendergast's result, we show that the principal is more likely to use output-based compensation earlier in the agent's tenure. This suggests a negative relationship between tenure and incentive.

We should carefully interpret our results. Our model indicates that firms are more likely to use input-based compensation for agents with longer tenure, which implies that the agent with longer tenure has less discretion over the choice of tasks. One possible criticism is that it is more intuitive to assume that an agent with longer tenure has more discretion over what tasks he or she performs. In fact, it is not unreasonable to assume that an agent becomes entrenched over time so as to avoid the disciplinary mechanisms of principal-agent relationship. In such cases, the agent gains more discretion over what he or she performs as tenure increases. However, such entrenchment considerations are beyond the focus of our model. As long as the employment contract works like the principal-agent model, and as long as the agent does not become entrenched as tenure increases, it is not necessary that the principal gives more discretion to the agent simply because the agent has a long tenure. We can also understand the input-based compensation contract as the principal and agent mutually agreeing to undertake a certain task.

# 4 Settings

The principal employs an agent to perform one of n possible tasks. At the beginning of the period, the principal and the agent jointly draw match parameters $\theta_i, i = 1, \dots n$ for each task. $i$ denotes the task. $\theta_i$ is interpreted as a task specific ability of the agent. For simplicity, we assume that these match parameters are distributed normally and independently but possibly with different means and variances. The distribution of those parameters are assumed to be common knowledge for both the principal and the agent.

When the agent undertakes task i at period t, the output depends on the match parameter $\theta_i$, random variable $\rho_{ti}$ and the effort $e_{ti}$. The output at period t is assumed to have the following form.

$$y_{it} = \theta_i + \rho_{it} + e_{it}$$

The randomness of $\rho_{it}$ capture the riskiness of task i at period t. For simplicity, we assume that $\theta_i$, $\rho_{jt}$ and $e_{it}$ are independent for all i, j, and t. The distribution of $\rho_{it}$ are common knowledge for both the principal and the agent. The distribution is task specific and time-invariant. For simplicity, we assume that they are drawn from normal distributions possibly with different means and variances. The cost of effort is given by $C(e_{ti})$, which is assumed to be a strictly convex and monotonically increasing function.

The basic premise of the model is that the agent has more information about the realized values of the match $\theta_i$ and task specific uncertainty $\rho_{ti}$ than the principal. We assume that the agent can directly observe the realized values of $\rho_{ti}$ and $\theta_i$ but the principal can only observe the realized value of output $y_{it}$. The principal can only infer the realized value of $\theta_i$ by repeatedly observing output over time. For each period, a wage contract is made before the agent sees the realized value of $\rho_{ti}$. The principal's problem is not only to induce the correct effort but also to induce the correct action, or to assign the correct task to the agent. The agent is assumed to have personal preferences for tasks he or she enjoys most. $B_i$ denotes the private benefit of doing the $i^{th}$ task.

Both the principal and the agent are assumed to be risk neutral. Let $u(w) = w$ be the utility function of the agent where $w$ is the total compensation. This is the same assumption as Prendergast (2002). The reason why we assume the agent's risk neutrality is to isolate the effect of uncertainty on the principal's decision to choose an incentive contract from the standard trade-off between incentive and uncertainty arising from the agent's risk aversion.[11]

The principal chooses either an input-based contract or an output-based contract. For an input-based contract, the principal chooses a task that the agent should perform, and rewards the agent depending on the effort that he or she exerts. (i.e., whether the agent is keeping busy ). For an output-based contract, the principal lets the agent choose the task and reward the agent based on the output. Since an output-based contract ties compensation to output, it is incentive-intensive compensation. When making an output-based contract, the principal incurs a fixed cost $m_y$, and when making an input-based contract, the principal incurs a fixed cost $m_e$. It is assumed that $m_y > m_e$. The fixed cost $m_y$ is the manifestation of the cost associated with making an incentive-contract in a multi-task setting. [12]

For the most part, we assume that once the agent is hired he or she will be hired

[11]Standard trade-off between incentive-pay and uncertainty is the following. Assume that agent is risk-avese. Under greater uncertainty, incentive intensive compensaiton should be accompanied by higher expected compensation in order to satisty the participation constraint of the agent. Therefore, to principal, high-powered incentive plan comes at cost of higher expected pay. Thus, if uncertainty is too high, the principal reduces the incentive-intensity of the compensation package.For example, see Aggarwal and Samwick (1999).

[12]Holmstrom and Milgrom (1991)argues that, since complex jobs have many dimensions of performance, some of which are poorly observed, rewarding on the observed dimensions can have harmful effects on the unobserved dimensions, making it more costly to make output-based contract.

forever. We relax this assumption by later introducing turnover into the model. Introduction of turnover does not alter our main result that incentive decreases with tenure.

The following part of the paper is organized as follows. In section 5, we present the simplest model where matching parameters are excluded from the model. This section is intended to clear up why we have a positive relationship between incentive and uncertainty. Section 6 introduces matching parameters into the model. This section shows that incentive is non-increasing with tenure. Section 7 incorporates turnover into the model.

# 5 The simplest case: Without match parameters

In this section, we present the simplest case in which the model does not include match parameters. This is to clear up the reason why the principal chooses output-based compensation when he or she faces large uncertainty. Output is given by

$$y_{ti} = \rho_{ti} + e_{ti}$$

Without match parameters, this model becomes one of the models described in Prendergast (2002). Prendergast (2002) shows that there can be a positive relationship between uncertainty and incentive pay. Let n=2, and assume that $\rho_{ti}, i = 1, 2$ are drawn independently from $N(0, \tau^2)$ for all t.

## 5.1 Optimal input-based contract at the initial period (t=0)

The reservation utility for the agent is normalized to zero. Assume that the private benefits for the agent are $B_1 = 0$, and $B_2 = B > 0$. The principal does not know for which task the agent has private benefit. Since both tasks look ex-ante identical to the principal, the principal randomly chooses the task and assign it to the agent. The principal's problem is given by

$$\max_{e_{0i},\ w(e_{0i}),\ i=1,\ 2} E[\rho_{0i} + e_{0i} - w(e_{0i})] - m_e \tag{13}$$

$s.t$

$$E[w(e_{0i}) - C(e_{0i})] \geq 0 \tag{14}$$

Notice, that for input based compensation, the effort level is observable. It can be easily shown that the constraint binds at the optimum, so that $E[w(e_{0i})] = C(e_{0i})$. Plug this into (13), and we can see that the principal chooses the optimal level of effort by maximizing $E[\rho_{0i} + e_{0i} - C(e_{0i})] - m_e$. By noting that the expectation of $\rho_{0i}$ is zero, and by taking the derivative of $C(e_{0i})$ with respect to the effort, we obtain the optimal choice of effort $e^*$, which is the solution to $C'(e^*) = 1$. The expected surplus for the principal is $e^* - C(e^*) - m_e$. The expected surplus for the agent is B/2 since the principal randomly chooses between the two tasks. The total surplus (sum of principal's surplus and agent's surplus) is, thus, equal to

$$e^* - C(e^*) - m_e + B/2 \tag{15}$$

## 5.2 Optimal output-based contract at the initial period (t=0)

Since the agent is risk neutral by assumption, the principal "sells" the firm at a fixed price $\alpha_0$. In other words, the principal offers a wage contract $w_{0i} = y_{0i} - \alpha_0$.

The utility maximization problem for the agent is given by

$$\max_{e_i,\ i=1,2} E[\rho_i + e_i - \alpha_0 - C(e_i)] \tag{16}$$

Since $\alpha_0$ is constant, the agent chooses the optimal effort by solving $C'(e) = 1$. This shows that the optimal effort level is the same as that for input-based compensation, $e^*$. After the contract is made, the agent observes the realized value of $\rho_i$. The agent then chooses the task so as to maximizes his or her utility $\rho_{0i} + e^* + B_i$. We assume that $B_i$ is so small that the agent ignores the private benefit upon choosing the task, that is, the agent chooses the task that maximizes $\theta_i + \rho_{0i}$. This is the same assumption made by Prendergast (2002). This simple example shows that output-based compensation induces the correct level of effort $e^*$ and the correct task, the task that maximizes output.

The principal chooses the optimal value of $\alpha_0$ in the following way. The principal knows that the agent chooses the task that produces higher output. Let $\widehat{\rho}_0 = max\{\rho_{01}, \rho_{02}\}$. Notice that $\widehat{\rho}_0$ is a second order statistic of two independent random samples from $N(0, \tau^2)$. Since both the agent and principal do not observe the realized values of $\rho_{0i}$ for $i$=1, 2 at the time of the contract, the output is a random variable for both the principal and agent, given by $\widehat{\rho}_0 + e_{0i}$. The principal's problem is, then, given by

$$\max_{e_{0i},\ \alpha_0\ i=1,2} E[\widehat{\rho}_0 + e_i - w(e_{0i})] - m_y \tag{17}$$

$s.t$

$$E[w(e_{0i}) - C(e_{0i})] \geq 0 \tag{18}$$

Since neither the agent nor the principal observe the realized value of $\rho_i$ at the time of contract, there is no information asymmetry. Thus the optimal choice of $\alpha_0$ is the value that binds the participation constraint which is given by

$$E[\widehat{\rho}_0 + e^* - \alpha_0] - C(e^*) = 0 \tag{19}$$

By noting that $E[\widehat{\rho}_0] = \sqrt{\frac{\tau^2}{\pi}}$, we have

$$\alpha_0 = \sqrt{\frac{\tau^2}{\pi}} + e^* - C(e^*) \tag{20}$$

By adding the agent's surplus, B/2, to (20), we obtain the sum of surpluses for both the principal and the agent,[13] which is given by

$$\sqrt{\frac{\tau^2}{\pi}} + e^* - C(e^*) - m_y + \frac{B}{2} \tag{21}$$

## 5.3 The condition where output-based contract is chosen at period 0

If the surplus from output-based compensation (21) exceeds the surplus from input-based compensation (15), the principal chooses output-based compensation. This condition is givey by

$$\sqrt{\frac{\tau^2}{\pi}} + e^* - C(e^*) - m_y + B/2 \geq e^* - C(e^*) - m_e + B/2$$

$$\Longrightarrow$$

$$\sqrt{\frac{\tau^2}{\pi}} \geq m_y - m_e \tag{22}$$

Notice that the larger $\tau$ is, the more likely it is for the principal to choose output-based compensation. This illustrate the positive relationship between uncertainty and

[13]We divide B by 2 since the chance that the task with higher realized value of $\rho_{0i}$ also has the private benefit is 1/2.

incentive. The more uncertain the business environment is, the greater the likelihood that output-based contract is chosen. The idea is that, if the business environment is uncertain to the principal, the principal has little to say about how the business should be conducted. Therefore, the principal delegates the business decision to the agent who observes the business environment more clearly. To keep the agent from abusing the delegated power (for example, the agent may choose a task based on personal preference which may not maximize output), the principal ties compensation to output. It can be seen that this simple example illustrates the positive relationship between uncertainty and incentive-based pay.

### 5.4 Condition where output-based compensation is chosen at period t

Because of the iid assumption about $\rho_{it}$, the decision rule at period t, $t \geq 1$ is the same as the decision rule at period 0, i.e., at period t, the principal chooses output-based compensation if the following holds.

$$\sqrt{\frac{\tau^2}{\pi}} \geq m_y - m_e \tag{23}$$

The principal is more likely to choose output-based compensation when the business environment is more uncertain.

## 6 A model with match parameters

The realized values of the match parameters are not directly observable to the principal. Thus, those parameters introduce additional uncertainty into the model. Match parameters can be interpreted as the agent's productivity of performing a specific task at a given firm and is referred to as an "ability" in this study. Due to such additional uncertainty, it becomes more difficult for the principal to assign the correct task. This makes it more likely for the principal to choose output-based compensation. However, the principal can infer the realized values match parameters by observing output $y_{ti}$ over time. This means that overall uncertainty decreases over time. Our idea to show a negative relationship between incentive and tenure is as follows. Suppose that, upon hiring a new agent, uncertainty is high enough for the principal to choose output-based compensation. By observing output over time, the principal's uncertainty about the agent decreases over time. If the uncertainty decreases sufficiently, the principal may switch to input-based compensation later in the agent's tenure. Since input-based compensation is less incentive-intensive, this shows a negative relationship between tenure and incentive. We formalize such an argument in this section.

We keep the assumption that once the agent is hired, he or she is hired forever. The principal hires an agent to undertake one of two possible tasks. At the beginning of the period, the principal and agent jointly draw match parameters for each task. $\theta_i$ is the match parameter for task $i$, $i$=1, 2. The match parameters are assumed to be time invariant. At period $t$, the output when the agent performs task i is given by

$$y_{ti} = \theta_i + \rho_{ti} + e_{ti} \tag{24}$$

We assume that match parameters are drawn independently from the same distribution $N(\mu, \sigma^2)$. All other settings are the same as the simplest model described in Section 6.

To simplify the notation, let

$$X_{ti} = \theta_i + \rho_{ti} \tag{25}$$

By the independence assumption of $\theta_i$ and $\rho_{ti}$, and by the fact that the principal does not observe the realized value of both parameters, the principal perceives $X_{ti}$ as a random variable drawn from $N(\mu, \sigma^2 + \tau^2)$.

## 6.1 Optimal input-based contract at period 0

At the beginning of the period, there is no previous information about the match parameters. The principal's problem is to solve the following.

$$\max_{e_{0i},\ w(e_{0i}), i=1,2} E[X_{0i} + e_{0i} - C(e_{0i})] - m_e \tag{26}$$

$$s.t$$

$$w(e_{0i}) - C(e_{0i}) \geq 0 \tag{27}$$

The constraint binds, i.e., $w(e_{0i}) = C(e_{0i})$. The principal, then, chooses the optimal effort level by maximizing

$$E\left[X_{0i} + e_{0i}\right] - C(e_{0i}) - m_e \tag{28}$$

The optimal level of effort is again given by $e^*$, the solution to $C'(e^*) = 1$. Since both $X_{01}$ and $X_{02}$ have the same distribution, the principal randomly chooses between two tasks. The sum of the surpluses for principal and agent is given by

$$\mu + e^* - C(e^*) - m_e + B/2 \tag{29}$$

## 6.2 Optimal output contract at period 0

Since this is the initial period, there is no previous information about $\theta_i$. Due to the risk neutrality assumption, the principal "sells" the firm at a fixed price $\alpha_0$. The output-based contract is of the form, $w_{0i} = y_{0i} - \alpha_0$. The agent's problem is given by

$$\max_{e_i,\ i=1,2} [X_{0i} + e_i - \alpha_0 - C(e_i) + B_i] \tag{30}$$

Clearly, the optimal effort is $e^*$, the solution to $C'(e) = 1$. The agent chooses the task that maximizes $X_{0i} + e^* - \alpha_0 - C(e^*) + B_i$. Again, we assume that $B_i$ is so small that the agent ignores the private benefit upon choosing a task, i.e., the agent choose a task so as to maximize $\theta_i + \rho_i$.

Let $Z_0 = min\{X_{01}, X_{02}\}$. At the time of the contract, both the principal and the agent do not observe $\theta_i$ and $\rho_i$. Thus both the principal and the agent perceive $Z_0$ as the second order statistic of two independent samples from $N(\mu, \sigma^2 + \tau^2)$. Thus the output is a random variable given by $y_{0i} = Z_0 + e_{0i}$ for both the principal and the agent at the time of the contract. At the optimal contract, the participation constraint binds. The optimal value of $\alpha_0$ can be found by solving the following.

$$E(Z_0 + e^*) - \alpha_0 - C(e^*) = 0$$

$$\Longrightarrow$$

$$\alpha_0 = \mu + \sqrt{\frac{\sigma^2 + \tau^2}{\pi}} + e^* - C(e^*) \tag{31}$$

The sum of the surpluses for both the principal and the agent is given by

$$\mu + \sqrt{\frac{\sigma^2 + \tau^2}{\pi}} + e^* - C(e^*) - m_y + \frac{B}{2} \tag{32}$$

## 6.3 The Condition where output-based compensation is preferable at period 0

By comparing (29) and (32), the condition in which the output-based compensation is preferred at period t=0 is given by

$$\sqrt{\frac{\sigma^2+\tau^2}{\pi}} \geq m_y - m_e \tag{33}$$

Again the decision of whether the principal chooses output-based compensation or input-based compensation depends on the level of uncertainty. However for this case, the uncertainty stems not only from the riskiness of the task $\tau$ but also from the uncertainty about the agent's ability $\sigma$. The more uncertain the principal is about the agent's ability, the greater the likelihood that the principal chooses output-based compensation.

The condition (33) is indicative of the principal's decision for later periods. Notice that the principal can infer the realized value of $\theta_i$ by repeatedly observing the output over time. This reduces the uncertainty stemming from the match parameters for later periods. This in turn makes it likely that the principal will switch to input-based compensation later on. To make this point clear, let us assume that

$$\sqrt{\frac{\sigma^2+\tau^2}{\pi}} \geq m_y - m_e \tag{34}$$

$$\sqrt{\frac{\sigma^2}{\pi}} < m_y - m_e \tag{35}$$

The condition (34) means that at the beginning of the period, the principal always prefer output based compensation. The condition (35) means that if there is no uncertainty in the agent's ability, the principal always prefers input-based compensation. This is the most interesting case since, if condition (34) is not satisfied, the principal chooses input-based compensation upon hiring the agent. Since uncertainty does not increase in our setting, for all the remaining period, the principal continues to choose input-based compensation. If condition (35) is not satisfied, the uncertainty of the business enviroment is so large, that even if the principal knows the agent's ability with certainty, the principal chooses output-based compensation at any period.

Our idea to show the inverse relationship between incentive and tenure is the following. As the principal repeatedly observe output over time, the uncertainty about the agent's ability (expressed by $\sigma$) decreases over time, reducing overall uncertainty. Therefore, it is more likely for the principal to choose input based compensation later in the agent's tenure.

### 6.4 Condition where output-based and input-based compensation in period 1

In period 1, the principal observes the output in the previous period, t=0. The principal has updated his belief about the distribution of $\theta_i$. Let $k_0 = \frac{\sigma^2}{\sigma^2+\tau^2}$. If the agent chooses task i in the previous period, the conditional distribution of $\theta_i$ is given by

$$[\theta_i|\theta_i + \rho_{0i} + e^* = y_0] \sim N[\mu + k_0(y_0 - \mu), k_0\sigma^2] \tag{36}$$

and if the agent did not choose task $i$ in the previous period, the conditional distribution of $\theta_i$ is given by

$$[\theta_i|\theta_i + \rho_{0i} + e^* \leq y_0] \tag{37}$$

$$\sim \quad \frac{1}{\Phi(\frac{y}{\sqrt{\sigma^2+\tau^2}})} \int_{\infty}^{t} [1 - \Phi(\frac{\theta_i - y}{\tau})] \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2}\frac{(\theta_i - \mu)^2}{\sigma^2}] d\theta_i \tag{38}$$

## 6.5 Optimal input-based compensation for period 1

At the beginning of period 1, the principal has already observed which task the agent chose in period 0, as well as the realized output $y_{0i_0} = \theta_{i_0} + \rho_{0i_0} + e^*$. $i_0$ denotes the task that the agent chose in period $t = 0$. Given this information, the principal's problem for input-based compensation is the following.

$$\max_{e_{1i}, w(e_{1i}),\ i=1,2} E[\theta_i + \rho_{1i} + e_{1i} | y_{0i_0}] - w(e_{1i}) - m_e \tag{39}$$

$s.t$

$$w(e_{1i}) - C(e_{1i}) \geq 0 \tag{40}$$

The participation constraint binds so that, again the optimal effort level is $e^*$. Notice that

$$E[\theta_i + \rho_{1i} + e^* | y_{0i_0}] - w(e^*) - m_e \tag{41}$$

$$= \quad E[\theta_i + \rho_{1i} | y_{0i_0}] + e^* - w(e^*) - m_e \tag{42}$$

$$= \quad E[\theta_i | \theta_{i_0} + \rho_{0i_0} + e^*] + e^* - w(e^*) - m_e \tag{43}$$

where $E[\theta_i|\theta_{i_0} + \rho_{0i_0} + e^*]$ should be understood as the shorthand notation of $E[\theta_i|\theta_{i_0} + \rho_{0i_0} + e^* = y_{0i_0}]$ if $i = i_0$ and $E[\theta_i|\theta_{i_0} + \rho_{0i_0} + e^* \leq y_{0i_0}]$ if $i \neq i_0$.

We introduce a new notation. Let $\theta_{1i} \sim [\theta_i|y_{0i_0}]$, i.e., $\theta_{1i}$ is a random variable having the distribution given by (36) or (37) depending on wether $i = i_0$ or $i \neq i_0$ respectively.

Now, let $S_1 = max\{[E[\theta_{11} + \rho_{11}], E[\theta_{12} + \rho_{12}]\}$. For input based compensation, the principal assigns a task that has higher expected output. Thus, the sum of the expected surplus of the principal and the agent is

$$S_1 + e^* - C(e^*) - m_e + B/2 \quad (44)$$

## 6.6 Optimal output based compensation at t=1:

For an output based contract, because of risk neutrality, the wage offer has the following form, $w_{1i} = y_{1i} - \alpha_1$. $\alpha_1$ is the price at which the principal "sells" the firm. Using the same argument as t=0, we can see that the agent chooses the effort level equal to $e^*$ and chooses the task that has a higher output level.

The agent has observed the realized value of $\theta_1$ and $\theta_2$ in the initial period $t$=0. Therefore, at $t$=1, the match parameters are known to the agent. On the other hand, the principal only has information about the distribution of these parameters conditional on the previous period's output. This implies that there is information asymmetry at the time of the contract. The principal has to choose the optimal $\alpha_1$ given this information asymmetry. This causes the following problems. The principal set the optimal price at which he or she sells the firm to the agent ($\alpha_1$) by making the participation constraint binding, i.e., $E[y_{1i} - \alpha_1 - C(e_{1i})] = 0$. Because of the information asymmetry, the expectation with respect to the principal's information is different from the expectation with respect to the agent's information. Optimal $\alpha_1$ with respect to the principal's information may be too small (constraint is slack) or may be too large (constraint is violated) with respect to the agent's information. It is in general impossible to always make the participation constraint binding. We find that characterizing the optimal $\alpha_1$ that takes into account such information asymmetry difficult. Therefore, we simply assume that the agent chooses $\alpha_1$ by satisfying the following.

$$E[max\{\theta_{11} + \rho_{11}, \theta_{12} + \rho_{12}\}] + e^* - \alpha_1 - C(e^*) = 0 \tag{45}$$

$$\Longrightarrow$$

$$\alpha_1 = E[max\{\theta_{11} + \rho_{11}, \theta_{12} + \rho_{12}\}] + e^* - C(e^*) \tag{46}$$

In other words, we assume that the principal chooses $\alpha_1$ as if the agent did not know the matching parameters. $\alpha_1$ computed this way can be understood as an approximation to the true optimal $\alpha_1$ that takes into account the information asymmetry. If the principal's uncertainty about the agent's ability is small, this should be a good approximation. Based on this approximation, the sum of the expected surplus for both the principal and the agent is given by

$$E[max\{\theta_{11} + \rho_{11}, \theta_{12} + \rho_{12}\}] + e^* - C(e^*) + B/2 \tag{47}$$

## 6.7 Condition where the output based contract is preferred at period 1

This is given by

$$E[max\{\theta_{11} + \rho_{11}, \theta_{12} + \rho_{12}\}] - S_1 > m_y - m_e \tag{48}$$

**Proposition 1** *The principal is more likely to choose input based compensation at t=1, that is*

$$E[max\{\theta_{11} + \rho_{11}, \theta_{12} + \rho_{12}\}] - S_1 < \sqrt{\frac{\sigma^2 + \tau^2}{\pi}} \tag{49}$$

**Proof** *See Appendix C*

Therefore, in the period t=1, the principal is more likely to use input-based compensation. The intuition is that, in the period 1, the principal has observed the output in the previous period and has a better idea about the agent's ability, decreasing the principal's overall uncertainty about the business outcome. Reduced uncertainty makes it more likely for the principal to choose input-based compensation.

## 6.8 Condition where output-based compensation is preferred at period t

We assume that the principal chooses output-based compensation until the (t-1)-th period. Let $y_{0i_0}, \ldots, y_{t-1i_{t-1}}$ be the the history of outputs observed by the principal. $i_s$ is the task chosen by the agent at period $s$. The distribution of $\theta_i$ conditional on the output history is given by

$$F_{ti}(x) = P(\theta_i \leq x | y_{0i_0}, \ldots, y_{t-1i_{t-1}}) \tag{50}$$

This notation needs explanation. Fix $i$ at period t. Continue to use the notation $i_s$ to denote the task chosen at period $s$. If $i = i_s$, then the notation $P(\theta_i \leq x | y_{si_s})$ means $P(\theta_i \leq x | \theta_{i_s} + \rho_{si_s} + e^* = y_{si_s})$. If $i \neq i_s$, the notation $P(\theta_i \leq x | y_{si_s})$ means $P(\theta_i \leq x | \theta_{i_s} + \rho_{si_s} + e^* \leq y_{si_s})$. Let $\theta_{ti}$ be a random variable having distribution given by $F_{ti}$ in (50). In other words, $\theta_{ti}$ is the principal's updated belief about $\theta_i$ given the history of output.

## 6.9 Input-based compensation in period t

Input-based compensation is to assign a task with higher expected value of $\theta_i + \rho_{ti}$. Expectation is taken conditional on information about the history of output. The agent chooses effort equal to $e^*$. Let $S_t = max\{E[\theta_{t1} + \rho_{t1}], E[\theta_{t2} + \rho_{t2}]\}$. The total surplus of input based compensation is given by

$$S_t + e^* - C(e^*) - m_e + B/2 \tag{51}$$

## 6.10 Output-based compensation at period t

Because of risk neutrality, the optimal output-based compensation is $w_{it} = y_{it} - \alpha_t$ where $\alpha_t$ is the price at which the principal "sells" the firm to the agent. Using exactly the same logic as t=1, we can show that the agent chooses the optimal effort $e^*$ and chooses the task that maximizes the output.

There is again a problem of information asymmetry in choosing the optimal $\alpha_t$. As in the case for t=1, we assume that the principal chooses $\alpha_t$ by binding the participation constraint with respect to the principal's information, and that the agent accepts the offer.

$$E[max\{\theta_{t1} + \rho_{t1}, \theta_{t2} + \rho_{t2}\}] + e^* - \alpha_t - C(e^*) = 0 \tag{52}$$

$$\Longrightarrow$$

$$\alpha_t = E[max\{\theta_{t1} + \rho_{t1}, \theta_{t2} + \rho_{t2}\}] + e^* - C(e^*) \tag{53}$$

Expectation is taken conditional on the information held by the principal. The sum of the expected surpluses for both the principal and the agent is given by

$$E[max\{\theta_{t1} + \rho_{t1}, \theta_{t2} + \rho_{t2}\}] + e^* - C(e^*) + B/2 \tag{54}$$

## 6.11 The relationship between incentive and tenure

By comparing the surplus for input based compensation (51) and the surplus for output-based compensation (54), the condition in which output-based compensation is preferred at period t is given by

$$E[max\{\theta_{t1} + \rho_{t1}, \theta_{t2} + \rho_{t2}\}] - S_t > m_y - m_e \tag{55}$$

Now, the following proposition shows that incentive decreases with tenure.

**Proposition 2** *Under the following conditions,*

$$\sqrt{\frac{\sigma^2 + \tau^2}{\pi}} \geq m_y - m_e \tag{56}$$

$$\sqrt{\frac{\tau^2}{\pi}} < m_y - m_e \tag{57}$$

*the principal chooses output-based compensation upon hiring a new agent, then switch to input-based compensation later on after observing the output for a sufficient number of periods.*

**Proof** *See Appendix (D)*

Since input-based compensation is less incentive intensive compensation, this proposition shows that incentive decreases with tenure under conditions (56) and (57). This proposition also suggests that in general, incentive is non-increasing in tenure. To see this, notice that if condition (56) is not met, the principal chooses input-based compensation upon hiring a new agent. Since uncertainty is non-increasing, the principal continues to choose input-based compensation at any period. Thus, incentive is unchanged over time. If condition (57) is not met, even if the principal knows the agent's ability with certainty, he or she chooses output-based compensation. Again, incentive is unchanged over time.

# 7 Incorporating turnover

So far we have assumed that once the agent is hired, he or she works forever at the firm. We will now relax this assumption by introducing turnover into the model. We consider a three-stage model. The first stage is the pre-contractual period. The principal and agent jointly draw match parameters and make an employment agreement. The second stage is the end of the initial period, where both agent have observed the initial period output. In the second stage, the principal decides whether or not to retain the agent. If the principal retains the agent, the principal makes the decision about whether to choose input-based compensation or output-based compensation. If the principal fires the agent, the principal finds a new agent (the principal goes back to the first stage). The third stage is the end of the second period. We assume that at the third stage, the true values of the match parameters are revealed. In this stage, the principal also makes the decision about whether to retain the agent or not. If the decision is to retain the agent, he or she also chooses the type of compensation. We can find decision rules for the principal by solving the problem backwards.

## 7.1 The Decision rule at the third stage:

Let Q be the pre-draw value and $J_3$ be the value function of the principal at the third stage. In this stage, the realized values of the match parameters are revealed. If the principal retains the agent at this stage, the agent will be hired forever. Furthermore, if the principal retains the CEO, by the assumptions (44) and (45), the principal always chooses input based compensation. The present discounted value of the total surplus is $\frac{1}{1-\beta}[m_2 + e^* - C(e^*) - m_e + B/2]$ where $\beta$ is the discount factor. If the principal decides to fire the agent, the principal goes back to the first stage and finds a new agent. The principal's value when finding a new agent is Q. The value function at the third stage is written as

$$J_3(m_2) = max\{\frac{1}{1-\beta}[m_2 + e^* - C(e^*) - m_e + B/2], Q\} \tag{58}$$

where $m_2 = E[\theta_1 + \rho_{21}, \theta_2 + \rho_{22}]$. Notice that, $\frac{1}{1-\beta}[m_2 + e^* - C(e^*) - m_e + B/2]$ is an increasing function in $m_2$ whereas Q is constant. Therefore, the decision rule for the principal is the following.

$$Keep\ the\ agent\ if\ m_2 \geq \overline{m}_2$$

$$Find\ a\ new\ agent\ if\ m_2 < \overline{m}_2$$

$$where\ \overline{m}_2\ satisfies\ \frac{1}{1-\rho}[\overline{m}_2 + e^* - C(e^*) - m_e + B/2] = Q$$

If the realized ability is too low to meet the threshold value $\overline{m}_2$, the principal fires the agent and find a new agent. Otherwise, the principal retains the agent forever.

## 7.2 The decision rule at the second stage

In the second stage, the principal has seen the output of the initial period. First, consider the value when the principal's decision is to retain the agent. Let $i_0$ denotes the task that the agent has chosen at $t = 0$. The principal makes a decision about whether to retain the agent at this stage. If the decision is to retain the agent, the principal also chooses either input-based compensation of output-based compensation. Note that, given the decision of the principal to retain the agent, the decision rule for choosing between input-based compensation and output-based compensation is the same as the case in which we do not have turnover in the model. Therefore, the introduction of turnover in the model does not alter the result that the principal is more likely to choose input based compensation when uncertainty is small. Since the uncertainty about the agent's ability is smaller in the second stage than in the initial stage, the principal is more likely to choose input-based compensation in the second stage than in the initial stage. This still suggests a negative relationship between incentive and uncertainty.

If the principal chooses input-based compensation, the surplus is given by equation (51). $S_1$ in the equation is a function of $y_{0i_0} = \theta_{i_0} + \rho_{1i_0}$. Let $v_e(y_{0i_0})$ be the surplus at t=1 for input-based compensation. The surplus at t=1 when choosing output based compensation is given by equation (54). Again, this is a function of $y_{0i_0}$. Let $v_e(y_{0i_0})$ and $v_y(y_{0i_0})$ be the surplus for input-based and output-based compensation respectively. The present discounted value given the decision to retain the agent is thus given by

$$max[v_e(y_{0i_0}), v_y(y_{0i_0})] + \beta E[J_3(m_2)]$$

.

Therefore, the value function at the second stage $J_2$ is a function of $y_{0i_0}$, given by

$$J_2(y_{0i_0}) = max\{max[v_e(y_{0i_0}), v_y(y_{0i_0}) + \beta E[J_3(m_2)],\ Q\} \tag{59}$$

Both $v_e(y_{0i_0})$ and $v_y(y_{0i_0})$ are increasing functions in $y_{0i_0}$. The decision rule for the principal at the second stage is given the

$$Keep\ the\ CEO\ if\ y_{0i_0} \geq \overline{y}_1$$

$$Find\ a\ new\ CEO\ if\ y_{0i_0} < \overline{y}_1$$

$$where\ \overline{y}_1\ satisfies\ max[v_e(\overline{y}_1), v_y(\overline{y}_1)] + \beta E[J_3(m_2)] = Q$$

## 7.3 Value in the initial stage:

In the initial stage, due to assumptions (56) and (57), the principal chooses output-based compensation. The value Q is thus given by

$$Q = \mu + \sqrt{\frac{\sigma^2 + \tau^2}{\pi}} + e^* - C(e^*) - m_y + \frac{B}{2} + \beta E[J_2(y_0)] \tag{60}$$

The value functions as a whole depends on all the parameters in the equation such as $\sigma$ and $\tau$. Given specific values of the parameters, Q can be computed by a recursive method using equations (58), (59) and (60).

Incorporating turnover in the model does not alter the most important result -that incentive decreases with tenure. This is because at each time period, the principal chooses between input-based and output-based compensation according to the rules given by equation (55) conditional on the decision of the principal to retain the agent.

Introduction of turnover merely adds an extra layer to the principal's decision, i.e., to decide whether to retain the CEO or not. In addition, the model suggests that a better performance leads to a lower likelihood of CEO turnover.

# 8 Implication of the model

We conclude this study by summarizing the implication of our model.

**Incentives are non-increasing with tenure** If condition (56) and (57) are satisfied, the principal chooses output-based compensation upon hiring a new agent but switches to input-based compensation later in the agent's tenure. This shows a negative relationship between incentive and tenure. If conditions (56) and (57) are not met, incentive is unchanged over time. Thus our model suggests that incentive is non-increasing with tenure.

**2. Incentives increase with the uncertainty about the agent's ability:** If the principal's uncertainty about the agent's ability is high, the principal is likely to choose output-based compensation, suggesting that there should be a positive relationship between incentive and uncertainty about the agent's ability. If we can find a proxy variable for such uncertainty, we can potentially test this hypothesis. For example, consider CEO compensation. It is reasonable to assume that a CEO's managerial ability is known with greater certainty if the number of years worked by the CEO at the company prior to becoming a CEO is large. In other words, the number of years worked prior to becoming a CEO serves as a proxy for prior knowledge about the CEO's ability held by the shareholders and the board of directors. We can potentially test this implication by testing the relationship between pay-for-performance sensitivity and years worked by the CEO prior to becoming a CEO.

**3. Incentives are non-decreasing with uncertainty in which the firm operates:** This result is from Prendergast (2002) and is the main result on which our model is based. If we find a proxy variable for such uncertainty, we can potentially test this hypothesis. For CEO compensation, the R&D to the book value of asset ratio has been used to proxy for such uncertainty. For example, see Garen (1994).

# 9 Conclusion

Although standard principal-agent theories imply that incentive intensity of compensation increases with tenure, prior empirical work of the relationship between pay-for-performance sensitivity in CEO compensation consistently finds a negative relationship. This becomes our motivation to develop a model that is consistent with the prior empirical findings on the relationship between incentive and tenure. Our model is an extension of Prendergast's (2002) which shows a *positive* relationship between incentive and tenure. We introduced additional uncertainty to Prendergast's model, i.e., the principal's uncertainty about the agent's ability. The CEO's ability is gradually revealed over time, decreasing the principal's uncertainty of the business outcome. Assuming Prendergast result, we show that the principal is more likely to choose input-based compensation for CEO's with longer tenure because of the reduced uncertainty. This suggests a negative relationship between incentive and tenure. These implications of the model can be tested using CEO compensation data. In addition to the negative relationship between incentive and tenure, this model also implies that incentive is non-increasing in the uncertainty about the agent's ability, and that incentive is non-increasing in the uncertainty in which the firm operates.

Chapter 3

**Empirical investigation of the relationship between tenure and incentive using CEO compensation data**

# 1 Introduction

Many prior studies relate CEO compensation and principal-agent theories whereby principal-agent theories have been supported. However, one of the standard results of principal agent theories that pay-for-performance sensitivity increases with tenure as CEO has not been supported by empirical studies. Murphy (1986) uses CEO compensation data between 1974 and 1984 and finds that pay-for-performance sensitivity is smaller for CEOs whose tenure is greater than 4.6 years than for CEOs with shorter tenure. This result means that pay-for-performance sensitivity decreases with CEO tenure, a result which contradicts the standard principal-agent theories. Gibbons and Murphy (1992) also use CEO compensation data between 1971 and 1989 and similarly find that pay-for-performance sensitivity decreases as CEO tenure increases, contradicting standard principal-agent theories.

Such results leave us unsure if CEO compensation is consistent with standard principal agent theories. In fact, I show in Chapter 2 that pay-for-performance sensitivity may be non-increasing with tenure as CEO. This model is an extension of Prendergast (2002) and is referred to as "extended-Prendergast" model in this study. There are two econometric problems in the previous studies by Murphy (1985) and Gibbons and Murphy (1992). First, their studies do not include stock option grants in the total compensation; therefore, a major part of the incentive effect may be missing from their results. Second, their studies ignore selection biases: Since turnover of CEO is common place, for each period, we only observe CEOs who survive up to that period. Ignoring such facts potentially causes selection biases. Since tenure and incentive could have a negative relationship, and since prior studies have econometric problems, it is worthwhile to re-examine the tenure-incentive relationship.

We seek to make the following contributions. (1) This study examines the relationship between pay-for-performance sensitivity and CEO tenure using more recent CEO compensation data. CEO tenure is the number of years as CEO. (2) We estimate such a relationship by estimating a joint system of equations that incorporate unobserved heterogeneity, which controls for selection biases. (3) This studies serves as a test of wether CEO compensation is explained by standard principal-agent theories or explained better by the "extended-Prendergast model" described in Chapter 2. In doing so, we also examine two additional relationships: (a) incentive and uncertainty about the managerial ability of CEOs, and (b) incentive and riskiness of the environment in which the firm operates. Standard principal-agent theories predict that we find a positive relationship for (a) and a negative relationship for (b). Our "extended-Prendergast" model predicts the opposite. (a) is tested by examining the relationship between pay-for-performance sensitivity and the number of years worked at the firm prior to becoming a CEO. (b) is tested by examining the relationship between pay-for-performance and R&D to asset ratio. Rationale of the methods for these tests is explained in Section 3.

To preview the results, we find that there is a strong and positive relationship between pay-for-performance sensitivity and CEO tenure. For CEOs with tenure fewer than or equal to six years, an improvement in firm performance from the median level to the 75th percentile level would only lead to a 0.06% increase in total compensation. On the other hand, for CEOs with tenure longer than or equal to seven years, the same improvement in total compensation would lead to an 8% increase in total compensation. This is strong support for standard principal-agent theories like Gibbons and Murphy's (1999), but rejects our "extended-Prendergast" model. For the additional test (a) we find little evidence of the existence of the relationship between pay-for-performance sensitivity and years worked prior to becoming CEOs. For the additional test (b), we did not find evidence of the existence of the relationship between pay-for-performance sensitivity and R&D to asset ratio. These additional tests do not provide further evidence to support or reject both standard principal-agent theories and our "extended-Prendergast" model.

# 2 Theories and prior empirical studies

Murphy (1986) characterizes the incentive contract for CEOs in a multi-period principal-agent setting. In his model, the agent works for a fixed number of periods. His model implies that pay rise is spread evenly throughout the remaining years, i.e., good performance is rewarded as a shift in the earning-tenure profile. This, in turn, implies that pay-for-performance sensitivity increases with tenure because, for later years, the reward is spread among fewer years. Thus, the same increase in output causes greater increase in annual compensation.[14] The econometric model is $\Delta ln(Salary + Bonus)_{it} = \alpha + \beta(Years\ as\ CEO) + \delta(Rate\ of\ return\ on\ common\ equity)_{it} + \epsilon_{it}$. Murphy uses CEO compensation data between 1974 and 1984. He divides the data into three sub-samples; the first sub-samples are for CEOs with tenure of 4.6 and less, the second sub-samples are for CEOs with tenure between 4.6 and 9.9 and the third sub-samples are for CEOs with tenures greater than 9.9. Murphy finds that $\delta$ is smaller for those with longer tenure. Thus, the estimated results show that pay-for-performance actually decreases with tenure, which is inconsistent with standard results of principal-agent theories.[15]

Gibbons and Murphy (1992) develop a model that characterizes the incentive contract when the agent has reputational career concern. In their model, good current performance positively affects the CEO's assessed ability, as assessed by the shareholders. The assessed ability is interpreted as "reputation" of the CEO. A good reputation about a CEO's ability causes an increase in his or her future compensation, thus creating incentives for CEOs to exert effort even without an explicit incentive contract. However, the reputational concern disappears as the CEO reaches his or her retirement. To supplement such a loss in incentive, firms may intensify pay-for-performance sensitivity as the CEO approaches his or her retirement. Gibbons and Murphy's model has two implications: (i) Pay-for-performance sensitivity increases as CEOs reach their retire-

[14] Therefore, the implication that pay-for-performance increases with tenure depends on the finiteness of a CEO's lifetime.

[15] They conclude that those results are supported by learning model instead.

ment, holding CEO tenure constant. (ii) Pay-for-performance sensitivity increases as CEO tenure increases, holding years to retirement constant. The second implication is relevant for our study.

The reason for the second implication is as follows. First, because of the agent's risk aversion, the standard trade-off between incentive and uncertainty holds, i.e., pay-for-performance sensitivity decreases with uncertainty of business performance. Suppose that the firm's output is given by $y_t = \eta + a_t + \epsilon_t$ where $\eta$ is an unobservable ability of a CEO (unobservable both by CEOs and the shareholders), $a_t$ is the effort and $\epsilon_t$ is the usual independent random shocks. Notice that there are two sources of uncertainty to both shareholders and the CEO, $\eta$ and $\epsilon_t$. Although $\eta$ is not directly observable (both by the shareholders and by the CEO), it can be inferred by repeatedly observing the firm's output over time. Because the uncertainty about the ability of the CEO decreases over time, so does the overall uncertainty of business outcome diminish. By the inverse relationship between incentive and uncertainty, the Gibbons and Murphy show that pay-for-performance sensitivity should increase with tenure.

Their econometric model to test the pay-for-performance tenure relationship is the following. $\Delta(Salary + Bonuses)_{it} = \beta_1(low\ tenure\ dummy) \times \Delta ln(Firm\ Value)_{it} + \sum_{n=1972}^{1988} \beta_n(n^{th}\ year\ dummy_{it}) \times \Delta ln(Firm\ value)_{it} + \beta'(other\ variables)_{it}$. The theory predicts that $\beta_1$ is negative, reflecting the negative relationship between pay-for-performance sensitivity and CEO tenure. They used CEO compensation data between 1971 and 1989 to estimate this model. They find a positive but statistically insignificant coefficient. Thus the implication of their theory that incentive and uncertainty has positive relationship is not strongly supported by the data. Nonetheless, he finds evidence that pay-for-performance sensitivity increases as the agent reaches his or her retirement, although the magnitude of increase in pay-for-performance sensitivity is surprisingly small. For the CEOs with cash earning of $562,000, a 10% change in shareholders wealth corresponds to $7,300 of CEO 's wealth change for CEOs more than three years from retirement, but $9,500 for CEOs fewer than thee years to retirement.

Note again that the basic reason why pay-for-performance sensitivity increases with tenure in Gibbons and Murphy (1992) is because of the standard inverse relationship between incentive and risk. Contrary to such standard results, Prendergast (2002) shows that, under some conditions, pay-for-performance sensitivity is non-decreasing with uncertainty. In his model, uncertainty dictates if the principal chooses “input-based compensation” or “output-based compensation”. For input-based compensation, the agent is rewarded by his or her input, i.e., if he or she is keeping busy. On the other hand, output-based compensation ties the compensation to the output, i.e., the firm’s value. It is assumed that making output-based compensation incurs a fixed cost to the firm so that the firm wants to avoid output-based compensation if necessary. Assuming that the agent observes the business environment better than the principal, if the uncertainty of the business environment is significant, the principal delegates the decision making power to the agent. However, to keep the agent from abusing the delegated power, the principal uses output-based compensation. In other words, under greater uncertainty, the benefit of delegating decision making power to the agent exceeds the cost of making output-based compensation. If the business environment is less uncertain, the principal keeps the decision making power to himself or herself and uses input-based compensation. Since output-based compensation ties compensation to output, we are likely to observe greater pay-for-performance sensitivity when uncertainty is large.

In Chapter 2, we extended the Prendergast model to a multi-period setting. We also incorporate a matching theory to his model. Basic intuition of our "extended-Prendargast Model" is simple: Uncertainty regarding the outcome of the business stems not only from the business environment but also from uncertainty regarding the CEO's ability. The CEO's ability is gradually revealed over time, thus decreasing the uncertainty of the business outcome. Using Prendergast's argument, we can show that the principal is more likely to use output-based compensation earlier in the CEO's tenure. This suggests a negative relationship between tenure and incentive.

In summary, standard principal-agent theories like Gibbons and Murphy's imply that there should be a positive relationship between tenure and pay-for-performance sensitivity. On the other hand, the extended-Prendargast model implies that pay-for-performance sensitivity is non-decreasing with tenure.

# 3 Testable implications

As is noted, the main objective of our study is to estimate the relationship between pay-for-performance sensitivity and CEO tenure using the most recent CEO compensation data. Examination of relevant literature in Section 2 reveals additional testable implications of principal agent theories, which are summarized in the this section.

**1. Incentive and tenure:** Standard principal-agent theories predict that pay-for-performance sensitivity increases with CEO tenure. On the other hand, our extented-Prendergast model predicts that pay-for-performance sensitivity is non-increasing with tenure.

**2. Incentive and the uncertainty about managerial ability:** Standard principal-agent model implies that pay-for-performance sensitivity increases if the uncertainty about the managerial ability of the CEO decreases. On the other hand, our extended-Prendergast model implies that pay-for-performance sensitivity is non-increasing with uncertainty about the agent's managerial ability, because greater uncertainty about the CEO's managerial ability is likely to make the principal choose output-based compensation.

It seems reasonable to assume that the managerial ability of CEOs is known with greater certainty if the number of years worked at the company prior to becoming a CEO is greater. In other words, the number of years worked prior to becoming a CEO serves as a proxy for prior knowledge held by the shareholders and the board of directors about CEO's managerial ability. Therefore, we can potentially test this implication by testing the relationship between pay-for-performance sensitivity and years worked prior to becoming a CEO.

**3. Incentive and riskiness of the business environment:** Standard principal-agent models imply that there is a negative relationship between pay-for-performance sensitivity and the riskiness of the environment in which the firm operates. On the other hand, our model predicts that pay-for-performance sensitivity decreases with risk. To test this, we need a proxy variable for the risk. The proxy variable that we use is R&D to the book value of asset ratio. The rationale for the choice is presented in Section 4.

# 4 Definitions of our key variables

To test the three implications described in the previous section, we start by defining total compensation, the proxy for the firm performance variable, and the proxy variable for the riskiness of the environment under which the firm operates.

**Total compensation:** Both Murphy (1986) and Gibbons and Murphy (1992) use (Salary + Bonuses) as the CEO compensation variable, thus omitting stock based compensation such as stock options. This is because, their compensation data is from Fobes Executive compensation survey, which does not have information about stock option grants. Instead, we use ExecComp as the main data source, which has a detailed break down of the entire compensation package, including Black-Sholes Value of stock options. Since the stock options are the largest components of CEO annual compensation, we define total compensation as the sum of salary, bonuses and Black-Sholes values of stock options.[16]

[16] We include stock appreciation rights into the stock options.

There are several advantages in defining total compensation this way. First, salary, bonuses and stock options are the three components of CEO compensation package that attract the most attention in the executive compensation literature. Second, these three components make up approximately 90% of the whole compensation package for each year of our sample. By focusing on these three components, we can eliminate relatively minor components such as severance pay which may be influenced by factors other than firm performance variables.

**Firm performance:** Our firm performance variable is the change in firm values normalized by the end of the year firm value, given by the formula below.

$$(Perform)_{it} = \frac{(firm\ value)_{it} - (firm\ value)_{it-1} + (divident\ payout)}{(firm\ value)_{it}} \tag{61}$$

We also include dividend payout since shareholders are rewarded not only by increases in firm values but also by dividend payout. As is clear from the definition, this variable captures the values added to the firm by the CEO for each year. This definition has the following attractive feature: if we estimate a linear relationship of the following $(Total\ comp)_t = \alpha_t + \beta_t (Perform)_t$, the total compensation increases (decreases) only when the firm value increases (decreases), which is consistent with standard principal-agent theories.

Our definition of firm performance deviates from the proxy variable used by Gibbons and Murphy (1992). Gibbons and Murphy use end-of-year firm value as the firm performance variable. A problem associated with this definition is that firm values tend to be large for large size firms and large size firms tend to pay large amounts. Therefore, if we regress the firm performance against the total compensation using Gibbons and Murphy's definition, we may pick up the firm size effect instead of the pay-performance relationship. Our definition, on the other hand, avoids such problems since our performance variable is a "normalized change".

To construct this performance variable, we need firm value information for two consecutive periods. Thus, for each CEO, $(Perform)_{it}$ is available only after the second period. This means that we discard CEOs who quit at the end of their first year. This potentially causes selection biases at the initial year, a problem addressed in our econometric models.

**Proxy variable for uncertainty:** To proxy for the uncertainty of business outcome, we use the ratio of R&D to the book value of assets. As Garen (1994) notes, the nature of R&D is that the outcome of the investment is less certain. Therefore, R&D to asset ratio can be used to proxy the riskiness of business outcome.

It should be noted that considerable number of firms do not report R&D figures. We assume that those firms do not make R&D expenditure. This is a standard assumption in the literature. For example, Garen (1992) and Bizjak et al. (1991) assume non-reporting R&D as zero. This assumptions can be justified since many firms who do not report such expenditures are firms that are not known as research intensive industries, such as retail or restaurant industries.

# 5 Data

This section first describe the data, then describes the sample criteria.

## 5.1 Data descriptions

Our data set is from ExecuComp and Compustat published by Standard and Poors. ExecuComp covers detailed information about the five most highly paid executives in each company within the S&P 500, S&P Midcap 400, and S&P SmallCap 600 firms. My study focuses on compensation only of CEOs. The sample covers the years 1993 to 2003. All the compensation figures in my study are deflated by the Consumer Price Index, with year 2003 as the base year.

The yearly average of total compensation is seen in Table 16. Again, the total compensation is defined as the sum of salary, bonuses and Black-Sholes values of stock options. The increase from 1993 to 2000 is rather dramatic. The average total compensation in 1993 is $1.93 million. This figure nearly quadrupled to $6.66 million in year 2000. However, average total compensation declined after year 2000. In fact, this is the year in which the information technology industry stock "bubble" burst. Table 17 shows the yearly averages separately for the three components. As can be seen, the value of stock options increases substantially to reach a peak in year 2000 and then declines sharply. Both salary and bonus show modest but steady increases through the whole sample period.

## 5.2 Sample Criteria

We treat each CEO-firm combination as a unique CEO. We also require that each individual became a CEO on or after year 1993, the year our sample period began, so that there are no left censored observations in our sample. The exclusion of the left censored observations is to eliminate the possibly severe left censoring biases.

After eliminating observations that do not match our sample criteria, we obtain an unbalanced panel data set that contains 3031 CEO-years of observation including 1075 corporations and 1450 CEOs from 1993 to 2003.

## 5.3 A note on CEO age

The CEO age variable needs some explanation. We found that the age variable in ExecComp (*p_age_2*) extremely incomplete. *p_age_2* shows the age of CEOs in the latest Corporate Proxy files. In our data set, we have 1430 CEOs. Of them, the *p_age_2* is missing for 1076 of these CEOs (76%). Moreover, the information is missing in non-random way: average exit rate for the observations with missing ages is 0.13, while it is 0.019 for observations whose ages are not missing. Age information is crucial in examining the relationship between incentive and tenure since the omission of age will confuse tenure effect with age effect. Thus, we hand-collected most of the age information. We have checked proxy statements for each corporation to check the CEO's age. As a result, among 1450 CEOs, we were able to collect age information for 1370 of them. Age variable shows the age of CEOs when the proxy statements are filed. Corporate proxy statements are available on line at Securities and Exchange Commission's home page (http://www.sec.gov). Table 18 shows the average exit rate of CEO according to age. As can be seen, the exit rate increases sharply at the age of 64.

### 5.4 A note on the construction of total compensation variable

Close examination of the ExecComp data set reveals that some CEOs do not quit at the end of the financial years. Rather, some CEOs leave before completing the last fiscal year. In such case, the CEO is rewarded for only a fraction of the fiscal year. However, we would like to use the annual equivalent amount of total compensation. Therefore, we adopt the following method of constructing the total compensation variable. Suppose we observe a CEO quitting his or her job. First, we check the number of months the CEO worked that year. If the number of months is equal or greater than 9 months, we compute the annual equivalent of total compensation by dividing the original compensation figure by (month worked)/12. If the number of months worked is fewer than 9, we disregard the incomplete year, and take the previous year as the final year.

## 6 Econometric models

Gibbons and Murphy's model (1993) assumes that CEO's total compensation is a linear function of firm performance. Thus, estimating the linear relationship of the following is appropriate.

$$(Total\ Comp)_{it} = constant + \beta(Perform)_{it} \tag{62}$$

The "extended-Prendergast" model assumes that firm uses compensation contract of the form, $(total\ comp) = output - constant$. This means that our model a priori assumes that $\beta = 1$. However, to empirically estimate the pay-for-performance relationship, let us assume a linear relationship given by equation (62).

We estimate three models. Model 1 is our basic model where a single total compensation equation is estimated using a fixed effect model. The problem associated with Model 1 is that this does not control for selection biases. Model 2 deals with such a problem by using a random effect model that includes two CEO turnover equations. Selection biases occur due to correlations among error terms between total compensation equation and the turnover equations. We include unobserved heterogeneity terms in the model to deal with such correlations. By the construction of data, it is possible that both Model 1 and Model 2 confound tenure effects with year effects as we only observe long-tenure CEOs during the later years of the sample period. Model 3 elaborates Model 2 to separate tenure effects from year effects. At the end of this section, we detail our choice of explanatory variables.

## 6.1 Model 1: Single equation fixed effect model

Our single equation fixed effect model is the following.

$$\begin{aligned} & \Delta log(Total\ Compensation)_{it} = \qquad (63) \\ + \quad & \beta_1 \Delta(Perform)_{it} I_{\{2 \leq tenure \leq 6\}} \\ + \quad & \beta_2 \Delta(Perform)_{it} I_{\{7 \leq tenure \leq 11\}} \\ + \quad & \beta_3 \Delta(Perform)_{it} I_{\{2 \leq tenure \leq 6\}} \times (years\ before\ CEO) \end{aligned}$$

$$+ \quad \beta_4 \Delta(Perform)_{it} I_{\{7 \le tenure \le 11\}} \times (years\ before\ CEO)$$

$$+ \quad \beta_5 \Delta(Perform)_{it} \times (R\&D\ Ratio)_{it}$$

$$+ \quad \Delta X_{it}^{other} \beta^{others} + \epsilon_i^{comp}$$

$i$ denotes individual CEO , $t$ denotes tenure as CEO (CEO tenure) where $t$=1 is the initial period and $\Delta$ denotes the first difference. As is noted earlier, $(Perform)_{it}$ is only available after the second period of tenure. This means that this regression discards the first period observation. Since we eliminate all the left censored observations, the maximum tenure is eleven.

To examine the relationship between pay-for-performance sensitivity and CEO tenure, we estimate separate pay-for-performance parameters, $\beta_1$ for shorter CEO tenure (tenure between 2 and 6), and $\beta_2$ for longer CEO tenure (tenure between 7 and 11). Standard principal-agent theories predict $\beta_1 < \beta_2$. Our "extended-Prendergast" model predicts the opposite.

"Years before CEO" is a variable that shows the number of years the individual has worked for the company prior to becoming a CEO. The coefficients for the interaction terms between $(Years\ before\ CEO)_{it}$ and $(Perform)_{it}$ are used to examine the relationship between incentive and uncertainty about the CEO's managerial ability ($\beta_3$ and $\beta_4$). Again, we allow such uncertainty to have different impacts for shorter tenure and longer tenure. This is because the effect of having prior knowledge of the CEO's managerial ability may diminish as more information about the managerial ability becomes available as the CEO stays with the firm. If uncertainty about CEO ability negatively affects pay-for-performance sensitivity like standard principal-agent theories, $\beta_3$ and $\beta_4$ should be positive. On the other hand, our "extended-Prendergast" model implies that those coefficients should be non-positive.

$\beta_5$ is the coefficient for the interaction term between return and R&D asset ratio. Thus, $\beta_5$ shows the relationship between incentive and uncertainty under which the firm operates. Standard theories predict $\beta_5$ to be negative while "extended-Prendergast" model predicts $\beta_5$ to be non-negative.

The problem with this model is selection biases. CEO Turnover is common; therefore, for each period, we observe individuals conditional on these CEOs not quitting up to that period. This causes a potential selection biase. Moreover, we discard the first period observation from this model since we do not have a $(Perform)_{it}$ variable until the second period. This causes a initial period selection biase. Model 2 deals with such selection biases using an unobserved heterogeneity model.

## 6.2 Model 2: Unobserved heterogeneity model to deal with selection biases.

To address the problem associated with the selection biases, suppose that CEO turnover is given by the following equation.

$$
\begin{aligned}
y_{it} &= \alpha' Z_i + \mu_{it} \\
&\quad where \\
&\quad CEO\ leaves\ at\ the\ end\ of\ the\ t^{th}\ tenure\ if\ y_{it} > 0 \\
&\quad CEO\ stays\ if\ y_{it} \leq 0
\end{aligned}
$$

To obtain an unbiased estimate for $\beta$ in equation (63), the error term should be uncorrelated with the explanatory variables. Let $X_{it}$ be the vector of all the explanatory variables in equation (63). Since we only use CEOs who survive up to period t, the conditional expectation of the correlation between the error term and the explanatory variables becomes

$$E[\Delta X_{it} \cdot \epsilon_{it}^{comp} | \alpha Z_{it} + \mu_{it} \leq 0] \tag{64}$$

Assuming common variables exist between $X_{it}$ and $Z_{it}$, if $\epsilon^{comp}$ and $\mu_{it}$ are correlated, (64) may not be zero. Thus, OLS estimation of equation (63) would produce biased results.

Correlation between the error terms is caused by the existence of unobserved heterogeneity: if there is an unobserved variable that affects both the total compensation equation (63) and CEO turnover, the error terms will be correlated. The following model deals with such correlations by explicitly incorporating an unobserved heterogeneity term in the model.

$$\begin{aligned} Compensation\ equation \quad &: \quad \Delta log(Total\ Compensation)_{it} = \Delta X_{it}'\beta \qquad (65) \\ &\quad +(\rho^{exit}\chi_i + \epsilon_{it}^{comp}) \\ Turnover\ equation \quad &: \quad y_{it} = Z_{it}'\alpha + (\rho^{exit}\chi_i + \mu_{it}), \quad t \geq 2 \qquad (66) \\ Initial\ period\ turnover \quad &: \quad y_i^{init} = W_i'\theta + (\rho^{init}\chi_i + \mu_i^{init}) \qquad (67) \end{aligned}$$

This system of equations includes total compensation equation, CEO turnover equation for tenure greater or equal to 2 , and initial period turnover equation.[17] $y_{it}$ in the

[17]Total compensation equation is defined for $t \geq 3$ due to the facts that we only have $(Perform)_{it}$

turnover equation (66) is a latent variable such that if $y_{it} > 0$, the CEO exit at the end of period t, and stay in the firm otherwise. We include $(Perform)_{it}$ in the turnover equation, thus turnover equation is defined only for CEO tenure greater than 2. $y_i^{init}$ in equation (67) is the latent variable such that if $y_i^{init} > 0$, the CEO quits at the end of initial period, and stay otherwise. $\mu_{it}$ and $\mu_i^{init}$ are assumed to be distributed as standard logistic.

Turnover equation controls for the selection biases for $t \geq 2$. As noted earlier, by the construction of data, our total compensation equation discards the first period observation. This causes potential selection biases at the initial period. Inclusion of the initial period turnover equation controls for such biases.

We assume that $\epsilon_{it}^{comp} \perp \epsilon_{it}^{mix} \perp \mu_{it} \perp \mu_i^{init}$. $\chi_i$ is a CEO-firm match specific unobserved explanatory variables that is assumed to be distributed as standard normal. $\rho^{comp}$, $\rho^{mix}$, $\rho^{exit}$ and $\rho^{init}$ are the factor loads. Since we do not observe $\chi_i$, this is a part of the error term in each equation. For instance, the error term in equation (65)is $(\rho^{mix}\chi_i + \Delta\epsilon_{it}^{mix})$. In other words, the correlation among error terms are captured by $\chi_i$ when factor loads are not all equal to zero. Thus, inclusion of turnover equation controls for the selection biases for the period $t \geq 2$. The initial turnover equation controls for the selection biases at the initial period. The estimation of the system of equations is by maximum likelihood. The likelihood function can be found in Appendix **??**.

from the second period and that we take the first difference.

## 6.3 Model 3: Separating tenure effects from year effects

In both Model 1 and Model 2, we separately estimate pay-for-performance sensitivity for CEOs with shorter tenure and for CEOs with longer tenure. A potential problem with such method is that, the coefficients may pick up "year effects." For example, suppose that pay-for-performance sensitivity has systematically changed for all the firms after the year 2000. By construction of the data, a considerable number of observations whose CEO tenure is between 7 and 11 are in the period between year 2001 and 2003. Thus, $\beta_1$ and $\beta_2$ may pick up such year effects rather than tenure effect.

The assumption that pay-for-performance sensitivity has changed after year 2000 is not unreasonable. As is described in Section 5, total compensation declines sharply after year 2000. This year coincides with the IT industry bubble burst. It is likely that such a drastic decline in total compensation is accompanied by a systematic change in corporate governance. To separate tenure effects from potential year effects, we replace the total compensation equation in (65) by the following equation.

$$\Delta log(Total\ Compensation)_{it} = \tag{68}$$

$$+ \quad \beta_{11}\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}I_{\{1993\leq year\leq 2000\}}$$

$$+ \quad \beta_{12}\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}I_{\{2001\leq year\leq 2003\}}$$

$$+ \quad \beta_{21}\Delta(Perform)_{it}I_{\{7\leq tenure\leq 11\}}I_{\{1993\leq year\leq 2000\}}$$

$$+ \quad \beta_{22}\Delta(Perform)_{it}I_{\{7\leq tenure\leq 11\}}I_{\{2001\leq year\leq 2003\}}$$

$$+ \quad \beta_{31}\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}I_{\{1993\leq year\leq 2000\}} \times (years\ before\ CEO)$$

$$+ \quad \beta_{32}\Delta(Perform)_{it} I_{\{2 \le tenure \le 6\}} I_{\{2001 \le year \le 2003\}} \times (years\ before\ CEO)$$

$$+ \quad \beta_{41}\Delta(Perform)_{it} I_{\{7 \le tenure \le 11\}} I_{\{1993 \le year \le 2000\}} \times (years\ before\ CEO)$$

$$+ \quad \beta_{42}\Delta(Perform)_{it} I_{\{7 \le tenure \le 11\}} I_{\{2001 \le year \le 2003\}} \times (years\ before\ CEO)$$

$$+ \quad \beta_{5}\Delta(Perform)_{it} \times (R\&D\ Ratio)_{it}$$

$$+ \quad \Delta X_{it}^{other} \beta^{others} + \epsilon_{it}^{comp}$$

This means that for CEOs with shorter tenure, we estimate two separate pay-for-performance parameters, $\beta_{11}$ for early years (year between 1993 and 2000) and $\beta_{21}$ for late years (year between 2001 and 2003). Similarly, for longer tenure, we estimate two separate pay-for-performance parameters, $\beta_{12}$ for early years, and $\beta_{22}$ for late years. By comparing the pay-for-performance sensitivities for shorter tenure and longer tenure either within early years or later years, we can avoid confounding tenure effects with year effects.

## 6.4 Choice of explanatory variables

Table 4 is the complete list of our explanatory variables. $\Delta X_{it}$ for $t \geq 3$ is the set of variables used for total compensation equation (65)[18]. $W_i$ is the variables for initial turnover equation (67).

The set of variables used for the turnover equation (66), $Z_{it}$, is $[\Delta X_{it}, X_{it}, X_{i2}, W_i]$ for $t \geq 2$. $\Delta X_{it}$ at t=2 is set at zero in the turnover equation. There are three vectors of level variables, $X_{it}$, $X_{i2}$ and $W_i$. We include $X_{it}$ since CEO turnover is affected not only by

[18]Constant variables or variables whose first differences are collinear with the constant term are omitted from $\Delta X_{it}$.

"growth" of these explanatory variables, but also by the level of those variables. $X_{i2}$ and $W_i$ are vectors of initial condition. variables.[19] By including initial condition variables, we can incorporate total compensation and option mix in a reduced form specification. Option mix is the proportion of stock options in the total compensation. In Chapter 1, we show that total compensation and option mix are important determinants of CEO turnover. See also Appendix **??**.

Inclusion of some variables needs explanation. The market to book asset ratio is commonly used as the proxy for growth opportunity for the following reason. The book value of an asset can be interpreted as the value of an asset that is already in place. On the other hand, market value of an asset is the value of an asset which will be in place in the future plus the value already in place. Therefore, the market to book asset ratio reflects the perceived growth opportunity of the firm. Since it is the perceived growth opportunity of the firm, this would affect total compensation.

Sales are typically used as proxy for firm size. Larger size is typically associated with greater compensation. Murphy (1985) shows the importance of controlling for the firm size when estimating wage equation. Size of the firm may also affect the increase in wage because managing a larger firm requires greater managerial skills, and to retain the talent, the firm may have to compensate him with greater wage increase. Firm size is also associated with greater scrutiny by media, which could affect the turnover of the CEOs.

Dividend yield is an inverse proxy for the investment opportunity set. Firms with

[19]Since we include those initial condition variables, we set $X_{it}$ equal to zero at t=2.

high investment opportunity spend more money for investment, thus cash returned to investors is typically small.

Macroeconomic trends also influence CEO compensation. For example, a good macroeconomic condition would lead to greater CEO compensation. Therefore, we want to control for such macroeconomic trends. One method is to include GDP figure. However, GDP figure may have different impacts for different industries. For example, an increase in GDP and a subsequent increase in price level may have opposite effect on export oriented industry and import oriented industry. Therefore, a better way to control for such trends is to incorporate the industry average of total compensation. Table 20 shows summary statistics of some of important variables.

# 7 Main empirical results

This section summarizes our empirical findings. We begin with general description of our estimation results in Section 7.1, then analyze the relationship between incentive and tenure in Section 7.2. In section 7.3, we interpret estimated pay-for-performance parameters. Section 7.4 exhibits the results for the relationship between incentive and uncertainty of the managerial ability of CEOs. This section also shows the results of the relationship between incentive and the riskiness of the environment under which a firm operates.

## 7.1 Estimation results across the models

Table 21 presents the estimated coefficients for the total compensation equation for Model 1 (fixed effect model), Model 2 (heterogeneity model) and Model 3 (unobserved heterogeneity model that separates year effects). The estimated coefficients for Model 1 and Model 2 are never substantially different from each other. In fact, the factor load for the total compensation equation is not significantly different from zero at 5% confidence level. Significant effects of heterogeneity only comes from the correlation between the turnover equation (66) and initial turnover equation (67). Note that, Model 1 is nested in Model 2 given the restriction that all the factor loads are equal to zero. However, $\chi^2$ test rejects such a hypothesis.[20] This means that we reject Model 1 in favor of Model 2. Note also that Model 2 is nested in Model 3 given the restriction that there are no year effects. Thus, we base our analysis on Model 2 and Model 3 only.

## 7.2 Incentive-tenure relationship:

For Model 2, estimated pay-for-performance parameters appear to be much larger for longer CEO tenure than shorter CEO tenure. $\widehat{\beta}_1$ is 0.002 while $\widehat{\beta}_2$ is 0.28. Hat is a notation to indicate that it is an estimated coefficient. Although $\widehat{\beta}_1$ appears to be imprecisely estimated, $\widehat{\beta}_2$ is still much larger than the upper bound of the 5% confidence interval for $\widehat{\beta}_1$. Thus, pay-for-performance sensitivity appears to be increasing with CEO tenure, which is consistent with standard principal-agent theories. To formally test if pay-for-performance increases with tenure, consider the null hypothesis $H_0: \ \beta_1 \geq \beta_2$

[20] $\chi^2$ test statistics for null hypothesis that all factor loads are zero is 25.55.

with an alternative hypothesis $H_1: \ \beta_1 < \beta_2$. Notice that,

$$\frac{(\widehat{\beta}_2 - \widehat{\beta}_1) - (\beta_1 - \beta_2)}{\sqrt{var(\widehat{\beta}_2) + 2cov(\widehat{\beta}_1, \widehat{\beta}_2) + var(\widehat{\beta}_2)}} \sim t_{N-K}$$

where N is the number of observations and K is the number of parameters to be estimated. This is a simple one-tail t test. The test statistics is 3.85, thus, null hypothesis is rejected at any conventional confidence level. The results of Model 2 strongly support standard principal-agent theories like Gibbons and Murphy's (1992)

Model 3 separates year effects from tenure effects, thus it is a more general model. First, let us compare the pay-for-performance parameters for the period between year 1993 and 2000. $\beta_{11}$ is the pay-for-performance parameter for shorter tenure, and $\beta_{21}$ for longer tenure. $\widehat{\beta}_{11}$ is 0.002 with standard error equal to 0.056 while $\widehat{\beta}_{21}$ is 0.54 with the standard error equal to 0.100. We reject the null hypothesis $H_0 : \beta_{11} \geq \beta_{21}$ at 2.5% confidence level.[21] Thus, for the period between year 1993 and 2000, we find strong evidence that pay-for-performance sensitivity increases with tenure.

For the period between year 2000 and 2003, the estimated pay-for-performance parameters are -0.03 for shorter tenure ($\widehat{\beta}_{12}$) and 0.16 for longer tenure ($\widehat{\beta}_{22}$). A negative estimate for $\beta_{12}$ may be counter intuitive, but the magnitude of the coefficient is negligibly small. Thus pay-for-performance parameters appear to be increasing with CEO tenure. However, both parameters are rather imprecisely estimated so that null hypothesis, $\beta_{12} \geq \beta_{22}$, cannot be rejected at 5% confidence level.

The longer-tenure pay-for-performance parameter for the period between 1993 and

[21] Test statistics is 4.543.

2000 ($\widehat{\beta}_{21}$) is larger than the longer-tenure pay-for-performance parameter for Model 2 ($\widehat{\beta}_2$) by more than 90%. This may indicate possible year effects. In fact, Model 2 is nested in Model 3 given the restriction of no year effects [22]. However, $\chi^2$ test statistic is 7.85. Since the critical value for 5% confidence level with degree of freedom 10 is 18.31, we fail to reject the "no year effects" hypothesis. Therefore, Model 2 becomes the most relevant for our analysis.

In sum, both Model 2 and Model 3 provides evidence that pay-for-performance sensitivity increases with CEO tenure. Therefore, our results strongly support standard principal-agent theories. In addition, existence of "year effects" is rejected, making Model 2 the most relevant model for our analysis. In the following, we compute the magnitude of pay-for-performance sensitivity in a more informative way. We base our analysis on Model 2.

## 7.3 Magnitudes of pay-for-performance sensitivities.

To interpret the pay-for-performance parameters in a more informative way, we compute the change in total compensation when firm performance changes from the median level to the 75th percentile level or to the 25th percentile level. Table 22 shows such results. First, consider the pay-for-performance sensitivity for shorter tenure. The median level of total compensation for CEOs at the 4th year of tenure is \$2.5 million. The estimation result, $\widehat{\beta}_1 = 0.0022$, suggests that, starting from the median total compensation, if firm

[22] More precisely, the restrictions are $\beta_{11}=\beta_{12}$, $\beta_{21}=\beta_2$, $\beta_{31}=\beta_{32}$ and $\beta_{41}=\beta_{42}$. To be consistent, we impose the same "no year effects" restrictions on the initial condition variables. The total number of constraints is ten.

performance improves from its median level (0.07) to its 75 percentile level (0.32), total compensation would increase by \$1,500. This is an increase in total compensation of about 0.06%. If firm performance deteriorates from the median to the 25 percentile level (-0.17), total compensation decreases by \$1,300, which is a 0.05% decrease in total compensation. This shows that for shorter tenure, pay-for-performance sensitivity is almost non-existent.

Now consider the pay-for-performance sensitivity for longer tenure. The median total compensation for CEOs with 10 years of experience is \$4 million. Starting at this level, total compensation increases by \$320,000 if firm performance improves from the median level (0.07) to the 75 percentile level (0.34). This is approximately 8% increase in total compensation. If firm performance deteriorates from its median level to the 25 percentile level (-0.17), total compensation would decrease by \$277,000, which is a decrease of 6.4%. This result again confirms that pay-for-performance sensitivity is significantly higher for CEOs with longer tenure, providing strong supporting evidence for standard principal-agent theories.

## 7.4 Relationship between incentive and uncertainty about managerial ability of CEOs, and between incentive and risk.

$\beta_3$ and $\beta_4$ in Model 2 shows the relationship between the incentive and uncertainty about a CEO's managerial ability. Standard principal-agent theories such as Gibbons and Murphy (1992) predict that both coefficients are positive. As can be seen from Table 21, we have positive estimates for $\beta_3$ (coefficient for shorter tenure), but negative estimates for $\beta_4$ (coefficient for the longer tenure). The magnitude of both coefficients are quite small as are their standard errors. In fact, if either coefficients are as large as the upper bounds of the 95% confidence intervals (0.01, 0.002), one would be hard pressed to consider these effects as substantial. For example, consider two CEOs, one with Years-before-CEO =0 (hired from outside); the other with Years-before-CEO =5 (having worked for the firm for 5 years). Based on the upper bound of 95% confidence interval for $\widehat{\beta}_3$ (0.01), an improvement in firm performance from the median level to the 75th percentile level would lead to an 0.06% increase in total compensation for the CEO with Years-before-CEO=0, while it only leads to 1.5% increase for the CEO with Years-before-CEO=5.[23] Pay-for-performance sensitivity is greater for CEOs with Years-before-CEO =5 only by 1.46 percentage points. This seems quite small. Therefore, we find little evidence of the existence of the relationship between incentive and uncertainty about the CEO's managerial ability.

$\beta_5$ shows the relationship between incentive and the riskiness of the environment in which the firm operates. Although the estimated coefficient is negative as standard principal-agent theories predict, the estimate is extremely small (-0.05). Even the upper bound of the 95% confidence interval estimate (0.38) means that the effect of R&D ratio on the pay-for-performance is almost non-existent: The median and 75 percentile levels of R&D ratios are 0.001 and 0.035 respectively. When firm performance improves from the median level to the 75 percentile level, total compensation for a CEO at a

[23]Change in total compensation is computed assuming that the CEO is at the 4th year of tenure, having median compensation of $ 2.5 million.

firm with median R&D increases by 0.06% while total compensation for a CEO at a firm with 75 percentile level R&D ratio increases by 0.04%. The difference in pay-for-performance sensitivity is only 0.02 percentage points. Therefore, we find little evidence of the existence of the relationship between incentive and uncertainty in which the firm operates.

In sum, the examination of relationships between (a) incentive and the uncertainty of CEO's managerial ability, and (b) incentive and risk do not provide further evidence to support or reject both standard principal-agent theories and our extended-Prendergast Model.

# 8 Conclusion

The main objective of this study is to examine the relationship between pay-for-performance sensitivity and CEO tenure. Using the most recent CEO compensation data, we find that pay-for-performance sensitivity is significantly greater for CEOs with longer tenure. For CEOs with tenure between 7 and 11 years, an improvement in firm performance from the median level to the 75th percentile level would lead to as much as an 8% increase in total compensation. On the other hand, for CEOs with tenure equal to or fewer 6 years, the same improvement in firm performance only leads to a 0.06% increases total compensation. Thus, pay-for-performance sensitivity dramatically increases with CEO tenure. This is a strong support for standard principal-agent theories like Gibbons and Murphy (1992), but rejects our "extended-Prendergast" model. In addition, we examine the relationship between incentive and uncertainty about a CEO's managerial ability, and the relationship between incentive and riskiness of the environment in which the firm operates. The magnitudes of the estimated coefficients for both relationships are very small. Such relationships appear to be non-existent or minute at most. Thus, these two additional tests do not provide further evidence to support or reject both standard principal-agents theories and our extended-Prendergast model.

Table 16: Yearly average of total compensation

| Year | # of obs | Average total comp |
|---|---|---|
| 1993 | 1330 | 1.93 |
| 1994 | 1495 | 2.26 |
| 1995 | 1572 | 2.21 |
| 1996 | 1639 | 2.89 |
| 1997 | 1699 | 3.59 |
| 1998 | 1627 | 4.09 |
| 1999 | 1653 | 5.36 |
| 2000 | 1656 | 6.66 |
| 2001 | 1590 | 5.91 |
| 2002 | 1575 | 4.21 |
| 2003 | 1494 | 3.46 |

Compensation figures are in million dollars

Table 17: Yearly average of salary, bonuses and options

| Year | # of obs | Salary | Bonus | Option |
|---|---|---|---|---|
| 1993 | 1330 | 0.63 | 0.50 | 0.79 |
| 1994 | 1495 | 0.63 | 0.55 | 1.09 |
| 1995 | 1572 | 0.63 | 0.56 | 1.03 |
| 1996 | 1639 | 0.63 | 0.67 | 1.59 |
| 1997 | 1699 | 0.62 | 0.66 | 2.31 |
| 1998 | 1627 | 0.62 | 0.65 | 2.81 |
| 1999 | 1653 | 0.63 | 0.75 | 3.98 |
| 2000 | 1656 | 0.64 | 0.78 | 5.24 |
| 2001 | 1590 | 0.65 | 0.67 | 4.60 |
| 2002 | 1575 | 0.66 | 0.70 | 2.84 |
| 2003 | 1494 | 0.67 | 0.83 | 1.95 |

Compensation figures are in million dollars

Table 18: Average exit rate of CEOs by Age

| Age of CEOs | # of obs | Average exit rate |
|---|---|---|
| 51 | 150 | 0.047 |
| 52 | 177 | 0.045 |
| 53 | 189 | 0.085 |
| 54 | 198 | 0.061 |
| 55 | 198 | 0.111 |
| 56 | 199 | 0.080 |
| 57 | 195 | 0.087 |
| 58 | 180 | 0.1 |
| 59 | 160 | 0.086 |
| 60 | 136 | 0.080 |
| 61 | 123 | 0.146 |
| 62 | 88 | 0.159 |
| 63 | 63 | 0.132 |
| 64 | 58 | 0.379 |
| 65 | 41 | 0.341 |
| 66 | 22 | 0.364 |
| 67 | 15 | 0.333 |

Table 19: Choice of explanatory variables

| Variables that affect total compensation ($X_{it}$) | Initial exit variables ($W_i$) |
|---|---|
| $(Perform)_{it} I_{\{2 \le tenure \le 6\}}$ | $log(Market\ to\ book\ asset\ ratio)_t$ |
| $(Perform)_{it} I_{\{7 \le tenure \le 11\}}$ | $(\%\ change\ in\ sales)_t$ |
| $(Perform)_{it} I_{\{2 \le tenure \le 6\}} \times (years\ before\ CEO)$ | $(Interlocking\ directirship\ dummy)_t$ |
| $(Perform)_{it} I_{\{7 \le tenure \le 11\}} \times (years\ before\ CEO)$ | $log(Sales)_t$ |
| $(return)_{it} \times (R\&D\ Ratio)_{it}$ | $(\frac{R\&D}{assets})_t$ |
| $log(Market\ to\ book\ asset\ ratio)_{t-1}$ | $Dummy\{R\&D = 0\}_t$ |
| $\Delta log(market\ to\ book\ asset\ ratio)_t$ | $log(stockpricevolatility)_t$ |
| $(Interlocking\ directirship\ dummy)_t$ | $(Yeardummies)_t$ |
| $(Inter\ locking\ directorship\ dummy)_{t-1}$ | Industry Dummies |
| CEO Tenure | $log(dividend\ yeild + 1)_t$ |
| $(CEO\ Tenure)^2$ | $(industry\ average\ total\ comp)_t$ |
| $log(sales)_t$ | $(industry\ average\ Option\ mix)_t$ |
| $log(sales)_{t-1}$ | Dummies for S&P 500 |
| $(\%\ change\ in\ sales)_t$ | Dummies for S&P Mid cap |
| $(\%\ change\ in\ sales)_{t-1}$ | Years before CEO |
| $(\frac{R\&D}{assets})_t$ | Dummy{(Years before CEO) is missing} |
| $(\frac{R\&D}{assets})_{t-1}$ | $(Age)_t$ |
| $(Age)_t$ | $(Age)^2_t$ |
| $(Age)^2_t$ | $Dummy\{Age\ over\ 64\}_t$ |
| $Dummy\{Ageover64\}_t$ | 3 years Return to equity |
| $Dummy\{R\&D = 0\}_t$ | Number of Board meetings |
| $Dummy\{R\&D = 0\}_{t-1}$ | $Dummy\{Age\ missing\}$ |
| $log(stockpricevolatility)_t$ | |
| $log(stockpricevolatility)_{t-1}$ | |
| Year dummies | |
| $(Interlocking\ Directorship\ Dummies)_t$ | |
| $(Interlocking\ Directorship\ Dummies)_{t-1}$ | |
| $log(dividend\ yeild + 1)_t$ | |
| $log(dividend\ yeild + 1)_{t-1}$ | |
| $(industry\ average\ total\ comp)_t$ | |
| $(industry\ average\ total\ comp)_{t-1}$ | |
| $(industry\ average\ Option\ mix)_t$ | |
| $(industry\ average\ Option\ mix)_{t-1}$ | |

Table 20: Data summary statistics (1993 - 2003)

| Variable | Mean | St div | Min | Max |
|---|---|---|---|---|
| Perform | 0.35 | 10.18 | -1.0 | 582.85 |
| $\frac{R\&D}{Assets}$ | 0.03 | 0.07 | 0 | 1.18 |
| Sales (in $ million) | 6.63 | 18.07 | $9.4 \times 10^{-5}$ | 257.16 |
| Age | 53.01 | 6.99 | 33 | 81 |
| CEO Tenure | 2.9 | 1.97 | 1 | 11 |
| Years before CEO | 5.20 | 9.31 | 0 | 46 |

Table 21: Estimation results by models

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| $\beta_1$: $\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}$ | 0.0022 | 0.0022 | – |
| | (0.00014) | (0.018) | |
| $\beta_2$:$\Delta(Perform)_{it}I_{\{7\leq tenure\leq 11\}}$ | 0.282 | 0.281 | – |
| | (0.188) | (0.068) | |
| $\beta_{11}$:$\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}I_{\{1993\leq year\leq 2000\}}$ | – | – | 0.0022 |
| | | | (0.020) |
| $\beta_{12}$:$\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}I_{\{2001\leq year\leq 2003\}}$ | – | – | -0.032 |
| | | | (0.041) |
| $\beta_{21}$: $\Delta(Perform)_{it}I_{\{7\leq tenure\leq 11\}}I_{\{1993\leq year\leq 2000\}}$ | – | – | 0.544 |
| | | | (0.012) |
| $\beta_{22}$: $\Delta(Perform)_{it}I_{\{7\leq tenure\leq 11\}}I_{\{2001\leq year\leq 2003\}}$ | – | – | 0.157 |
| | | | (0.105) |
| $\beta_3$: $\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}$ | 0.005 | 0.005 | – |
| $\times(years\ before\ CEO)$ | (0.002) | (0.0028) | |
| $\beta_4$:$\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}$ | -0.004 | -0.005 | – |
| $\times(years\ before\ CEO)$ | (0.009) | (0.007) | |
| $\beta_{31}$: $\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}I_{\{1993\leq year\leq 2000\}}$ | – | – | 0.0038 |
| $\times(years\ before\ CEO)$ | | | (0.003) |
| $\beta_{32}$: $\Delta(Perform)_{it}I_{\{2\leq tenure\leq 6\}}I_{\{2001\leq year\leq 2003\}}$ | – | – | 0.0097 |
| $\times(years\ before\ CEO)$ | | | (0.007) |
| $\beta_{41}$: $\Delta(Perform)_{it}I_{\{7\leq tenure\leq 11\}}I_{\{1993\leq year\leq 2000\}}$ | – | – | -0.013 |
| $\times(years\ before\ CEO)$ | | | (0.013) |
| $\beta_{42}$: $\Delta(Perform)_{it}I_{\{7\leq tenure\leq 11\}}I_{\{2000\leq year\leq 2003\}}$ | – | – | -0.005 |
| $\times(years\ before\ CEO)$ | | | (0.023) |
| $\beta_5$: $\Delta(Perform)_{it}\times(R\&D\ Ratio)_{it}$ | -0.044 | -0.047 | -0.020 |
| | (0.136) | (0.219) | (0.526) |
| $\rho^{comp}$ | – | 0.069 | 0.073 |
| | | (0.045) | (0.046) |
| $\rho^{exit}$ | – | 2.274 | 5.078 |
| | | (0.559) | (0.502) |
| $\rho^{init}$ | – | -13.645 | -17.440 |
| | | (7.520) | (22.007) |

Inside the brackets are standard errors. For Model 1, they are robust standard errors. Likelihood for Model 2 is -3403.604. Likelihood for Model 3 is -3399.678.

Table 22: Magnitudes of Pay-for-performance sensitivity

| CEO tenure | Change in total Comp when Perform changes from the median to 25th percentile level | Change in total comp when Perform changes from the median to 25th percentile level |
|---|---|---|
| Shorter tenure<br>(4th year of tenure) | -$1,300<br>(-%0.05 change) | +$1,500<br>(+%0.06 change) |
| Longer tenure<br>(10th year of tenure) | -$277,000<br>(-%6.4 change) | +$320,000<br>(+% 8 change) |

## Appendices

# A Variable definitions

$$\begin{aligned}
\Delta shareholders' wealth &= \frac{(firm\ value)_t - (firm\ value)_{t-1} + dividend\ payout}{(firm\ value)_{t-1}} \\
(Market\ to\ book\ asset\ ratio)_t &= [Assets - (Total\ common\ equity) \\
&\quad + (Share\ outstanding) \times (Share\ closing\ price)] \\
&\quad \div Asset
\end{aligned}$$

If CEO i is in industry j and number of firms in the industry j at time t is $N_{jt}$ then

$$\begin{aligned}
(Industry\ average\ total\ comp)_{it} &= \frac{1}{N_{jt}-1} \sum_{s\ is\ in\ industry\ j,\ j \neq i} (Total\ compensation)_{st} \\
(Industry\ average\ Option\ mix)_{it} &= \frac{1}{N_{jt}-1} \sum_{s\ is\ in\ industry\ j,\ j \neq i} (Option\ mix)_{st} \\
R\&D\ ratio &= \frac{R\&D}{Assets}
\end{aligned}$$

**Interlocking directorship :** This generally involves one of the following situations.

- The officer serves on the board committee that makes his compensation decisions.
- the officer serves on the board (and possibly compensation committee) of another company that has an executive officer serving on the compensation committee of the indicated officer's company.
- the officer serves on the compensation committee of another company that has an executive officer serving on the board (and possibly compensation committee ) of the indicated officer's company.

**Outside succession dummy:** This takes the value 1 if the individual is employed by the firm for 2 years or less prior to becoming an CEO.

## B Sample criteria

We treat CEO-firm combination as a unique CEO. If an individual works for two firms, he or she counts as two CEOs. When a CEO works for multiple firms during the sample period, two situations can arise. The first case is where the CEO works for multiple firms at the same time. The second case is where the CEO quits the first firm and moves to the second, so that there is no overlap in working periods. The first case is likely to occur due to mergers. To eliminate a possible merger effect in turnover, I drop all the CEOs that fall under this category. For the second case, I drop the CEOs who worked at three or more firms, so that in my sample, the number of firms at which each individual works is at most two. The reason for such elimination is that it is important to keep some degree of homogeneity in the sample since this is the study of CEO duration.

In addition, we require that for each CEO the observations be of consecutive years; if any variables are missing in a particular year, all the observations after the first missing observation are dropped. There are two reasons for this requirement. First, we want to avoid imputing values for missing observations. Second, missing observations sometimes occur because a CEO has a "skipped" year, i.e. a situation where a whole year is missing between the year of the CEO's first appearance in the sample and the last year in the sample. This may occur because a CEO leaves the firm and is rehired later. Requiring the observation to be consecutive eliminates such a possibility, keeping some homogeneity in the samples.

Finally, when the CEO exits the firm, it is not necessary that the CEO completes the last fiscal year. If the CEO complete full 9 months of the fiscal year, we treat the actual year as the year of separation. For such a case, we compute the full year equivalent amount of total compensation as $(Actual\ total\ comp) \times \frac{12}{Months\ worked}$. If the CEO does not complete full 9 months, we treat the previous year as the year of separation.

# C Proof of proposition 1

Notice that, $max\{\theta_{1i} + \rho_{1i}, i = 1, 2\}$ is the second order statistic of two independent samples. One of the arguments has distribution (36) and the other has distribution (37).

We can show that (49) holds. Let Y, Y' be independent random variables from $N(\mu, \sigma^2 + \tau^2)$. We have $E[max\{Y - \mu, Y' - \mu\}] = \sqrt{\frac{\sigma^2+\tau^2}{\pi}}$. It suffices to show that $E[max\{\theta_{1i} + \rho_{1i}, i = 1, 2\}] - S_1 < E[max\{Y - \mu, Y' - \mu\}]$

We show this in two steps. First, let W and Z be independent random variables having the distribution functions (36) and (37) respectively. Let V be a random variable drawn from $N(\mu, \sigma^2 + \tau^2)$. We have,

$$E[max\{W - S_1, Z - S_1\}] = E[max\{\theta_{1i} + \rho_{1i}, i = 1, 2\}] - S_1 \qquad (69)$$

We treat $S_1$ as a constant rather than a random variable at this point. The random variable V stochastically dominates the random variable Z since $P(\theta_i \leq t | \theta_i + \rho_{0i} \leq y_0) \geq P(\theta_i \leq t)$. This means that $E[max\{W - S_1, V - S_1\}]$ also stochastically dominates

$E[max\{W - S_1, Z - S_1\}]$, i.e., $E[max\{W - S_1, V - S_1\}] \geq E[max\{W - S_1, Z - S_1\}]$.

Both $max\{W - S_1, Z - S_1\}$ and $max\{Y - \mu, Y' - \mu\}$ are the second order statistics of two independent random variables from non-identical normal distributions. By noticing that $W - S_1$ has zero mean but lower variance than $y - \mu$ and that $Z - S_1$ has the same variance as $Y - \mu$ but lower mean, we can see that [24]

$$E[max\{W - S_1, V - S_1\}] < E[max\{Y - \mu, Y' - \mu\}] \tag{70}$$

By gathering all the results, we have

$$E[max\{\theta_{1i} + \rho_{1i}, i = 1, 2\}] - S_1 = E[max\{W - S_1, V - S_1\}] \tag{71}$$

$$\leq E[max\{W - S_1, V - S_1\}] < E[max\{Y - \mu, Y' - \mu\}] \tag{72}$$

$$= \sqrt{\frac{\sigma^2 + \tau^2}{\pi}} \tag{73}$$

This proves the proposition.

# D Proof for proposition 2

The distribution of $\theta_{ti}$ is not normal because of the conditioning sets of the following form, $\theta_{si} + \rho_{si} + e^* \leq y_s$. If all the conditioning sets had the following form $\theta_{si} + \rho_{si} + e^* = y_s$, the $\theta_{ti}$ would have a normal distribution. Since it is easier to deal with a normal distribution, we define $\phi_{ti}$ by eliminating the conditioning set of the form,$\theta_{si} + \rho_{si} + e^* \leq y_s$ from $\theta_{ti}$. For example, if

$$\theta_{3i} \sim [\theta_1 \mid \theta_1 + \rho_{0i} + e^* = y_0,\ \theta_1 + \rho_{1i} + e^* \leq y_1,\ \theta_1 + \rho_{2i} + e^* = y_2]$$

[24] if $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$, and if X, Y are independent, then $E[max\{X, Y\}] = \mu\Phi(\frac{\mu_X - \mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}}) + \mu_Y \Phi(\frac{\mu_Y - \mu_X}{\sqrt{\sigma_X^2 + \sigma_Y^2}}) + \sqrt{\sigma_X^2 + \sigma_Y^2} \frac{1}{\sqrt{2\pi}} exp[-\frac{1}{2} \frac{(\mu_X - \mu_Y)^2}{\sigma_X^2 + \sigma_Y^2}]$. This is increasing in either $\mu_X$, $\mu_Y$,$\sigma_X$ or $\sigma_Y$.

then

$$\phi_{3i} \sim [\theta_1 | \theta_1 + \rho_{0i} + e^* = y_0,\ \theta_1 + \rho_{2i} + e^* = y_2]$$

$\theta_{ti}$ is stochastically dominated by $\phi_{ti}$. Therefore, we have the following relationship.

$$E[max\{\theta_{t1} + \rho_{ti}, \theta_{t2} + \rho_{t2}\}] - max\{E[\theta_{t1} + \rho_{t1}], E[\theta_{t2} + \rho_{t2}]\}$$

$$\leq$$

$$E[max\{\phi_{t1} + \rho_{ti}, \phi_{t2} + \rho_{t2}\}] - max\{E[\theta_{t1} + \rho_{t1}], E[\theta_{t2} + \rho_{t2}]\}$$

Claim 1: There exist an integer $T > 0$ such that $E[max\{\phi_{t1} + \rho_{ti}, \phi_{t2} + \rho_{t2}\}] - max\{E[\theta_{t1} + \rho_{t1}], E[\theta_{t2} + \rho_{t2}]\} < \sqrt{\frac{\tau^2}{\pi}}\ for\ all\ t > T$

Proof

Write

$$A_t = E[max\{\phi_{t1} + \rho_{ti}, \phi_{t2} + \rho_{t2}\}] - max\{E[\theta_{t1} + \rho_{t1}], E[\theta_{t2} + \rho_{t2}]\}$$

$$B_t = E[max\{\phi_{t1} + \rho_{ti}, \phi_{t2} + \rho_{t2}\}] - max\{E[\phi_{t1} + \rho_{t1}], E[\phi_{t2} + \rho_{t2}]\}$$

Notice that

$$E[max\{\phi_{t1} + \rho_{ti}, \phi_{t2} + \rho_{t2}\}] \tag{74}$$

$$= s_{t1}\Phi(\frac{s_{t1} - s_{t2}}{\sqrt{Q}}) + s_{t2}\Phi(\frac{s_{t2} - s_{t1}}{\sqrt{Q}}) + \sqrt{Q}\frac{1}{\sqrt{2\pi}}exp[-\frac{1}{2}\frac{(s_{t1} - s_{t2})^2}{Q}] \tag{75}$$

$$where \tag{76}$$

$$Q = \Omega_{t1} + \Omega_{t2} + 2\tau^2 \tag{77}$$

From (75) we can see that [25]

$$E[max\{\phi_{t1} + \rho_{ti}, \phi_{t2} + \rho_{t2}\}] < max\{s_{t1}, s_{t2}\} + \sqrt{\frac{max\{\Omega_{t1}, \Omega_{t2}\} + \tau^2}{\pi}} \tag{78}$$

Therefore, we have

$$B_t < \sqrt{\frac{max\{\Omega_{t1}, \Omega_{t2}\} + \tau^2}{\pi}} \tag{79}$$

$$\rightarrow \sqrt{\frac{\tau^2}{\pi}} \ as\ t \rightarrow \infty \tag{80}$$

To show (80), we use the fact that $max\{\Omega_{t1}, \Omega_{t2}\} \rightarrow 0$.

Next, we prove that

$$|A_t - B_t| \rightarrow 0\ as\ t \rightarrow \infty \tag{81}$$

To see this, notice that

$$|A_t - B_t| = max\{E[\theta_{t1} + \rho_{t1}], E[\theta_{t2} + \rho_{t2}]\} - max\{E[\phi_{t1} + \rho_{t1}], E[\phi_{t2} + \rho_{t2}]\} \tag{82}$$

Thus, it suffices to show that

$$E[\theta_{t1} + \rho_{t1}] - E[\phi_{t1} + \rho_{t1}] \tag{83}$$

$$= |E[\theta_{t1}] - E[\phi_{t1}]| \rightarrow 0 \quad as\ t \ \rightarrow \ \infty \tag{84}$$

To see this, notice that $E[\theta_{t1}]$ and $E[\phi_{t1}]$ have the following form.

$$E[\theta_{t1}] = E[\theta_1|C_t, D_t] \tag{85}$$

$$E[\phi_{t1}] = E[\theta_1|C_t] \tag{86}$$

[25] Notice that first two terms of (75) are a convex combination.

where $C_t$ is the correction of the sets of the following form $\{\theta_i + \rho_{si} + e^* = y_{s_i}\}$. $D_t$ is the collection of the sets of the form, $\{\theta_i + \rho_{si} + e^* \leq y_{s_i}\}$ Since $E[\phi_{t1}] \to \theta_1$, which is constant, it is also true that $E[\theta_{t1}] \to \theta_1$. Therefore, (84) holds.

We have shown that $A_t \to B_t$ and that $B_t \to \sqrt{\frac{\tau^2}{\pi}}$. Notice that $A_t = E[max\{\theta_{11} + \rho_{11}, \theta_{12} + \rho_{12}\}] - S_1$. Thus, the proposition is proved.

# References

[1] Aggarwal, Rajesh K and Samwick, Andrew A. "The Other Side of the Trade-Off: The impact of Risk on Executive Compensation." The Journal of Political Economy, Vol.107, Issue 1 (Feb., 1999), 65-105.

[2] Anderson, Mark C; Banker, Rajiv D and Ravindran, Sury. "Executive Compensation in the Information Technology Industry." Management Science, Vol.46, No.4, Information Technology Industry, (Apr.,2000), 530-547.

[3] Bac, Mehmet. "On-the-Job Specific Training and Efficient Screening." Journal of Labor Economics, Vol.18, No.4 (Oct.,2000),681-701.

[4] Bertrand, Marianne and Mullainathan, Sendhil. "Are Ceos Rewarded For Luck? The Ones Without Principals Are." The Quarterly Journal of Economics, MIT Press, vol. 116(3), (August, 2001) pages 901-932.

[5] Bizjak, John; Brickley, James; and Coles, Jeffery. "Stock-Based Incentive Compensation, Asymmetric Information and Investment Behavior." Manuscript. Salt Lake City: Univ. Utah, October 1991.

[6] Bollen, Kenneth; Guilkey, David K. and Mroz, Thomas A. "Binary Outcomes and Endogenous Explanatory Variables: Tests and Solutions with an Application to the Demand for Contraceptive Use in Tunisia." Demography, Vol.32, No.1 (Feb., 1995), 111-131

[7] Bryan, Stephen; Hwang, LeeSeok, and Lilien, Steven. "CEO stock-based Compensation: An Empirical Analysis of Incentive-intensity, Relative Mix, and Economic Determinants." The Journal of Business, Vol.73, No.4 (Oct., 2000), 661-693.

[8] Eaton, Jonathan and Rosen, Harvey S. "Agency, Delayed Compensation, and the Structure of Executive remuneration." The Journal of Finance, Vol.38, No.5(Dec.,1983),1489-1505.

[9] Fizel, John L and Louie,Kennith K.T. "CEO Retention, Firm Performance and Corporate Governance." Manegerial And Decision Economics, Vol.11, No.3 (Jul., 1990), 167-176.

[10] Garen,John E. "Executive Compensation and Principal-Agent Theory." The Journal of Political Economy. Vol. 102, No.6 (DEC., 1994), pp. 1175-1199

[11] Gaver, J.J. and Gaver, K.M. "Additional Evidence on the Association between the Investment Opportunity Set and Corporate Financing, Dividend, and Compensation Policies". Journal of Accounting and Economics, 16(January/April/July., 1993): 127-160.

[12] Gibbons, Robert and Murphy, Kevin J. "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." Journal of Political Economy, Vol. 100, No. 3 (Jun., 1992) , pp. 468-505

[13] Goolsbee, Austan."What Happens When You Tax the Rich? Evidence from Executive Compensation." The Journal of Political Economy, volume 108, Issue 2 (Apr., 2000), 352-378.

[14] Hall, Brian j and Liebman, Jeffrey. "Are CEOS Really paid Like Bureaucrats?" The Quarterly Journal of Economics, Vol. 113, No.3 (Aug., 1998), 653-691.

[15] Hall, Brian J and Murphy, Kevin J. "Optimal Exercise Prices for Executive stock Options." The American Economic Review, Vol.90, No.2, Paper and Proceedings of the One Hundred Twlfth Annual Meeting of the American Economic Association (May, 2000), 209-214.

[16] Hallock, Kevin F. "Reciprocally Interlocking Boards of Directors and Executive Compensation". J. OF FINANCIAL AND QUANTITATIVE ANALYSIS, (September, 1997)

[17] Heckman, James J and Singer, Burton. "Longitudinal analysis of labor market data." Cambridge University Press, New York, 1985.

[18] Holmstrom, Bengt and Milgrom, Paul. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design". Journal of Law, Economics, and Organization, Vol. 7, Special Issue: [Papers from the Conference on the New Science of Organization, January 1991] (1991) , pp. 24-52

[19] Inderst,Roman. "Screening in a Matching Market." The Review of Economics Studies, Vol.68, No.4 (Oct., 2001), 849-868.

[20] Jensen, Michael and Murphy, Kevin j. “Performance Pay and Top-Management Incentives.” The Journal of Political Economy, Volume 98, Issue 2 (Apr., 1990 A), 225-264.

[21] Jensen, Michael C. and Murphy, Kevin J. ”CEO Incentives: It’s Not How Much You Pay, But How”. Michael C. Jensen, FOUNDATIONS OF ORGANIZATIONAL STRATEGY, Harvard University Press, 1998; Harvard Business Review, No. 3, (May-June 1990 B) Available at SSRN:

[22] Lewis, Tracy R and Sappington, David E.M. “Optimal Contracting with Private Knowledge of Wealth and Ability.” The review of Economics studies. Vol. 68. No.1 (Jan., 2001), 21-44.

[23] Jovanovic, Boyan. “Job Matching and the Theory of Turnover.” The Journal of Political Economy, Vol.87,No.5, Part 1 (Oct., 1979),972-990.

[24] Kang, Jun-Koo and Shivdasani, Anil.“Firm Performance, Corporate Governance, and Top Executive Turnover in Japan.” Journal of Financial Economics, 38 (1995) 29-58.

[25] Lamber,Richard A; larcker, David F and Verrecchia, Robert E.“Portfolio Consideration in Valuing Executive Compensation.” Journal of Accounting Research, Vol.29,N0.1 (Spring, 1991),129-149.

[26] Ljungqvist, Lars and Sargent, Thomas J. “Recursive Macroeconomic Theory”. (1995) MIT Press.

[27] Murphy, Kevin J. “Corporate Performance and Managerial Remuneration. An Empirical Analysis.” Journal of Accounting and Economics 7 (1985) 11-42.

[28] Murphy, Kevin J. “Incentive, Learning, and Compensation: A Theoretical and Empirical Investigation of Managerial Labor Contracts.” The Rand Journal of Economics, Volume 17, Issue 1 (Spring, 1986), 59-76.

[29] O’Flaherty,Brendan and Siow,Aloysiu.“On the Job Screening ,Up or Out Rules, and Firm Growth.” The Canadian Journal of Economics, Vol.25, No.2 (May., 1992), 346-368.

[30] O'Flaherty,Brendan and Siow,Aloysiu.,"Up-or-out Rules in the Market for Layers." Journal of Labor Economics, 1995, val.13, no4.

[31] O'Flaherty, Brendan and Siow, Aloysius. "Up-or-Out Rules in the Market for Lawyers." Journal of Labor Economics, Vol. 13, No. 4 (Oct., 1995) , pp. 709-735

[32] Rust, John (1987). "Optimal replacement of GMC Bus Engine: An Empirical Model of Harold Zurcher." Econometrica, Vol.55, No.5 (Sep., 1987), 999-1033

[33] Richardson, Sylvia and Green, Peter. "On Bayesian Analysis of Mixtures with and Unknown Number of Components." Journal of Royal Statistical Society.Series B (Methodological) Vol1. No.4 (1997), pp. 731-792.

[34] Park, June S; Peters, Michael H and Tang, Kwei. "Optimal Inspection policy in Sequential Screening." Management Science, Vol.37, No.8 (Aug., 1991), 1058-1061.

[35] Prendergast, Canice and Stole, Lars. "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning." The Journal Of Political Economy, Vol.104, No.6 (Dec., 1996), 1105-1134.

[36] Prendergast, Canice. "The Tenuous trade-off between Risk and Incentives." The Journal of Political Economy, Oct 2002; 110,5, P1071-P1102.

[37] Prendergast, Canice (C). "What Trade-off of Risk and Incentives?" The American Economic Review, Vol. 90, No. 2, Papers and Proceedings of the One Hundred Twelfth Annual Meeting of the American Economic Association (May, 2000), 421-425.

[38] Slade,Margaret E. "Multitask Agency and Contract Choice: An Empirical Exploration." International Economics Review, Vol.37, No.2 (May,1996), 465-486.

[39] Stigliz, Joseph E. "The theory of "Screening," Education, and the Distribution of Income." The American Economics Review, Vol.65, No.3(Jun., 1975), 283-300.

[40] Tadelis, Steven. "The Market for Reputations as an Incentive Mechanism." The Journal of Political Economy; Aug 2002; 110, 4 pg. 854.

[41] Tirole, Jean. "Corporate Governance." Econometrica, Vol.69, N0.1 (Jan., 2001), 1-35

[42] Warner, J.B., Watts, R.L. and K.H.Wruck. "Stock Prices and Top Management Changes." Journal of Financial Economics 20 (1988), 461-492.

[43] Yabushita, Shiro. "Theory of Screening and the Behavior of the Firm: Comment." The American Economics Review, Vol.73, No.1,(Mar.,1983),242-245.

[44] Yermack, David. "Good Timing: CEO Stock Option Awards and Company News Announcements." The Journal of Finance, Vol.52, No.2 (Jun.,1997), 449-476.