Promotion, Turnover and Compensation in the Executive Market*

George-Levi Gayle, Limor Golan, Robert A. Miller
Tepper School of Business, Carnegie Mellon University

October 2008

Abstract

This paper is an empirical study of the market for managers, more specifically the effects of agency, human capital, and preferences on their promotion, tenure, turnover and compensation. From a large longitudinal data set compiled from observations on executives and their publicly listed firms, we construct a career hierarchy and report on its main features. Our summary results motivate a dynamic competitive equilibrium model, whose parameters we identify and estimate. Controlling for heterogeneity amongst firms, which differ by size and sector, and also managers, whose backgrounds vary by age, gender and education, our estimates are used to evaluate how important moral hazard and job experience are in jointly determining promotion rates, turnover and compensation.

1 Introduction

Chief executives are paid more than their subordinates, and internal promotions with the firm are positively correlated with wage growth.[1] Since high ranking executives are almost always drawn from the lower ranks, usually from within the firm, it is tempting to conclude that part of the reward from working hard in a low rank is the chance of promotion to earn rents. Theory provides several possible explanations, ranging from human capital acquired on lower level job, to superior ability being revealed with experience leading to wage dispersion, or as the prize in a tournament played by lower ranked executives to induce hard work.[2] The premise of all these explanations is the commonly held opinion that the CEO is better off than those he supervises. Yet several studies, conducted with data on executive compensation and returns from publicly traded firms, show quite conclusively that CEO compensation is more sensitive to the excess returns of firms than the compensation of lower ranked executives.[3] Thus at the upper levels of the career ladder,

*We thank the participants of the 2008 World Congress on National Accounts and Economic Performance Measures for Nations 2008, the seminar participants in Carnegie Mellon and Stony Brook for comments and suggestions. This research is supported by the Center for Organizational Learning, Innovation and Performance in Carnegie Mellon University and National Science Foundation Grant Award SES0721098. Preliminary and Incomplete.

[1]See Lazear (1992), Baker, Gibbs and Holmstrom (1994a), McCue (1996)

[2]See Prendergast (1999), Gibbons and Waldman (1999) and Neal and Rosen (2000) for surveys.

[3]See Margiotta and Miller (2000) and Gayle and Miller (2008a, 2008b).

differently ranked jobs do not have the same characteristics. Whether one job is more desirable than another depends on the probability distribution of financial compensation that generates his income, as well as its nonpecuniary costs and benefits.

To the best of our knowledge, no one has attempted to quantify how much a CEO receives as a rent from human capital in management and leadership, and how much he is compensated for receiving a more volatile income. A small but growing literature on the structural estimation of moral hazard models investigates the empirical relationship between the principal's return and the agent's compensation, in order to quantify how incentives are used for inducing agents to work in the interests of their principals and truthfully revealing their hidden information.[4] These studies find that estimates of the higher risk premium necessary to compensate a CEO for a more uncertain income relative to the second in command are of the same order of magnitude as differences in expected compensation. Such findings do not resonate with common opinion, because they imply the CEO receives very little pecuniary rent from his promotion to that position. Published work does not, however, integrate human capital and its behavioral consequences into an optimal contracting framework, confounding any attempt to gauge the degree of on-the-job training provided at lower ranks relative to the nonpecuniary value of holding a job at any given rank. More generally, the empirical importance of human capital in the executive labor market, and the role of promotions in this process, is unclear.[5]

This paper is an empirical study of the effects of incentives, human capital, and preferences of managers, with goal of explaining the differences in the promotion, tenure, job turnover and compensation structure across managers using a dynamic competitive equilibrium model. Our data contain background information on executives, including age, gender, education, executive experience and the types of firms they work for, plus detailed information on their compensation and the financial returns of their firms and their rank within a career hierarchy. We identify and estimate a dynamic equilibrium model to analyze and disentangle the effects of competition in the market for managers using data on internal promotions, job turnover and the compensation of executives. estimate. Controlling for heterogeneity amongst firms and managers, our estimates are used to evaluate how important moral hazard and job experience are in jointly determining promotion rates, turnover and compensation.

The model is set up in the next two sections. Executives choose job, firm and effort level every period. They have preferences over jobs, particularly, effort is costly. These taste parameters vary across jobs and firms. In addition, every period managers privately observe a firm-job specific taste shock. The effort level is private information as well. While working they accumulate firm-specific and general human capital. We assume human capital accumulation on a job is greater when the manager exerts effort. The rate of human capital accumulation varies across jobs and firm as well, therefore, working in some firms and jobs may increase the manager's stock of human capital. Firms offer contracts which provide incentives for managers to exert effort. Because exerting effort increases the manager's stock of human capital, future promotion prospects provide incentives.[6] Thus,

[4] Ferrall and Shearer (1999), Margiotta and Miller (2000), Dubois and Vukina (2005), Bajary and Khwaja (2006), Duflo, Hanna, and Ryan (2007), D'Haultfoeviller and Fevrier (2007), Einav, Finkelstein and Schrimpf (2007), Nekipelov (2007), Gayle and Miller (2008a,b,c).

[5] Frydman (2005) finds evidence on the increase importance of general skills in executive compensation.

[6] Gibbons and Murphy (1992) develop and empirically test a model of optimal contracts in the presence

variation in compensation across firms and jobs partially reflect the different opportunities to accumulate human capital and different promotion prospects. In addition, managers' age and rank imply differences in career concerns affecting the optimal compensation schemes. The markets for executives is competitive. Managers have different stocks of human capital and compensation adjusts to clear the market for each skill set.

Identification of the parameters of the model is analyzed in Section 4, while our data is described in Section 5, where we define the job hierarchy and wage compensation. Our measure of compensation is comprehensive, and includes salary and bonus, stock and option grants, retirement benefits, as well as income directly attributable to holding securities in the firm in lieu of a widely diversified portfolio. The compensation data is augmented with data on the titles of the executives, along with their professional and demographic background compiled from the Marquis "Who's Who" . Compensation of the executives are sensitive to fluctuations in the abnormal returns. In fact, the firm's excess return (over and above the market's return) is the most important determinant of managerial compensation, suggesting the importance of incentives and moral hazard. We find that in fact the higher the executive's rank in the firm, the more sensitive his compensation to the abnormal return. We also find that firm turnover is positively correlated with promotions and higher compensation.

Estimation is discussed in Section 6, while some preliminary estimates from the structural estimation are reported in the final section. We used four metrics to assess how much agency problems in executive markets are mitigated by their career concerns. Two of these measure the impact of an executive shirking rather than working, while the other two focus on the cost of eliminating the moral hazard problem. We find that firms are prepared to pay hardly anything to eliminate the moral hazard problem at the lower ranks, but that at the upper levels, the risk premium paid to executives for accepting an uncertain income stream that depends on the firm's abnormal returns, are considerably greater. Career concerns greatly ameliorate the moral hazard problem for lower level executives, but their importance declines monotonically with promotion through the ranks. Overall our empirical findings, based on a large sample of executives employed by a broad cross section of publicly traded firms, demonstrate that the design of the hierarchy and the promotion process are important tools, used in conjunction with compensation schemes, for disciplining employees and aligning their interests to the goals of the organization.

2 The Model

Our model analyzes promotion, turnover and executive compensation, where shareholders are subject to moral hazard from choices made by their managers, and are also worse informed than their managers about the value of their job matches. Promotions and career prospects vary across firms, job rank and work experience. Each period firms advertise for executives by offering them contracts that depend on the position within the firm, the applicant's known characteristics, and the firm's random return next period. In our framework, the demand for executives is exogenously fixed by success rate probabilities firms must achieve through their hiring recruiting policies, an assumption we relax in our companion paper on organization capital. Executives privately observe their individual

of career concerns in the market for CEOs.

taste shocks for the period, select a contract or quit the workforce, choose a work routine that is not observed by the firm's directors, as well as their consumption for the period. They accumulate general and firm-specific human capital while working and are compensated at the beginning of next period after the firm's returns are realized. The managers are risk averse and sequentially maximize their expected lifetime utilities, Shareholders minimize expected summed and discounted compensation costs subject to achieving the exogenous success rate and inducing management to pursue the interests of the firm (work diligently) as opposed to some other goal (shirk).

2.1 Human Capital

Executives are infinitely lived and work at most for $T \leq \infty$ periods. There are J identical firms in the market indexed by $j \in \{0, ..., J\}$, with $j = 0$ representing retirement, which we assume is an absorbing state. There are K different types of positions within each firm, indexed by $k \in \{1, ..., K\}$. We let $I_{jkt} \in \{0, 1\}$ indicate the manger's job, his rank k at firm j in time period $t \in \{0, 1, ...\}$, where $I_{0kt} = 1$ means the executive is retired. The choices are mutually exclusive so for each time period t before retirement:

$$\sum_{j=1}^{J} \sum_{k=1}^{K} I_{jkt} = 1$$

There are two activities in the firm, called working and shirking, denoted by $l_t \in \{0, 1\}$,where $l_t \equiv 0$ means the manager shirks in period t and $l_t \equiv 1$ means the manager works.

We investigate two scenarios. The first is that the manager accumulates general and firm-specific capital at a rate that depends on the type of firm and the manager's effort level, in particular that human capital is only accumulated if the manager works diligently. Thus the specific human capital of a manager entering period t in the j^{th} firm is:

$$h_{jt} = \sum_{j=1}^{J} \sum_{k=1}^{K} I_{jkt} l_{t-s}.$$

Since l_t is private information, h_{jt} is also the private information of the manager. The second scenario we consider is that the specific human capital of a manager accumulates independently of work effort and only depends on his employment history, which is public knowledge:

$$h_{jt} = \sum_{j=1}^{J} \sum_{k=1}^{K} I_{jkt}.$$

In both cases his general human capital is the sum of experience in all firms:

$$h_{0t} = \sum_{j=1}^{J} h_{jt}$$

Thus $h_{0t} = (h_{0t}, \ldots, h_{Jt})$ is observed by everyone in one scenario but is private information in the other.

2.2 Preferences of Managers

The preferences of managers are characterized by the discounted sum of a time additively separable constant absolute risk aversion utility function, which is multiplicative in consumption and nonpecuniary factors. Human capital affects both the productivity of the firm, as discussed below, and also enter preferences directly, through the ease with which tasks are accomplished. Thus the preference parameters of a manager depend on his employer and rank d_{mkt}, his experience h_t, his activity choice l_t, and other demographic characteristics (such as age, gender and education), which are all consolidated within a vector denoted by z_t. Our empirical framework assumes that z_{t+1} is fully determined by z_t and the choices (j, k, l) made at t, an assumption we make in this section essentially for convenience. We represent the preference parameters by the mapping $\alpha_{jkl}(z_t)$. We normalize the parameters associated with retirement. That is if $j = 0$, then $\alpha_{jmk} = \alpha_0 = 1$ for all j and k; and $\varepsilon_{0kt} = \varepsilon_{0t}$ for all k. For any other job choice $j \neq 0$ we assume that the disutility comes from work, or $\alpha_{1jk}(z_t) > \alpha_{0jk}(z_t)$ for all z_t. An individual taste shock indexed by time, firm, and position, denoted by ε_{jkt}, also affects current utility. To summarize life-time utility is

$$-\sum_{t=1}^{\tau}\sum_{j=1}^{J}\sum_{k=1}^{K}\beta^t d_{jkt}\left[\alpha_{0jkt}(1-l_t)+\alpha_{1jkt}l_t\right]\exp\left(-\rho c_t-\varepsilon_{jkt}\right)-\sum_{t=\tau+1}^{\infty}\beta^t\exp\left(-\rho c_t-\varepsilon_{0t}\right) \tag{1}$$

where β is the subjective discount factor and ρ is the constant absolute risk aversion parameter and $\tau \leq T$ is the retirement date, defined as:

$$\tau = \sum\nolimits_{t=1}^{\infty}\prod\nolimits_{s=1}^{t}\left(\sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K} d_{jkt}\right)$$

We assume there exists a complete set of markets for all publicly disclosed events relating to commodities, with price measure Λ_t defined on F_t and derivative λ_t. This implies that consumption by the manager is limited by a lifetime budget constraint, which reflects the opportunities she faces as a trader and the expectations she has about her compensation. The lifetime wealth constraint is endogenously determined by the manager's work activities. By assuming markets exist for consumption contingent on any public event, we effectively attribute all deviations from the law of one price to the particular market imperfections under consideration. Let e_t denote the endowment at date t. We also measure $w_{mk,t+1}$, the manager's compensation for employment at rank k for firm n in period t, in units of current consumption. To indicate the dependence of the consumption possibility set on the set of contingent plans determining labor supply and effort, we define $E_0[\bullet \,|l]$ as the expectations operator conditional on work and effort level choices throughout the manager's working life. The budget constraint can then be expressed as

$$E_t(\lambda_{t+1}e_{t+1}) + \lambda_t c_t \leq \lambda_t e_t + E_t(\lambda_{t+1}w_{jkt+1}|l_t, I_{jkt}) \tag{2}$$

2.3 Firm Technology

The actions of a firm's managers affect the distribution of their firm's abnormal returns, defined as the residual component of returns not priced by aggregate factors. Thus the abnormal returns in period t to the j^{th} firm, denoted by x_{jt}, is a random variable that depends on the managers' efforts in the previous period. Conditional on the effort vector of the executive branch $\{l_{jkt}\}_{k=1}^{K}$, the characteristics of management (including their collective executive experience,) and the firm (such as size, capital structure, human capital of executives and industrial mix), x_{jt} is independently distributed across both firms and periods. Let z_{jt} denote the set of characteristics defining the firm and its management in period t, including the composition of the management $(I_{j1t}, ...I_{jKt})$ and their experience $(h_{j1t}, ...h_{jKt})$. Also let $f(x|l_{j1t}, ...l_{jKt}, z_{jt})$ denote the probability density function for x_t, conditional on effort levels $(l_{j1t}, ...l_{jKt})$. For convenience we adopt the notation:

$$f(x|l_{j1t}, ...l_{jKt}, z_{jt}) = \begin{cases} f_j\left(x|z_{jt}\right) \text{ if } \sum_{k=1}^{K} I_{jkt}l_{jkt} = \sum_{k=1}^{K} I_{jkt} \\ f_{jk}\left(x|z_{jt}\right) \text{ if } \sum_{k=1}^{K} I_{jk't}l_{jkt} = \sum_{k=1}^{K} I_{jk't} - 1 \text{ and } l_{jkt} = 0 \end{cases}$$

A conflict of interest between owners and managers arises because whereas , managers prefer to shirk rather than work, so $\alpha_{1jk}\left(z_t\right) > \alpha_{0jk}\left(z_t\right)$, but the expected return to firms is increasing in the number of its executives who work, meaning that for all (j, k):

$$\int xf_j\left(x|z_{jt}\right)dx > \int xf_{jk}\left(x|z_{jt}\right)dx$$

The objective function of firm $j \in \{1, \ldots, J\}$ is the expected value of discounted abnormal returns net of managerial compensation, given by the expression:

$$E_0\left[\sum\nolimits_{t=1}^{\infty} \lambda_t \left(x_{jt}\vartheta_{j,t-1} - \sum\nolimits_{k=1}^{K} I_{jkt}w_{jkt}\right)\right]$$

where ϑ_{jt} denotes the value of the j^{th} firm at time t,

3 Optimization and Competitive Selection

At the beginning of every period, x_{jt}, the returns from the working in the previous period are revealed to everyone, human capital and other state variables are updated to z_t and the managers are compensated w_{jkt} according to the previous period's schedule. Executives privately observe realizations of preference shocks, ε_{jkt} being privately revealed to each manager and choose consumption. Also the demand for executives P_{jkt} and their effort L_{jkt} is assigned to each firm. Firms advertise one-period contract offers to the executive market, in the form of a compensation schedule w_{jkt+1}, executives accept their most attractive offer, I_{jkt}, or quits (whichever leads higher expected lifetime utility), and selects consumption c_t. Finally each executive then chooses an effort level l_t which he privately observes. The realization of the returns outcome $x_{j,t+1}$ is revealed at the beginning of the next period, and the process repeats itself until period $T+1$ when all remaining managers retire.

Shareholders (firms) and managers solve three optimizing problems. Managers sequentially maximize expected lifetime utility with respect to consumption given the current choice of work compensation schedules and the future job transition probabilities, a single agent dynamic choice problem. Given the supply of managers as a function of future compensation schedules and participation probabilities, firms minimize their expected compensation from paying executive management subject to two constraints, achieving an exogenously determined target probability for successfully hiring at each rank, and conditional on an acceptance at a given rank inducing the designated effort level. From the conditional valuation function that solves the first problem, and given the contracts offered by employers, executives optimally choose job, firm and effort level for the executive, a discrete choice problem. The solution to these three optimization problems determine the stochastic process governing the supply of executives. We define an equilibrium selection for this model by a compensation schedule for each executive type such that the endogenously determined supply of executives follows the same stochastic process as the exogenously given demand for executives. We prove that an equilibrium selection exists in our model.

3.1 Consumption and Saving

We first solve for the stochastic sequence of consumption savings choices as a function of the compensation plans offered by different firms and the career choices the manager makes in response. He optimally chooses his consumption stream when his income accrues from a sequence of lotteries with prizes of $w_{jk,t+1}(z,x)$, and also nonpecuniary (scaled utility) benefits from participating of $\alpha_{ljkt}\left[\exp\left(-\varepsilon_{mkt}^{*}\right)|j,k,l,z\right]$, where ε_{jkt}^{*} is the value of the period t disturbance when lottery (j,k,l) is selected. Which lottery he plays in period t is determined by another lottery,defined by its probabilities $p_{jklt}(z)$.

Let $V_t(z)$ denote the indirect utility function found by substituting the (stochastic) consumption sequence c_t^o into the utility that sequentially maximizes the expected utility function (1) subject to the budget constraints (2) by optimally choosing when presented with compensation schedules $w_{jk,t+1}(z,x)$ and taking the probabilistic choice rule defined by generically by the mapping $p_{jklt}(z)$, which denotes the probability that the manager selects (j,k) in period t given characteristics z and integrating over $\varepsilon_{jkt} \equiv (\varepsilon_{0t}, \varepsilon_{11t}, \ldots, \varepsilon_{JKt})$:

$$V(z) = -\sum_{t=1}^{\infty}\sum_{j=1}^{J}\sum_{k=0}^{K}\beta^t p_{jklt}(z)\left[\alpha_{0jkt}(1-l_t)+\alpha_{1jkt}l_t\right]\exp\left(-\rho c_t^o\right)E\left[\exp\left(-\varepsilon_{mkt}^{*}\right)|j,k,l,z\right]$$

Similarly we define $V_{jklt}(z)$ as the conditional valuation as the optimized value of expected lifetime utility starting at t given discrete choice (j,k,l) and state z and idiosyncratic shock ε_{jkt}. Appealing to Bellman's principle:

$$V_{jklt}(z_t,\varepsilon_{jkt}) = \left[\alpha_{0jkt}(1-l_t)+\alpha_{1jkt}l_t\right]\exp\left(-\rho c_t^o\right)\exp\left(-\varepsilon_{jkt}\right) - \beta V\left(z_{t+1}^{(j,k,l)}\right) \quad (3)$$

and from their respective definitions

$$V(z_t) = -\sum_{j=1}^{J}\sum_{k=0}^{K}\sum_{l=0}^{1} p_{jklt}(z_t)\,E\left[V_{jklt}(z_t,\varepsilon_{jkt})\,|j,k,l,z\right] \quad (4)$$

Given the compensation schedules and their career choices, which we In order to derive the solution to the optimal consumption decision we start out with the conditional valuation function for working one period at time t and then retiring and dying at $n+1$, where the nonpecuniary parts of utility from working are ε_{mkt} (is the expected conditional valuation of this unobserved nonpecuniary benefit, and α_k treated as a parameter, where

α_0 is also estimated as a parameter. For notational ease denote by $z_{mt} = (h_{mt}, h_t)$, assume that z_{mt} has finite support $\mathbb{Z}$, let b_t denoted the period t price of a infinitely lived bond, and a_t the price of a security that pays off the (random) dividend $(\ln\lambda_s - s\ln\beta - \ln\lambda_t)$ is period s. Let:

$$\upsilon_{jk,t+1}(z,x) \equiv \exp\left(-\rho w_{jk,t+1}(z,x)/b_{t+1}\right)$$

denote the risk adjusted utility weight for receiving compensation $w_{jk,t+1}(z,x)$ at the beginning of period $t+1$ for working (j,k) in period t when the manager's characteristics are z, the bond price in period $t+1$ is b_{t+1}, and the abnormal return to the firm is x. For all $s \in \{1,\ldots,T\}$ we set $A_0(z_t) \equiv 1$ for all (j,k) and recursively define $A_s(z_t)$ as:

$$A_s(z_t) = \sum\nolimits_{(j,k,l)} p_{jkl}(z_t)\,\alpha_{jklt}^{1/b_t}\left[A_{s-1}\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} E[\exp\left(-\varepsilon_{jkt}^{*}/b_t\right)|z_t]\left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}} \tag{5}$$

where ε_{jkt}^{*} is the value of the period t disturbance when (j,k) is selected, b_t is the current price of a perpetual bond at t, and $z_{t+1}^{(j,k,l)}$ is the value of the state variables in period $t+1$ induced by the choices (j,k,l) in period t. As a preliminary result we provide the following extension to the infinite horizon case $T=\infty$.

Lemma 1 *Suppose (z_t,b_t) is a stationary Markov process and, conditional on (z_t,b_t), ε_{jkt}^{*} is identically and independently distributed. Then there exits an $A(z_t)$ which uniquely satisfies the fixed point:*

$$A(z_t) = \sum\nolimits_{(j,k,l)} p_{jkl}(z_t)\,\alpha_{jklt}^{1/b_t}\left[A\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} E[\exp\left(-\varepsilon_{jkt}^{*}/b_t\right)|z_t]\left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}}$$

and $A_s(z_t)$ uniformly converges to an $A(z_t)$as $s\to\infty$.

The value function and the optimal consumption rule for the manager is formed from $A_s(z_t)$ as Lemma 2 shows.

Lemma 2 *The conditional valuation functions take the form:*

$$V_{jklt}(z_t) = -b_t\alpha_{jklt}^{1/b_t}\left[A_{s-1}\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}}\exp\left(-\varepsilon_{jkt}^{*}/b_t\right)\left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}}\exp\left(-\frac{a_t+\rho e_t}{b_t}\right)$$

Lower values of $A_{s-1}(z_t)$ in the conditional valuation functions are associated with higher investment value. From Lemma 2 and the equations for $A_s(z_t)$ and $V(z)$, given by (4) and (5) we obtain the following formula for the unconditional valuation function:

$$V(z) = -A_s(z_t)\,b_t\exp\left(-\frac{a_t+\rho e_t}{b_t}\right) \tag{6}$$

Thus $A_s(z_t)$ can be interpreted as a standardized unconditional valuation function in which the endowment is $e_t = a_t/\rho$.

Differentiating the right side of (3) after substituting in (6) we obtain optimal consumption in terms of $A_s(z_t)$, namely:

$$\begin{aligned} c_t^o &= \frac{e_t}{b_t} + \frac{t}{\rho}\log\beta + \frac{a_t}{\rho b_t} \\ &\quad -\rho^{-1}\left\{\log\left(E_t\left[\upsilon_{jk,t+1}\right]\right) + \left(1-b_t^{-1}\right)\varepsilon_{jkt}^{*}\log\alpha_{jklt} - \log\left[A_{s-1}\left(z_{t+1}^{(j,k,l)}\right)\right]\right\} \end{aligned}$$

The first three terms of the formula for optimal consumption are familiar, spending the interest on the endowment, discounting consumption over time due to impatience and adjusting for aggregate risk. The next term (inside the parenthesis) reflects the certainty equivalent of the compensation, while the two final terms stem from the effects of non-pecuniary features of the job on the marginal utility of consumption, and the investment value of the job through the shifts in the compensation and perks of future jobs.

3.2 Promotion and Turnover

Job choices are found by choosing (d_t, l_t) to maximize $V_{jklt}(z_t)$ or minimizing the logarithm of its absolute value. Then dividing through by b_t and adding the constant $(a_t + \rho e_t)$ we are left with the problem of finding:

$$\underset{(d_t,l_t)}{\arg\min} \sum\nolimits_{(j,k,l)} d_{jkt} \left\{ \begin{array}{c} \log \alpha_{jklt} + (b_t - 1) \log \left[A_{s-1} \left(z_{t+1}^{(j,k,l)} \right) \right] \\ + (b_t - 1) \log \{ E[\upsilon_{jk,t+1} | z_t, l_t] \} - \varepsilon_{jkt} \end{array} \right\}$$

For example if ε_{jkt} is identically and independently distributed Type I Extreme Value, then conditional expectation of the disturbance, the recursion for $A_s(z_t)$ and the choice probabilities simplify it the formulas given in following lemma.

Lemma 3 *If ε_{jkt} is independently and identically distributed as extreme value Type I with location and scale parameters* $(0, 1)$, *then*

$$E\left[\exp\left(\varepsilon_{jkt}^{*}/b_t\right)\right] = p_{jklt}^{-1/b_t} \Gamma\left[(b_t - 1)/b_t\right]$$

$$A_s(z_t) = \sum\nolimits_{(j,k,l)} \alpha_{jklt}^{\frac{1}{b_t}} \left[p_{jkl}(z_t) \left[A_{s-1}\left(z_{t+1}^{(j,k,l)} \right) \right] \{ E[\upsilon_{jk,t+1} | z_t, l_t] \} \right]^{1-\frac{1}{b_t}} \Gamma\left[(b_t - 1)/b_t\right]$$

and:

$$p_{jklt}(z_t) = \frac{\alpha_{jklt} \left[A_{s-1}\left(z_{t+1}^{(j,k,l_t)} \right) \right]^{(b_t-1)} \{ E[\upsilon_{jk,t+1} | z_t, l_t] \}^{(b_t-1)}}{1 + \sum_{j'=1}^{J} \sum_{k'=1}^{K} \alpha_{j'k'l't} \left[A_{s-1}\left(z_{t+1}^{(j',k',l')} \right) \right]^{(b_t-1)} \left\{ E[\upsilon_{j'k',t+1} | z_t, l'] \right\}^{(b_t-1)}} \quad (7)$$

3.3 Compensation

In our framework we assume firms cannot commit to long term multiperiod contracts with their executive staff. If human capital depends on employment history but not on effort, then the optimal long term contract decentralizes to a sequence of short term contracts, obviating the need to consider anything but one period contracts.[7] However if human capital is a function of unobserved effort, then there are benefits to shareholders from committing. In that scenario, the optimal long term compensation contract takes into account the signals a firm receives from abnormal returns about previous firm specific unobserved investments in human capital made by its workers. Absent a commitment device, firms engage in sequentially optimal short term one period contracts of the type we now proceed to analyze.

[7] The proof of this statement follows arguments developed in Fudenberg, Holmstrom and Milgrom (1990).

Each firm in our model sets its compensation schedule to minimize the expected costs of its executive management subject to a recruiting and retention success rate, which we define in terms of the proportion of the time its positions are filled with managers who are induced to work diligently. Shareholders of the j^{th} firm seek to attract at minimal expected cost a manager having characteristics z for its k^{th} ranked position, where the contract is lucrative enough to fill the position with probability $P_{jkt}(z)$ and, conditional on filling the position, the contract aligns the manager's incentives with shareholders' interests. In our first scenario, where human capital is unobserved by shareholders because it depends on past effort, we assume that shareholders are certain that managers have accumulated the human capital prescribed by the equilibrium path. In this respect shareholders act as if human capital is observed. The solution to this problem is found by minimizing expected compensation subject to an incentive contract, which induces a manager accepting the job offer to work diligently, the value and a participation constraint, which successfully attracts the executive with that probability.

The incentive compatibility condition is satisfied if and only the value of working diligently, $V_{jk1t}(z_t)$, exceeds $V_{jk0t}(z_t)$ the value of shirking, or simplifying:

$$E[\upsilon_{jk,t+1}(x)\, g(x, z_t)\, |z_t] \leq \left[\frac{\alpha_{1jkt}}{\alpha_{0jk0t}}\right]^{1/(b_t-1)} \frac{A_{s-1}\left(z_{t+1}^{(j,k,1)}\right)}{A_{s-1}\left(z_{t+1}^{(j,k,0)}\right)} E[\upsilon_{jk,t+1}(x)\, |z_t]$$

where the expectation is taken over the density associated with working hard and:

$$g_{jk}(x, z) \equiv f_{jk}(x|z) / f_j(x|z)$$

is defined as the likelihood ratio of density function for abnormal returns to the j^{th} firm when everyone except the k^{th} manager works diligently compared to density when everyone works diligently.

When human capital is public information $A_{s-1}\left(z_{t+1}^{(j,k,1)}\right) = A_{s-1}\left(z_{t+1}^{(j,k,0)}\right)$ because human capital accumulation does not depend on effort, and in this case the incentive compatibility constraint reduces to the constraint that applies when there are no human capital in the model at all:

$$E[\upsilon_{jk,t+1}(x)\, g(x, z_t)\, |z_t] \leq \left[\frac{\alpha_{1jkt}}{\alpha_{0jk0t}}\right]^{1/(b_t-1)} E[\upsilon_{jk,t+1}(x)\, |z_t]$$

When human capital accrues only if the manager is diligent, instead of weighting the utility of returns by α_{jk1t} alone, a parameter reflecting the distaste for work, or the utility of nonpecuniary benefits, the value of the job as a source of human capital $A_{s-1}\left(z_{t+1}^{(j,k,1)}\right)$, conditional on effort, also enters as a scaling factor through the difference between $z_{t+1}^{(j,k,1)}$ and $z_{t+1}^{(j,k,0)}$. Human capital is associated with higher compensation and better work conditions in the future, implying:

$$A_{s-1}\left(z_{t+1}^{(j,k,1)}\right) < A_{s-1}\left(z_{t+1}^{(j,k,0)}\right)$$

The more upper level job opportunities for increased pay and nicer working conditions draw upon human capital gained from hard work in lower level jobs, the more the ratio

$A_{s-1}\left(z_{t+1}^{(j,k,1)}\right)/A_{s-1}\left(z_{t+1}^{(j,k,0)}\right)$ offsets $\alpha_{1jkt}/\alpha_{0jk0t}$, which measures how much shirking is preferred to working when compensation is independent of effort. If career concerns are so important in motivating executives that:

$$A_{s-1}\left(z_{t+1}^{(j,k,1)}\right)/A_{s-1}\left(z_{t+1}^{(j,k,0)}\right) < \left(\alpha_{0jkt}/\alpha_{1jk0t}\right)^{1/(b_t-1)}$$

then the incentive compatibility constraint is nonbinding, obviating the reason for tying remuneration to the abnormal returns of the firm.

The participation constraint requires the firm to offer a compensation package that is successful with probability of at least $P_{jkt}(z)$. Since the probability of accepting a package is $p_{jkt}(z)$ we require that $p_{jkt}(z) \geq P_{jkt}(z)$. Substituting the equation for $p_{jkt}(z)$ into this inequality then yields:

$$\frac{\alpha_{jklt}\left[A_{s-1}\left(z_{t+1}^{(j,k,l_t)}\right)\right]^{(b_t-1)}\left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{(b_t-1)}}{1+\sum_{j'=1}^{J}\sum_{k'=1}^{K}\alpha_{j'k'l't}\left[A_{s-1}\left(z_{t+1}^{(j',k',l')}\right)\right]^{(b_t-1)}\left\{E[\upsilon_{j'k',t+1}|z_t,l']\right\}^{(b_t-1)}} \geq P_{jkt}(z)$$

Given the compensation packages offered in the other positions (j',k'), it is convenient to express the participation constraint as:

$$U_{jk}(z) \leq \alpha_{ljkt}E[\upsilon_{jk,t+1}|z_t,l]\left[A_{s-1}\left(z_{t+1}^{(j,k,l_t)}\right)\right] \tag{8}$$

where

$$U_{jk}(z)^{(b_t-1)} \equiv \left[\frac{P_{jkt}(z)}{1-P_{jkt}(z)}\right]\left\{1+\sum_{\substack{j'=1\\ j'\neq j}}^{J}\sum_{\substack{k'=1,\\ k'\neq k}}^{K}\alpha_{j'k'l't}\left[A_{s-1}\left(z_{t+1}^{(j',k',l')}\right)\right]^{(b_t-1)}\left\{E[\upsilon_{j'k',t+1}|z_t,l']\right\}^{(b_t-1)}\right\}$$

The right hand side of (8) is expected lifetime utility from the nonpecuniary features from accepting the (j,k) contract, aside from its individual taste shock component, the certainty equivalent of its compensation, and the standardized value function from continuing optimally next period with human capital $z_{t+1}^{(j,k,l_t)}$. The left side reflects the value of taking one of the outside options available to the manager and its implied continuation values.

Both constraints are linear in $\upsilon_{jk,t+1}$, and the objective function, the expected wage bill $E_t(w_{mkt+1})$ can be expressed as a concave function of $\upsilon_{jk,t+1}$, namely $E_t(\ln\upsilon_{mkt+1})$. Consequently the Kuhn Tucker Theorem applies, so the Lagrangian for the problem in which the firm elicits diligent work for the k^{th} rank can be written as

$$\begin{aligned}&E_t[\ln(\upsilon_{j,k,t+1})|z_t]+\eta_1\left[U_{jk}(z)-\alpha_{jk1t}\left[A_{s-1}\left(z_{t+1}^{(j,k,1)}\right)\right]E[\upsilon_{jk,t+1}|z_t,l]\right]\\ &+\eta_2E_t\left\{\upsilon_{j,k,t+1}\left[g(x,z_t)-\left(\frac{\alpha_{jk1t}}{\alpha_{jk0t}}\right)^{1/(b_t-1)}A_{s-1}\left(z_{t+1}^{(j,k,1)}\right)/A_{s-1}\left(z_{t+1}^{(j,k,0)}\right)\right]|z_t\right\}\end{aligned} \tag{9}$$

This formulation mimics the static model where there is no human capital, meaning $A_{s-1}(z_{j,k,l,t+1}) = 1$, and only one job choice, or $U_{jk}(z) = 1$. Consequently the short term contract cost minimization problem for the dynamic model where there are multiple jobs is solved by making the appropriate substitutions in the static model for one job type and appealing to Proposition 4 in Margiotta and Miller (200 page 682).

Lemma 4 *In the equilibrium where firms elicit high effort for all managers in the hierarchy, the optimal contract is*

$$w_{j,k,t+1}(x,z) = \frac{b_{t+1}}{\rho} \left\{ \begin{array}{l} (b_t - 1)^{-1} \left[\log U_{jk}(z) - \log(\alpha_{jk1t}) \right] \\ + \log \left[1 - \eta g_{jk}(x, z_t) + \eta \left(\frac{\alpha_{1jkt}}{\alpha_{0jkt}} \right)^{1/(b_t - 1)} A_{s-1}\left(z_{t+1}^{(j,k,1)} \right) / A_{s-1}\left(z_{t+1}^{(j,k,0)} \right) \right] \end{array} \right\} \tag{10}$$

where η *is the unique positive root to*

$$\int \left[\frac{f_{jk}(x, z_t)}{\eta(\alpha_{1jk}/\alpha_{0jk})^{1/(b_t-1)} A_{s-1}\left(z_{t+1}^{(j,k,1)} \right) / A_{s-1}\left(z_{t+1}^{(j,k,0)} \right) - \eta g_{jk}(x, z_t)} \right] dx = 1 \tag{11}$$

Equation 10 implies an expected utility level $U_{jk}(z)$ required to attract a manager with characteristics z to a job k in firm j with probability $P_{jk}(z)$. The expected utility increases in the outside options of the manager. As discussed above, the expected cost to the firm depends on the promotion probabilities and the continuation value attached to the job relative to the continuation values of working in other jobs. Firm-specific human capital accumulated on the job, should increase the value of working in the firm relative to the outside options and therefore, reduces the expected cost of the contract to the firm. General human capital increases the outside option.

When human capital does not depend on effort, the $A_{s-1}\left(z_{t+1}^{(j,k,1)} \right)$ and $A_{s-1}\left(z_{t+1}^{(j,k,0)} \right)$ terms cancel each other in (10) and (11). Consequently the only difference in the optimal contract distinguishing a model with human capital that depends on past job choices, from a model of pure moral hazard without any human capital, is the fixed component of compensation $U_{jk}(z)$. Regardless of whether human capital is observed or not, jobs associated with promotion prospects to higher paid jobs command an offsetting negative compensating differential, reflected in lower values of $U_{jk}(z)$. But when human capital is not observed, the prospect of promotion also ameliorates incentive problems, and would predict that lower level jobs on fast promotion tracks require less incentive pay.

3.4 Competitive Selection

To close the model we impose a competitive selection condition, which requires the supply of executives to follow the same stochastic process as their demand. Matching up conditional distributions is weaker than requiring the market for each executive to clear in every period, although the two objects converge to each other in large markets, by the law of large numbers.

In terms of our notation we define a competitive selection for a probability choice vector of demands by firms is P as a $w(z)$ solving every firm's cost minimization problems, such that the supply of executive at $w(z)$ is P when firms and managers have rational expectations. More formally, let $\Omega(P)$ denote the vector of compensation contracts that arise at time period t from the cost minimization problem, as a mapping of P. Thus $w(z) \equiv \Omega(P)$. Also let $\Psi(w(z))$ denote the choice operator that induces executive supply, which we denote by p. Thus $p \equiv \Psi[w(z)]$. By the definition of Ω and Ψ it immediately follows that $p = \Psi[\Omega(P)]$, and the executives make a competitive selection if $p = P$

for the compensation schedule $\Omega(P)$. Appealing to a fixed point theorem, the following Lemma establishes that a competitive selection exists.

Lemma 5 *There exists $w^*(z)$ satisfying the equalities $P = \Psi[w^*(z)] = \Psi[\Omega(P)]$.*

4 Identification and Testing

Our model is identified and estimated from longitudinal data on executive compensation, the firm's abnormal returns, and the transition choices executives make each period conditional on the values of their state variables, factors that affect their current and future payoffs. The primitives of our model are characterized by the probability density function of abnormal returns when every manager is diligent, denoted by $f_j(x|z)$; the density when only the k^{th} ranked executive shirks which we express as $f_{jk}(x|z)$; the coefficient of absolute risk aversion ρ; the parameters affecting tastes for diligent work including the effects of human capital $\alpha_{1jk}(z)$; the nonpecuniary benefits of shirking $\alpha_{0jk}(z)$; and the probability distribution of the idiosyncratic disturbance term to preferences $h(\varepsilon)$. The $f_j(x|z)$ density is identified directly from the data on abnormal returns. Moreover since $f_{jk}(x|z) \equiv f_j(x|z)/g_{jk}(x|z)$ is recovered from $f_j(x|z)$ and $g_{jk}(x|z)$, it is convenient to frame our discussion of those remaining parameters in terms of $\alpha_{0jk}(z)$, $\alpha_{1jk}(z)$, $h(\varepsilon)$ and $g_{jk}(x|z)$.

In Section 3 we demonstrated that the minimization problem solved by each firm in our framework has the same structure as the minimization problem a firm would solve if there was no human capital and no alternative job opportunities, that is where $A_s(z) \equiv U_{jk}(z) \equiv 1$. The only difference in the optimal contract distinguishing our model from this specialization emerges from interpreting the mappings $\alpha_{1jk}(z) A_s\left(z_{t+1}^{(j,k,1)}\right)$ and $U_{jk}(z)$. Gayle and Miller (2008c) analyze identification and estimation of pure moral hazard models where $A_s(z) \equiv U_{jk}(z) \equiv 1$. Given a regularity condition on $g_{jk}(x|z)$, they show that both $g_{jk}(x|z)$ and $\alpha_{0jk}(z)$ are exactly identified if and only if ρ and $\alpha_{1jk}(z)$ are identified from the competitive selection equations. Restating their results for our framework, $\alpha_{0jk}(z)$ and $g_{jk}(x|z)$ are exactly identified if and only if $\alpha_{1jk}(z)$, $h(\varepsilon)$ and ρ, and hence $A_s(z)$ and $U_{jk}(z)$, are identified from the competitive selection equations. Thus establishing identification in this framework amounts to proving $\alpha_{1jk}(z)$, $h(\varepsilon)$ and ρ are identified from the competitive selection equations.

In this section we formalize the arguments given in the previous paragraph. Following Gayle and Miller (2008c) we begin by proving that $g_{jk}(x|z)$ and $\alpha_{0jk}(z)$ are identified if ρ and $\alpha_{1jk}(z)$ are known, the former from the estimated compensation schedule, the latter from the incentive compatibility condition. Then we extend their results by showing that $\alpha_{1jk}(z)$ and ρ are identified providing one of the executive types faces a genuine choice between two different jobs at some point during his life cycle.

4.1 Parameters Characterizing Shirking

To estimate $g_{jk}(x|z)$ we impose a regularity condition that for all (j,k,z) there exists some finite return $\overline{x}$ such that $g_{jk}(x'|z) = 0$ for all $x' > \overline{x}$. This assumption implies that, should the firm performance at the end of the period be truly outstanding, then shareholders would be certain that all the executives had worked diligently during the period. Appealing to Proposition 3.1 of Gayle and Miller (2008c), we can then prove $g_{jk}(x|z)$ is identified if ρ is known. Substituting the expression derive for $g_{jk}(x|z)$ into Since the incentive compatibility condition yields an expression for $\alpha_{0jk}(z)/\alpha_{1jk}(z)$ in terms of $A_{s-1}(z)$, $\upsilon_{k,j,t+1}(x,z)$ and $g_{jk}(x|z)$, it immediately follows from their respective definitions and the expression for $g_{jk}(x|z)$ that if $\alpha_{1jk}(z)$, $h(\varepsilon)$ and ρ are known, then $\alpha_{0jk}(z)$ can be inferred.

Lemma 6

$$g_{jk}(x|z) = \frac{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - \upsilon^{-1}_{k,j,t+1}(x,z)}{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - E_t[\upsilon^{-1}_{k,j,t+1}(x,z)|z]}$$

$$\alpha_{0jk}(z) = \alpha_{1jk}(z)\left\{\left[\frac{A_{s-1}\left(z^{(1,j,k)}\right)}{A_{s-1}\left(z^{(0,j,k)}\right)}\right]\left[\frac{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - E\left[\upsilon_{j,k,t+1}(x,z)\,|l=1\right]^{-1}}{\upsilon^{-1}_{j,k,t+1}(\overline{x},z) - E_t[\upsilon^{-1}_{k,j,t+1}(x,z)|z]}\right]\right\}^{1-b_t} \tag{12}$$

The form of the shirking parameter estimates take in (12) suggests a test that differentiates the scenarios we have entertained about whether human capital is observed or not. If human capital is unobserved, and in addition a variable denoted z_{t0} enters $A_{s-1}(z)$ but not $\alpha_{0jk}(z)/\alpha_{1jk}(z)$, then from (12), variation in z_{t0} does not affect the quantity:

$$\left[\frac{A_{s-1}\left(z^{(1,j,k)}\right)}{A_{s-1}\left(z^{(0,j,k)}\right)}\right]\left[\frac{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - E\left[\upsilon_{j,k,t+1}(x,z)\,|l=1\right]^{-1}}{\upsilon^{-1}_{j,k,t+1}(\overline{x},z) - E_t[\upsilon^{-1}_{k,j,t+1}(x,z)|z]}\right]$$

Accordingly we can test the joint null hypothesis, that human capital is not observed and $\alpha_{0jk}(z)/\alpha_{1jk}(z)$ does not depend on z_{t0}, which implies:

$$cov\left\{\left[\frac{A_{s-1}\left(z^{(1,j,k)}\right)}{A_{s-1}\left(z^{(0,j,k)}\right)}\right]\left[\frac{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - E\left[\upsilon_{j,k,t+1}(x,z)\,|l=1\right]^{-1}}{\upsilon^{-1}_{j,k,t+1}(\overline{x},z) - E_t[\upsilon^{-1}_{k,j,t+1}(x,z)|z]}\right], z_{t0}\right\} = 0$$

4.2 Parameters Characterizing Diligence

This leaves $\alpha_{1jk}(z)$, $h(\varepsilon)$ and ρ to analyze from the participation equation and the choice probabilities. We now prove by an induction that $\alpha_{1jk}(z)$ and ρ are identified if at least one type of executive faces a choice between lotteries at some stage in his career . . . *more to come*

5 Data

The data for our empirical study was compiled from three sources. From Standard & Poor's ExecuComp database we extracted records on the job title and compensation of the eight highest paid executives in the S&P 500, Midcap, and Smallcap firms for the years 1992 through 2006 inclusive. Data on the employer firms were supplemented by the S&P COMPUSTAT North America database and monthly stock price data from the Center for Securities Research (CRSP) database. We matched the names, birth dates and gender of 16,300 executives from 1800 firms with information in Who's Who to augmented their records with biographical data. The resulting data set gives us unprecedented access to detailed firm characteristics, including accounting and financial data, along with their managers' characteristics, namely the main components of their compensation, including pension, salary, bonus, option and stock grants plus holdings, their socio-demographic characteristics, including age, gender, education, and a description of their career history through the five ranks and firms.

This section summarizes the aggregate features of our data set. We present estimates of the distribution of abnormal returns and show how they vary with executive characteristics. We estimate elasticities of compensation with respect to returns and measures of executive experience. Finally we investigate, empirically, how experience and other background variables affect job transitions and thus define the career paths of executives. In this way we describe the variation in the data that supports the identification and estimation of our model of executive compensation and career choice.

5.1 Summary Statistics

Most of the characteristics of the executives and firms in our sample require no explanation, but the construction of several variables merit comment. The sample of firms was initially partitioned into three industrial sectors by GICS code. Sector 1, called primary, includes firms in energy (GICS:1010), materials (1510), industrials (2010,2020,2030), and utilities (5510). Sector 2, consumer goods, comprises firms from consumer discretionary (2510,2520,2530,2540,2550) and consumer staples (3010,3020,3030). Firms in health care (3510,3520), financial services (4010,4020,4030,4040), information technology and telecommunication services (410, 4520, 4030, 4040, 5010) comprise Sector 3, which we call services. In our sample 37 percent of the firms belong to the primary sector, 28 percent to the consumer goods sector, and the remaining 35 percent to the services sector. Firm size was categorized by total employees and total assets, the median firm in each size category determining whether the other firms are called large or small. The sample mean value of total assets is \$18.2 billion (2000 US) with standard deviation \$76.2 billion, while the sample mean number of employees is 23,659 with standard deviation 65,702.

Table 1 describes the characteristics of management by sector and firm size. Jobs were assigned to a rank using the hierarchy ordering we developed in our work on gender discrimination.[8] At 27 percent, Rank 2 is the most commonly observed rank, which reflects the diversity of promotion schemes across firms. By way of contrast, the top and bottom ranks each only contribute 6 percent to the sample population. The distribution of ranks across the three sectors is roughly independent but small firms, as measured by either assets of employment, have a greater proportion of their executives congregating in the lower ranks, with 30 percent versus 20 in the bottom two ranks. Four measures of experience were included to capture the potential of on-the-job training. Executive experience is the number of years elapsed since the manager was first recorded as one of the top eight paid executives in the sample. Tenure is years spent working at the employee's current firm. We also tracked the number of moves the manager made throughout his career in different jobs and ranks, as well as the number of moves since becoming an executive. Promotion

[8] See Gayle, Golan and Miller (2008).

is a indicator variable for whether the manager was promoted recently or not.

The mean age of executives is almost 54 years with a standard deviation of about 9. Only 4 percent of the sample are female, ranging between 3 percent in the primary sector and 5 percent in the consumer sector. Roughly speaking, formal education is uniformly distributed evenly between bachelor degree or less, professional certification (in accounting or law for example), MBA, some other Master's degree, and Ph.D. The distribution is approximately independent of firm size and sector, ranging from 15 percent with an MS/MA in the consumer sector to 27 percent in small firms by employee for professionally certified executives.

Tenure in the firm averages about 14 years, about 40 years less than age, with standard deviation of about 11, two years more. The sectors are ranked the same way with respect to age and tenure; similarly firms with small assets have both the oldest executives and the longest tenure. In these respects average age, firm sector and size are almost sufficient statistics for average tenure, giving the deceptive appearance at this level of aggregation that executives within firms follow a well defined career track. Averaging across the sample, there are two rank and/or firm turnover moves per observation, one of which has occurred since acquiring executive status. About one third of executives have been promoted within the last two years.

The most important differences between the executives across firm size and sector relate to their compensation. Regardless of which measure is used, the mean salary and bonus in small firms is about two thirds the mean in large firms, about half the total compensation, with standard deviations about one third smaller.[9] This suggests that similarly named positions in small firms are not comparable to their analogues in large firms and may help explain differences between internal and external transitions.

Summarizing differences across firm type, the consumer sector has the lowest percent of executives with advance degrees and the highest percent of female executives, while the service sector has the lowest average tenure and the highest promotion rate and highest total compensation. Total compensation is roughly twice as large in large firms (using both measures), promotion and turnover rates are greater, tenure is lower, and there are more executives holding MBA degrees.

Table 2 describes the characteristics of executives by rank. The average age between Rank 1 and 3 declines from 60 to 52, but is more or less constant as rank falls off further. Similarly average tenure is roughly constant in the lower and middle ranks at 14 but rises to 15 and 17 for Ranks 2 and 1 respectively. The average gap between Ranks 1 and 3 in executive experience is 6 years. To summarize, relative to the lower ranks, Ranks 1 and 2 are 8 years older, with only 6 years more executive experience and just 2 years more tenure, late bloomers hired by the firm late in their career. Not that they are likely to move more than those who do not reach the top levels; although 8 years older the they average the same number of past moves, before and after becoming an executive.

Females form a very small fraction of the executive sample, and they are not uniformly distributed by rank. By a factor of two to three, females congregate in the lower executive

[9] We followed Antle and Smith (1985, 1986), Hall and Liebman (1998), Margiotta and Miller (2000) and Gayle and Miller (2008a, 2008b) by using total compensation to measure executive compensation. Total compensation is the sum of salary and bonus, the value of restricted stocks and options granted, the value of retirement and long term compensation schemes, plus changes in wealth from holding firm options, and changes in wealth from holding firm stock relative to a well diversified market portfolio instead.

ranks relative to males; 2 percent of the top two ranks are females, while 6 percent of Ranks 5 and 6 are female. With regard to the education background variables, the two most striking features are that there is higher percent (out of total executives in the rank) of executives with MBA degrees in the top 4 ranks, the percent of executive with another Masters degree or a Ph.D. is greater in the bottom there ranks, and there is a larger percent of executives with professional certification in the bottom 4 ranks.

Average total compensation and the salary components rise from Rank 7, are maximized at Rank 2, at levels that are more than twice as high as the corresponding figures for Rank 7, and decline. The salary component for Rank 1 is only eclipsed by Rank 2, but it is an open question whether the total financial compensation package offered for a Rank 1 position is more or less desirable than the offer for a Rank 5 position. Although the average compensation \$2.7 million for Rank 2 exceeds the Rank 5 mean by almost \$400,000, the standard deviation for the former is more than twice that of the latter. For example, if all compensation variation observed in the data was resolved before an executive accepted a position, implying the standard deviation simply reflects heterogeneity in fixed pay contracts, then there would be many Rank 5 positions that pay better than many Rank 2 positions. Alternatively if all the variation in compensation was resolved after the executive accepted his job, implying the standard deviation is a measure of the income uncertainty, the executive would prefer Rank 5 to Rank 1 position if he was sufficiently risk averse.

5.2 Abnormal returns

We defined the abnormal returns of the firm as the residual component of returns that cannot be priced by aggregate factors the manager does not control. In an optimal contract compensation to the manager might depend on this residual in order to provide him with appropriate incentives, but it should not depend on changes in stochastic factors that originate outside the firm, which in any event can be neutralized by adjustments within his wealth portfolio through the other stocks and bonds he holds. More specifically, letting ϑ_{jt} denote the value of the j^{th} firm at time t, the gross abnormal return attributable to all the executives' actions is the residual

$$x_{jt} \equiv \left(\vartheta_{jt} + d_{jt} + \sum_{k=1}^{K} w_{jkt}\right) / \vartheta_{jt-1} - \pi_t \tag{13}$$

where π_t is the return on the market portfolio in period t and d_{jt} is the dividend. This study assumes that x_t is a random variable that depends on the managers' efforts in the previous period but, conditional on the effort vector of the executive branch $\{l_{jkt}\}_{k=1}^{K}$, is independently and identically distributed across both firms and periods.[10]

We now show how abnormal returns depends on the experience and the other characteristics of the executives.

5.3 Compensation

We estimated annual excess returns for firms in equation 13 from the data, and then computed, conditional on the state variables, a nonparametric estimator of total com-

[10] In our sample the mean abnormal return is -0.005 with standard deviation 0.6, and we do not reject the null hypothesis that it is uncorrelated with the stock market.

pensation from our imputed values compiled from the data, which we assume is the sum of true compensation and independent measurement error. We used Kernel methods to nonparametrically estimate $w^o_{2mk}(x,z)$, the compensation schedule for diligent work, for each (m,k,z) as:

$$w^{(N)}_{2mk}(x,z)=\frac{\sum_{s=1,s\neq n}^{N}\sum_{t=1}^{T}w_{st}I\left\{I_{mkst}=1,z_{st}=z,\right\}K\left(\frac{x_{mt}-x}{\delta_{xN}}\right)}{\sum_{s=1,s\neq n}^{N}\sum_{t=1}^{T}I\left\{I_{mkst}=1,z_{st}=z,\right\}K\left(\frac{x_{mt}-x}{\delta_{xN}}\right)}$$

Table 3 reports OLS and LAD results from regressing how compensation varies with firms' and executives' characteristics. The (conditional) level effects are given in the first two columns of estimates, their interactions with abnormal returns in the second two. Controlling for background demographics and tenure more or less leaves intact the qualitative rank ordering on total compensation we found in Table 3. Total compensation to Ranks 6 and 7 differ by a statistically insignificant amount, and then rises with promotion, spiking at Rank 2, compensation to Rank 1 falling between Ranks 3 and 4. In contrast the unconditional means and standard deviations reported in Table 3, however, the results from the regression analysis separate the effects of excess return, which induces uncertainty to manager's total compensation, from the background variables that determine observed heterogeneity. Note that Rank 1 is more affected by excess returns than every rank except 2. Thus Rank 1 has a lower (OLS) or the same (LAD) estimated mean and more dependence on abnormal returns than Rank 3, while Rank 2 has a higher mean but more dependence than Rank 3. Therefore Rank 3 offers a superior total compensation package to Rank 1, and for sufficiently risk averse executives, a more attractive compensation package than the Rank 2. Continuing in this vein, dependence on excess returns is declining in the remaining middle or lower ranks.

All the firm size and sector variables have significant coefficients except the OLS estimator of the level effect distinguishing the consumer from service sector. None of the background variables for executives interact significantly in the OLS regression, but almost all have significant level effects irrespective of estimator. A notable exception are the coefficients relating to gender. The OLS estimator indicates that gender has no effect on compensation level or its dependence on abnormal returns, whereas the LAD estimator implies there is a small positive level effect of $91,731 and significantly reduced dependence on abnormal returns, both factors making an executive positions more attractive to females relative to males.

With respect to education the OLS results show, that after controlling for the other observed differences, Ph.D. and MBA graduates earn more than $300,000 in excess of executives with undergraduate degrees only, who earn $386,793 more than those with professional certification only. Compensation is quadratic in age as is the case in wage regressions for many occupations. Tenure, executive experience and the number of past moves have statistically significant effects on compensation but are small and inconsequential in magnitude. More noteworthy is the large estimated sign-on bonus associated with turnover, $551,859 for LAD and $994,989 for OLS.

Overall our results suggest that after controlling for rank and firm type, there are significant returns from acquiring general human capital in formal education, but little from firm specific capital that is measured in terms of tenure within any one job and/or experience acquired at a variety of jobs. Similarly gender is not a useful predictor of wages given the other executive's and other characteristics and the nature of the job. To summarize, aside from formal education, job transitions and the abnormal returns of their own firms are the main drivers determining how wealthy executives become.

5.4 State Variables and Conditional Choice Probabilities

We denote the state variables relevant for the n^{th} manager at the time t by z_{nt}, one of $Z < \infty$ possible characteristics, the ranks by $r \in \{1, \dots, R\}$ and the firm types by $s \in \{1, \dots, S\}$. In our model $z_{n,t+1}$, the n^{th} manager's state variables in the period $t+1$, are fully determined by z_{nt}, the type of firm he transitions to, denoted s_{nt}, and his rank next period, r_{nt}, by a mapping $z_{n,t+1} \equiv f(z_{nt}, r_{nt}, s_{nt})$, which we define in the next section. Our theory models the transition of z_{nt} to $z_{n,t+1}$ through the competitive equilibrium choices of (r_{nt}, s_{nt}), a stochastic process that generates the data. The structural estimation of our theoretical framework uses as input reduced form estimates of $P(r_{nt}, s_{nt} | z_{nt})$, the probability of (r_{nt}, s_{nt}) conditional on z_{nt}.

We report our estimates for the reduced form of our model. Since R and S are finite, and we assume Z is a finite set, it follows that in principle cell estimators could be used to recover $P(r_{nt}, s_{nt} | z_{nt})$. Although our sample size, 59,066, is very large compared with all previous studies of this market, the comprehensive detail that accompanies each observation also greatly magnifies the total number of cells RSZ,needed to estimate the model, so this procedure is not feasible. For example only 5 percent of the observations in our sample are female, and none of them have doctorates and head small firms. Many smoothing algorithms are asymptotically equivalent. We used multinomial logits to estimate the reduced form, because of their computational tractability in recovering the structural parameters, because the logit estimates are easy to interpret, and because they illustrate how the variation in our data is used to estimate the underlying structure. For expositional convenience we decomposed $P(r_{nt}, s_{nt} | z_{nt})$ into

$$P(r_{nt}, s_{nt} | z_{nt}) \equiv P(r_{nt} | z_{nt}, s_{nt}) P(s_{nt} | z_{nt})$$

and separately estimated $P(s_{nt} | z_{nt})$, the probability of firm type selected as a function of the state variables, from $P(r_{nt} | z_{nt}, s_{nt})$, the selection of rank conditional on both the state variables and also the firm selected.

Table 4 presents our estimates of $P(s_{nt} | z_{nt})$. The columns refer to the type of firm chosen conditional on moving from the current employer, and the state variables are defined by the rows. The omitted (column) choice is to remain employed with the current firm one more period, and the base line (row) category is a college educated Rank 1 executive employed in a firm of type 1.

MBAs go to 7. MSMAs and Ph.D.'s don't transit as much, as we saw in the previous table. controlling for other state variables we now also see that no degree executives also do not move as much as the college educated group. Female behave the same as males. Similarly tenure and male have no significant effects on the probability of an external move. Older execs are more likely to leave and conditional on leaving are less likely to go 3 than the other types.

Perhaps the most striking feature of this table is that when executives move they join firms similar to the ones they left, that is defined in terms of sector and size. Furthermore conditional on moving to a firm of different size, they are more likely to join a firm in the same sector as the one they left. Broadly speaking, the bottom rows, referring to the rank of the executive at the beginning of the period, show that highly ranked executives are less likely to move than the lower ranked ones, evident form the fact that the estimated coefficients increase in each row.

The final column of Table 4 reports on the probability of leaving the sample for at least two years and never returning, a condition we call retirement. The higher the rank the less likely the probability of retirement, indicated by the decreasing sequence of coefficients on rank. Possibly for very different reasons, executives and those without formal qualifications are less likely to exit this sample than groups with other formal education. The indicator variable for gender has a far bigger impact than any of the education variables. Mirroring female labor supply more generally, women in this highly select and lucrative market are more likely to withdraw from it than their male colleagues and competitors. Finally there are significant sector differences.

Finally our estimates of $P(r_{nt}|z_{nt}, s_{nt})$ are presented in Table 5. It shows female executives with a doctorate are more likely to select into the bottom rank. The conditional choice probability estimates shed light on the effects of tenure and age. Here we see that, controlling for all other state variables, last period employer, and this year's employer as well, Rank 2 executives are in fact older than Rank 1 executives, signified by the higher coefficient estimate. Given values of the other observed factors, lower ranked employees have more tenure. The highest coefficients invariably show staying in the same rank is the most likely outcome, and an executive in the lowest rank is more likely to move to Rank i than Rank $i+1$. Similarly Rank 4 executives are more likely to be demoted than be promoted to Rank 3, evident from the estimated coefficients in Table 4. The results in Table 4 show that relative to other executives, turnover for a Rank 2 manager is more likely than external promotion.

6 Estimation

The bulk of this section lays out a sequential estimator for the model and reports our estimates of its several components. Given the minimal movement in bond prices over this period, we assumed the bond price is constant, setting $b_t = b$ for all t. For computational ease and to reduce the computational burden we estimated the stationary infinite horizon model. The taste and human capital parameters ρ and $\alpha_{1jk}(z)$ are estimated from the participation constraint, exploiting the idea that when risk averse managers make rational choices between different uncertain outcomes or lotteries they are revealing their attitude towards risk. The likelihood ratio $g_{jk}(x|z)$ is estimated from the curvature in the optimal compensation schedule. The shirking parameters $\alpha_{0jk}(z)$ are estimated from the incentive compatibility constraint, which reflects the fact that shareholder compensate managers just enough to deter them from engaging in activities that do not maximize the value of the firm.

6.1 Utility and Human Capital

In our estimation we assumed the disturbance is standard Type 1 Extreme Value, and in our initial estimates. Substituting the estimators for $P_{jk}(z)$ and $w^o_{2jk}(z)$ obtained from the previous section into

$$E_t\left[U_{jk}(z) - \alpha_{1jk}\left(z\right)A_{s-1}\left(z\right)\upsilon_{j,k,t+1}(x,z)\left|z\right.\right] = 0$$

Our estimator of the α_{2jk} parameters and the ρ, based on the competitive selection equations, is $\sqrt{NT}$ consistent and asymptotically normal, the covariance differing from the standard formula only because the choice probabilities and the compensation schedule are estimated in the first two steps. But rather than form JKZ orthogonality conditions from the conditional expectation functions, we formed a GMM estimator from the implied covariances

$$E\left\{\left[U_{jk}(z) - \alpha_{2jklt}A_{s-1}\left(z\right)\upsilon_{j,k,t+1}(x,z)\right]z\right\} = 0$$

using the counting variables, tenure, executive experience and age as instruments, after substituting in an approximating function $\widehat{U}_{jk}(z)$ for $U_{jk}(z)$. The former differs from the latter only because consistent estimators for $P_{jk}\left(z\right)$ and $w^*_{jk}\left(z\right)$ are used instead of their true values. The remaining background variables, categorical variables signifying educational background and gender, were also used as conditioning variables in forming the orthogonality functions for the estimator. Second, when forming the recursion that defines $A_s\left(z\right)$, used in the definitions of $U_{jk}(z)$ and $\widehat{U}_{jk}(z)$, we exploited the fact that, given the manager's choice, the transition of z_{nt} to z_{nt+1} is deterministic. Using the definition of conditional probability and sufficiency:

$$\begin{aligned}
\Pr\left(z'\left|z_{nt}, l_{ljk}, I_{mkt}\right.\right) &\equiv \Pr\left[I_{j,'k't+1} = 1, z_{n,t+1} = z'\left|z_{nt}, I_{jkt} = 1, l_{ljk} = 1\right.\right] \\
&= \Pr\left[I_{j,'k't+1} = 1\left|z_{nt+1} = z', z_{nt}, I_{jkt} = 1, l_{ljk} = 1\right.\right]\Pr\left[z_{nt+1} = z'\left|z_{nt}, I_{j,'k't+1} = 1\right.\right] \\
&= \Pr\left[I_{j,'k't+1} = 1\left|z_{nt+1} = z', z_{nt}, I_{jkt} = 1, l_{ljk} = 1\right.\right]I\left\{z_{nt+1} = z'\left|z_{nt}, I_{j,'k't+1} = 1\right.\right\}
\end{aligned}$$

6.2 The Remaining Parameters

We used nonparametric methods to recover

$$g_{jk}(x|z) = \frac{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - \upsilon^{-1}_{k,j,t+1}(x,z)}{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - E_t[\upsilon^{-1}_{k,j,t+1}(x,z)|z]}$$

and:

$$\alpha_{0jk}\left(z\right) = \alpha_{1jk}\left(z\right)\left\{A_{s-1}\left(z^{(0,j,k)}\right)/A_{s-1}\left(z^{(1,j,k)}\right)\left[\frac{\upsilon^{-1}_{k,j,t+1}(\overline{x},z) - E\left[\upsilon_{j,k,t+1}(x,z)\left|l=1\right.\right]^{-1}}{\upsilon^{-1}_{j,k,t+1}(\overline{x},z) - E_t[\upsilon^{-1}_{k,j,t+1}(x,z)|z]}\right]\right\}^{1-b_t}$$

7 Investment versus Moral Hazard

In the concluding section to this paper we assess how much agency problems in executive markets are mitigated by their career concerns. Two of the four metrics we use measure the impact of an executive shirking rather than working. We estimated how much abnormal returns would fall if shareholders failed to incentivize one of its executives but continued to pay the other according to the optimal schedule. This is one measure of how much a firm stands to lose by ignoring the moral hazard problem. The executive, on the other hand, is much more concerned with the compensating differential between diligence and shirking. We computed the compensating differential to an executive from following his interests (shirking) rather than acting according to the interests of the shareholders (working diligently). The other two metrics focus on the cost of eliminating the moral hazard problem. We report on how much the firm pays to induce diligence in the presence of human capital investment, a risk premium for eliminating the moral hazard problem. Finally we calculate how much more a firm would have to pay if executives were not motivated by career concerns, ambition that helps to internalize what would otherwise be a more substantial moral hazard problem.

Each metric was computed using the structural estimates obtained from the previous section, by executive rank, averaged over firm type and executive background. Thus successive rows in Table 6 report a sample average for the rank and its standard deviation, conditional on optimal behavior by the rest of the management team. For the purposes of comparisons with other studies in this literature we also report the estimated risk aversion parameter, the top entry. Quite plausible, and comparable to previous estimates found, we note that an executive with exponential utility and risk aversion parameter of 0.45 would be willing to pay \$217, 790 to insure against an actuarially fair gamble that offers a loss of \$1 million with probability one half and a gain of \$1 million with probability one half.

The first metric is an average over $\tau_{1mk}(z)$, the expected gross loss in the value of the firm of type m in percentage terms if a rank k executive with background z tends his own interests for one year, instead of maximizing the expected value of the firm, that is before netting out the decline in expected compensation all executives would incur from the deteriorating financial performance of the firm. When all executives work diligently, by definition abnormal returns have mean zero, meaning $E[x] = 0$. Thus $\tau_{1mk}(z)$ is found by integrating abnormal returns conditional on the executive in question shirking, when every other executive works diligently:

$$\tau_{1mk}(z) \equiv E\left\{x\left[1 - g_{mk}(x, z)\right]\right\} = -E\left[x g_{mk}(x, z)\right]$$

We interpret $\tau_{1mk}(z)$ as a measure of the executive's span of control, because it indicates his potential impact on the firm from behaving irresponsibly. Not surprisingly we find Rank 2 executives exercise the greatest span of control; at 11 percent per year, a chief executives can drive the value of firm equity down to less than half its current value in 8 years, shareholders willing. Similarly, the result that the estimated span of control declines through the middle and lower ranks, confirms our intuition. More remarkable is our finding that executives in Ranks 2 and 3 have a greater span of control than those in Rank 1, as do many in Rank 4.

Taking the manager's perspective rather than the firm's, the compensating differential between working hard and shirking, which we denote by $\tau_{2mk}(z)$, is measured by differencing $w^0_{1mk}(z)$, the manager's reservation certainty equivalent wage to shirk, from $w^0_{2mk}(z)$, the manager's reservation certainty equivalent wage to work diligently under perfect monitoring. Derived from the participation constraint, these certainty equivalents can be expressed as:

$$w^0_{1mk}(z) = \frac{b_{t+1}}{\rho} \log(\alpha^{t+1,1}_{mkt}(z)) + \frac{b_{t+1}}{\rho(b_t - 1)} \log(\alpha_{1mk}/U^E_{mk}(z_m))$$

and

$$w^0_{2mk}(z) = \frac{b_{t+1}}{\rho}\log(\alpha^{t+1,2}_{mkt}(z)) + \frac{b_{t+1}}{\rho(b_t-1)}\log(\alpha_{2mk}/U^E_{mk}(z_m))$$

Thus

$$\begin{aligned}\tau_{2mk}(z) &\equiv w^0_{2mk}(z) - w^0_{1mk}(z)\\ &= \frac{b_{t+1}}{\rho}\log(\alpha^{t+1,2}_{mkt}(z)/\alpha^{t+1,1}_{mkt}(z)) + \frac{b_{t+1}}{\rho(b_t-1)}\log(\alpha_{2mk}/\alpha_{1mk})\end{aligned}$$

If a manager does not maximize the value of the firm, he gains utility from the nonpecuniary benefits of pursuing his own interests, but does not acquire so much human capital, and thus reduces his chances of higher wages and better positions in the future.

The first factor would also arise in a static model of pure moral hazard where there are no career concerns, and in our formulation does not depend on the executives background characteristics:

$$\tau^{PM}_{2mk} \equiv \frac{b_{t+1}}{\rho(b_t-1)}\log(\alpha_{2mk}/\alpha_{1mk})$$

Our estimates in Table 6 show that contemporaneous nonpecuniary shirking/working benefit differential associated with the Rank 2 position, at \$2.48 million, exceed those associated with any of the other ranks, but that the annual differential from the Rank 1 position is the next highest. Thus Rank 1 has a lesser span of control than Rank 3, but more nonpecuniary benefits. Again these benefits decline through the middle and lower ranks.

The second factor determining $\tau_{2mk}(z)$ reflects those dynamic features of our framework relating to career concerns

$$\tau^{H}_{2mk}(z) \equiv \frac{b_{t+1}}{\rho}\log(\alpha^{t+1,2}_{mkt}(z)/\alpha^{t+1,1}_{mkt}(z))$$

Here we find that, on average, the benefits of human capital accumulation decline monotonically with rank, and that compared with τ^{PM}_{2mk}, are much less dispersed throughout the population of firm types and executive backgrounds. At the lower ranks these benefits are quite considerable. On average a Rank 5 executive is willing to forego \$1.88 million per year because of the greater opportunities working diligently versus shirking affords him, while a Rank 1 executive only values the human capital component of the compensating differential at \$400,000 million per year.

By inspection the compensating differential $\tau_{2mk}(z)$ is the sum of these two factors

$$\tau_{2mk}(z) = \tau^{H}_{2mk}(z) + \tau^{PM}_{2mk}$$

Our estimates imply the compensating differential for every rank except the second is about \$2 million per year, but exceeds \$3 million per year for Rank 2 executives.

How much a firm would be willing to eliminate moral hazard is measured by $\tau_{3mk}(z)$. Under a perfect monitoring scheme shareholders would pay a manager the fixed wage of $w^0_{2mk}(z)$, and thus eliminate the risk premium they pay him in the form of a favorable lottery over the outcome of abnormal returns to induce diligent work. Hence the expected value of a perfect monitor to shareholders, denoted $\tau_{3mk}(z)$, is the difference between expected compensation under the current optimal scheme and $w^0_{2mk}(z)$, or:

$$\begin{aligned}\tau_3 &\equiv E\left[w_{mk}\left(x\right)|z\right] - w^0_{2mk}(z) \\ &= E\left[w_{mk}\left(x\right)|z\right] - \frac{b_{t+1}}{\rho}\log(\alpha^{t+1,2}_{mkt}(z)) - \frac{b_{t+1}}{\rho(b_t-1)}\log(\alpha_{2mk}/U^E_{mk}(z_m))\end{aligned}$$

Our findings in Table 6 show that the firms are prepared to pay hardly anything to eliminate the moral hazard problem at the lower ranks, but that at the Ranks 1 and 3, the benefits of a perfect monitor are considerably more. Curiously, the average risk premium paid to Ranks 1 and 3, \$1.6 million and \$1.7 million respectively, are quite close, despite the fact that the other measures of moral hazard are not.

As one final check on the relevance of human capital to resolving moral hazard problems in the executive market, we estimated the extra premium shareholders would pay to eliminate the moral hazard problem if the benefits of acquiring human capital was ignored by an executive, say because neither the organizational structure nor the market rewarded his diligence. In our model this is represented by:

$$\tau_{4mk}(z) \equiv \frac{b_{t+1}}{\rho}\log(\alpha^{t+1,2}_{mkt}(z))$$

The estimates in Table 6 show that career concerns greatly ameliorate the moral hazard problem for lower level executives but their importance declines monotonically with promotion through the ranks, bordering on irrelevance for many Rank 1 executives.

8 Appendix

Proof of Lemma 1. For the metric space $B\left(z\right)$ we define a distance metric $d\left(B_0\left(z\right), B_1\left(z\right)\right)$ and its norm $\left\|B_0\left(z\right)\right\|$, and without loss of generality rescale utility by choosing the taste parameters α_{jklt} such that:

$$\left\|\alpha^{1/b_t}_{jklt} E[\exp\left(-\varepsilon^*_{jkt}/b_t\right)|z_t]\left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}}\right\| \leq k$$

for some real number $k \in (0,1)$. Defining the operator:

$$\Gamma\left[B\left(z_t\right)\right] \equiv \sum\nolimits_{(j,k,l)} p_{jkl}\left(z_t\right)\alpha^{1/b_t}_{jklt}\left[B\left(z^{(j,k,l)}_{t+1}\right)\right]^{1-\frac{1}{b_t}} E[\exp\left(-\varepsilon^*_{jkt}/b_t\right)|z_t]\left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}}$$

it follows that:

$$
\begin{aligned}
& d\left(\Gamma\left[B_0\left(z_t\right)\right], \Gamma\left[B_1\left(z_t\right)\right]\right) \\
= & \left\| \begin{array}{l} \sum_{(j,k,l)} p_{jkl}\left(z_t\right) \alpha_{jklt}^{1/b_t}\left[B_0\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} E[\exp\left(-\varepsilon_{jkt}^*/b_t\right)|z_t] \left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}} \\ -\sum_{(j,k,l)} p_{jkl}\left(z_t\right) \alpha_{jklt}^{1/b_t}\left[B_1\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} E[\exp\left(-\varepsilon_{jkt}^*/b_t\right)|z_t] \left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}} \end{array} \right\| \\
= & \left\| \sum_{(j,k,l)} p_{jkl}\left(z_t\right) \alpha_{jklt}^{1/b_t} \left\{ \begin{array}{c} \left[B_0\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \\ -\left[B_1\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \end{array} \right\} E[\exp\left(-\varepsilon_{jkt}^*/b_t\right)|z_t] \left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}} \right\| \\
\leq & \left\| p_{jkl}\left(z_t\right) \right\| \left\| \sum_{(j,k,l)} \alpha_{jklt}^{1/b_t} \left\{ \begin{array}{c} \left[B_0\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \\ -\left[B_1\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \end{array} \right\} E[\exp\left(-\varepsilon_{jkt}^*/b_t\right)|z_t] \left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}} \right\| \\
\leq & \left\| \alpha_{jklt}^{1/b_t} E[\exp\left(-\varepsilon_{jkt}^*/b_t\right)|z_t] \left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}} \right\| \left\| \sum_{(j,k,l)} \left\{ \begin{array}{c} \left[B_0\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \\ -\left[B_1\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \end{array} \right\} \right\| \\
\leq & k \left\| \sum_{(j,k,l)} \left\{ \left[B_0\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} - \left[B_1\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \right\} \right\| \\
\leq & kd\left(\left[B_0\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}}, \left[B_1\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} \right) \\
\leq & kd\left(\left[B_0\left(z_{t+1}^{(j,k,l)}\right)\right], \left[B_1\left(z_{t+1}^{(j,k,l)}\right)\right] \right) \\
\equiv & kd\left(\left[B_0\left(z_t\right)\right], \left[B_1\left(z_t\right)\right]\right)
\end{aligned}
$$

This proves Γ is a contraction operator. By Banach's theorem there exists a unique fixed point which we denote by $A(z_t)$, the limit of $\Gamma^s[B_0(z_t)]$ for any starting point $B_0(z_t)$. Setting $B_0(z_t) = 1$ and $A_s(z_t) = \Gamma^s[B_0(z_t)]$ proves the lemma. ■

Proof of Lemma 2. For all $s \in \{1, \ldots, T\}$ we set $A_0(z_t) \equiv 1$ for all (j,k) and recursively define $A_s(z_t)$ as:

$$
A_s\left(z_t\right) = -b_t \sum_{(j,k,l)} P_{jkl}\left(z_t\right) \alpha_{jklt}^{1/b_t}\left[A_{s-1}\left(z_{t+1}^{(j,k,l)}\right)\right]^{1-\frac{1}{b_t}} E[\exp\left(-\varepsilon_{jkt}^*/b_t\right)|z_t] \left\{E[\upsilon_{jk,t+1}|z_t,l_t]\right\}^{1-\frac{1}{b_t}}
$$

where ε_{jkt}^* is the value of the period t disturbance when (j,k) is selected, b_t is the current price of a perpetual bond at t, and $z_{t+1}^{(j,k,l)}$ is the value of the state variables in period $t+1$ induced by the choices (j,k,l) in period t. From Proposition 1 of Margiotta and Miller (2000, page 678), the value function solving the consumption savings problem at retirement date $T+1$ is:

$$
V_0(z) \equiv -b_{T+1} \exp\left[-\left(a_{T+1} + \rho e_{T+1}\right)/b_{T+1}\right]
$$

where (a_{T+1}, e_{T+1}) are defined in the text. Suppose a manager works in firm and rank coordinate pair (m,k) at time T for one period and then retires. After selecting (m,k) he chooses consumption and next period's endowment (c_T, e_{T+1}) optimally to maximize:

$$
\begin{aligned}
& -\alpha_{jklT}\exp\left(-\varepsilon^*_{jkT}\right)\exp\left(-\rho c_T\right) - E_T[b_{T+1}\exp\left(-\frac{a_{T+1}+\rho e_{T+1}}{b_{T+1}}\right)\upsilon_{jk,T+1}|z_T, l] \\
= & -\alpha_{jklT}\exp\left(-\varepsilon^*_{jkT}\right)\exp\left(-\rho c_T\right) - A_0\left(z^{(j,k,l)}_{T+1}\right)E_T[b_{T+1}\exp\left(-\frac{a_{T+1}+\rho e_{T+1}}{b_{T+1}}\right)\upsilon_{jk,T+1}|z_T, l]
\end{aligned}
$$

subject of his budget constraint. From Equation (15) of Margiotta and Miller (2000, page 680), the solution to this problem yields a value function of:

$$
V_{jklT}\left(z_T\right) \equiv -b_T\alpha^{1/b_T}_{jklT}\left[A_0\left(z^{(j,k,l)}_{T+1}\right)\right]^{1-\frac{1}{b_T}}\exp\left(-\varepsilon^*_{jkT}/b_T\right)\left\{E[\upsilon_{jk,T+1}|z_T, l]\right\}^{1-\frac{1}{b_T}}\exp\left(-\frac{a_T+\rho e_T}{b_T}\right)
$$

Integrating over $(\varepsilon_{11T}, \ldots, \varepsilon_{MKT})$, the idiosyncratic disturbance vector that is revealed at the beginning of the period, and the resulting job and effort level choice (j, k, l) yields

$$
\begin{aligned}
V_T\left(z_T\right) & \equiv -b_T\sum\nolimits_{(j,k,l)}P_{jkl}\left(z_T\right)V_{jklT}\left(z_T\right) \\
& = -\exp\left(-\frac{a_T+\rho e_T}{b_T}\right)b_T\sum\nolimits_{(j,k,l)}P_{jkl}\left(z_T\right)\alpha^{1/b_T}_{jklT}\left[A_0\left(z^{(j,k,l)}_{T+1}\right)\right]^{1-\frac{1}{b_T}}E\left[\exp\left(-\varepsilon^*_{jkT}/b_T\right)|z_T\right]\{E[\upsilon_{j} \\
& = -\exp\left(-\frac{a_T+\rho e_T}{b_T}\right)A_1\left(z_T\right)
\end{aligned}
$$

The proof is completed with an induction showing that for all $s \in \{1, \ldots, T-1\}$:

$$
V_{jklt}\left(z_t\right) \equiv -b_t\alpha^{1/b_t}_{jklt}\left[A_{s-1}\left(z^{(j,k,l)}_{t+1}\right)\right]^{1-\frac{1}{b_t}}\exp\left(-\varepsilon^*_{jkt}/b_t\right)\left\{E[\upsilon_{jk,t+1}|z_t, l_t]\right\}^{1-\frac{1}{b_t}}\exp\left(-\frac{a_t+\rho e_t}{b_t}\right)
$$

and

$$
V_{T-s+1}\left(z_t\right) = -\exp\left(-\frac{a_t+\rho e_t}{b_t}\right)A_s\left(z_t\right)
$$

Suppose the equation is true for all $r \in \{1, \ldots, s\}$ for $s < T-1$. Then the solution to the consumption savings decision at time period $t = T-s-1$ is found by maximizing:

$$
\begin{aligned}
& -\alpha_{jklt}\exp\left(-\varepsilon^*_{jkt}\right)\exp\left(-\rho c_t\right) - A_{s-1}\left(z^{(j,k,l)}_{t+1}\right)E_t[b_{t+1}\exp\left(-\frac{a_{t+1}+\rho e_{t+1}}{b_{t+1}}\right)\upsilon_{jk,t+1}|z_t, l_t] \\
= & -\alpha_{jklt}\exp\left(-\varepsilon^*_{jkt}\right)\exp\left(-\rho c_t\right) - E_t[b_{t+1}\exp\left(-\frac{a_{t+1}+\rho e_{t+1}}{b_{t+1}}\right)A_{s-1}\left(z^{(j,k,l)}_{t+1}\right)\upsilon_{jk,t+1}|z_t, l_t]
\end{aligned}
$$

with respect to (c_t, e_{t+1}). Substituting t for T and $A_{s-1}\left(z^{(j,k,l)}_{t+1}\right)\upsilon_{jk,t+1}$ for $\upsilon_{jk,T+1}$ in Equation above follows directly. Integrating over $(\varepsilon_{11t}, \ldots, \varepsilon_{JKt})$, the idiosyncratic disturbance vector that is revealed at the beginning of the period, and the resulting job and effort level choice (j, k, l) yields:

$$
\begin{aligned}
V_{T-s+1}\left(z_t\right) & \equiv \sum\nolimits_{(j,k,l)}P_{jkl,T-s+1}\left(z_t\right)V_{jklt}\left(z_t\right) \\
& = -\exp\left(-\frac{a_t+\rho e_t}{b_t}\right)b_t\sum\nolimits_{(j,k,l)}P_{jkl,T-s+1}\left(z_t\right)\alpha^{1/b_t}_{jklt}A_{s-1}\left(z^{(j,k,l)}_{t+1}\right)E\left[\exp\left(-\varepsilon^*_{jkt}/b_t\right)|z_t\right]\{E[\upsilon_{j} \\
& \equiv -\exp\left(-\frac{a_t+\rho e_t}{b_t}\right)A_s\left(z_t\right)
\end{aligned}
$$

as required, the third line following from the recursive definition of $A_s(z_t)$. ■

Proof of Lemma 3. For notational convenience define:

$$W_{jklt} \equiv \log \alpha_{jklt} + (b_t - 1) \log \left[A_{s-1} \left(z_{t+1}^{(j,k,l)} \right) \right] + (b_t - 1) \log \{ E[\upsilon_{jk,t+1} | z_t, l_t] \}$$

Then (j,k) is chosen if:

$$\varepsilon_{jkt} + W_{jklt} \geq \varepsilon_{j'k't} + W_{j'k'lt}$$

for $l = l_t$. Let $G(\varepsilon_{11t}, \ldots, \varepsilon_{JKt})$ denote the probability distribution function for $(\varepsilon_{11t}, \ldots, \varepsilon_{JKt})$ and $G_{jk}(\varepsilon_{11t}, \ldots, \varepsilon_{JKt})$ its derivative with respect to ε_{jkt}. Since $G(\varepsilon_{11t}, \ldots, \varepsilon_{JKt})$ is the product of independently distributed standard Type 1 Extreme value probability distributions in our model :

$$G_{jk}(\varepsilon_{11t}, \ldots, \varepsilon_{JKt}) = \exp(-\varepsilon_{jkt}) \prod_{(j',k')} \exp\left[-\exp\left(-\varepsilon_{j'k't}\right)\right]$$

Using the well known fact that:

$$W_{jklt} - W_{j'k'lt} = \log p_{jkt} - \log p_{j'k't}$$

it now follows that :

$$\begin{aligned}
& G_{jk}(\varepsilon_{jkt} + W_{jklt} - W_{11lt}, \ldots, \varepsilon_{jkt} + W_{jklt} + W_{JKlt}) \\
= \; & \exp(-\varepsilon_{jkt}) \prod_{(j',k')} \exp\left[-\exp\left(-\varepsilon_{jkt} + W_{j'k'lt} - W_{jklt}\right)\right] \\
= \; & \exp\left[-\varepsilon_{jkt} - \sum_{(j',k')} \exp\left(-\varepsilon_{jkt} + W_{j'k'lt} - W_{jklt}\right)\right] \\
= \; & \exp\left[-\varepsilon_{jkt} - \exp(-\varepsilon_{jkt}) \left\{ \sum_{(j',k')} \exp\left(\log p_{j'k't} - \log p_{jkt}\right) \right\}\right] \\
= \; & \exp\left[-\varepsilon_{jkt} - \exp(-\varepsilon_{jkt} - \log p_{jkt}) \left\{ \sum_{(j',k')} \exp\left(\log p_{j'k't}\right) \right\}\right] \\
= \; & \exp\left[-\varepsilon_{jkt} - \exp(-\varepsilon_{jkt} - \log p_{jkt})\right]
\end{aligned}$$

From Equation the conditional choice probability for (j,k) can be expressed as

$$p_{jkt} = \int_{-\infty}^{\infty} G_{jk}(\varepsilon_{jkt} + W_{jklt} - W_{11lt}, \ldots, \varepsilon_{jkt} + W_{jklt} + W_{JKlt}) \, d\varepsilon_{jkt}$$

Hence the probability density function of $\varepsilon_{jkt}^* \equiv I_{jk}\varepsilon_{jkt}$ is Type 1 extreme value with location parameter $-\log p_{jkt}$ and unit scale parameter since:

$$\begin{aligned}
h\left(\varepsilon_{jkt}^*\right) &= p_{jkt}^{-1} \frac{\partial}{\partial \varepsilon_{jkt}^*} \left[\int_{-\infty}^{\varepsilon_{jkt}^*} G_{jk}(\varepsilon_{jkt} + W_{jklt} - W_{11lt}, \ldots, \varepsilon_{jkt} + W_{jklt} + W_{JKlt}) \, d\varepsilon_{jkt} \right] \\
&= p_{jkt}^{-1} \exp\left[-\varepsilon_{jkt}^* - \exp\left(-\varepsilon_{jkt}^* - \log p_{jkt}\right)\right] \\
&= \exp\left[-\varepsilon_{jkt}^* - \log p_{jklt} - \exp\left(-\varepsilon_{jkt}^* - \log p_{jkt}\right)\right]
\end{aligned}$$

To derive:

$$E\left[\exp\left(\varepsilon_{jkt}^*/b_t\right)\right]$$

we draw from Equations (15) and (17) of Chapter 21 of Johnston and Kotz (1970, pages 277 - 278) proving that the moment generating function for ε^*_{jkt} is:

$$E\left[\exp\left(t\varepsilon^*_{jkt}\right)\right] = \exp\left(-t\log p_{jklt}\right)\Gamma\left(1-t\right)$$

Setting $t = b_t^{-1}$ this simplifies to:

$$E\left[\exp\left(\varepsilon^*_{jkt}/b_t\right)\right] = \exp\left(-\log p_{jklt}^{1/b_t}\right)\Gamma\left[\left(b_t-1\right)/b_t\right] = p_{jklt}^{-1/b_t}\Gamma\left[\left(b_t-1\right)/b_t\right]$$

■

Proof of Lemma 5. Let Ψ denote the choice mechanism of the executive, a mapping from wages into choice probabilities derived from the discrete problem the executive solves. Thus $p = \Psi\left[w\left(z\right)\right]$. Let Ω denote the wage mechanism of the firm, derived from the cost minimization problem it solves. Thus $w\left(z\right) = \Omega\left(P\right)$. Form the composite operator $\Gamma\left[w\left(z\right)\right] \equiv \Omega\left\{\Psi\left[w\left(z\right)\right]\right\}$. We seek to show $\Gamma\left[w\left(z\right)\right]$ has a fixed point, or that there exists a wage function $w^*\left(z\right)$ satisfying $w^*\left(z\right) \equiv \Gamma\left[w^*\left(z\right)\right]$. To prove $\Gamma\left[w\left(z\right)\right]$ has a fixed point we apply a standard fixed point theorem. Hence from the definition of Γ and $w^*\left(z\right)$ it now follows that $w^*\left(z\right) \equiv \Omega\left\{\Psi\left[w^*\left(z\right)\right]\right\}$. Thus:

$$P \equiv \Psi\left[w^*\left(z\right)\right] \equiv \Psi\left(\Omega\left\{\Psi\left[w^*\left(z\right)\right]\right\}\right) \equiv \Psi\left[\Omega\left(P\right)\right]$$

as required by the lemma. ■

References

[1] **Antle, R. and A. Smith** "An Empirical Investigation of the Relative performance Evaluation of Corporate Executives," *Journal of Accounting Research*, 24 pp. 1-39, 1986.

[2] **Antle, R. and A. Smith** "Measuring Executive Compensation: Methods and as Application, " *Journal of Accounting Research*, 23 pp. 296-325, 1985.

[3] **Bajari, P. and A. Khwaja,** "Moral Hazard, Adverse Selection and Health Expenditures: A Semiparametric Analysis," NBER Working Paper No. W12445, August 2006.

[4] **D'Haultfoeviller X. and P. Fevrier,** "Identification and Estimation of Incentive Problems: Adverse Selection," Working paper, September 2007.

[5] **P. Dubois and, T. Vukina**, "Optimal Incentives under Moral Hazard and Heterogeneous Agents: Evidence from Production Contracts Data", Working paper, December 2005.

[6] **E. Duflo, R. Hanna, and S. Ryan**, "Monitoring Works: Getting Teachers to Come to School", Working paper, MIT, November 2007.

[7] **L. Einav, A. Finkelstein and P. Schrimpf**, "The Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market," NBER Working Paper No. 13228, July 2007.

[8] **Ferrall C. and Shearer B.** "Incentives and Transactions Costs Within the Firm: Estimating an Agency Model Using Payroll Records," *Review of Economic Studies*, 66, 2, 309-338, 1999.

[9] **Frydman, Carola**. "Rising Through the Ranks: The Evolution of the Market for Corporate Executives, 1936-2003." Columbia University, 2005.

[10] **Fudenberg, Drew, Bengt Holmstrom and Paul Milgrom**. 1990. " Short-Term Contracts and Long-Term Agency Relationships." *Journal of Economic Theory*, Vol. 50, pp. 1-31.

[11] **Gayle, George-Levi and Miller, Robert A.** 2008a. "Has Moral Hazard become a More Important Factor in Managerial Compensation?" forthcoming, *American Economic Review*.

[12] **Gayle, George-Levi and Miller, Robert A.** 2008b. " The Paradox of Insider Information and Performance Pay" forthcoming, *CESifo Economic Studies*.

[13] **Gayle, George-Levi and Robert A. Miller.** 2008c "Identifying and Testing Generalized Moral Hazard Models of Managerial Compensation." Tepper school of Business, Carnegie Mellon University.

[14] **Gayle, George-Levi, Limor Golan and Robert A. Miller.** 2008 "Do Female Executives Bump into Glass Ceilings?" Tepper school of Business, Carnegie Mellon University.

[15] **Gibbons R. and K. J. Murphy.** "Optimal Incentive Contract in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy, 1992, vol. 100 (3), pp 468-505*

[16] **Gibbons, R**. **and M. Waldman** "Careers in Organizations: Theory and Evidence," *Handbook of Labor Economics* Vol. 3b. pp 2373–2437, 1999

[17] **Hall, Brian J. and Jeffrey B. Liebman.** 1998. "Are CEOS Really Paid Like Bureaucrats?" *The Quarterly Journal of Economics,* August 1998, CXIII pp. 653-680.

[18] **Johnson, Norman L. and Samuel Kotz** 1970. *Continuous Univariate Distriburtions-1,* John Wiley and Sons, New York.

[19] **Lazear E.** 1992**.** "The job as a concept," in W. Bruns, ed., *Performance Measurement, Evaluations and Incentives.* Harvard University Press, Boston, MA, pg.183-215

[20] **Margiotta, Marry M. and Robert A. Miller.** 2000**.** "Managerial Compensation and The Cost of Moral Hazard." *International Economic Review,* 41 (3) pp. 669-719.

[21] **Masson, R**. "Executive Motivations, Earnings, and Consequent Equity Performance," *Journal of Political Economy*, 79 pp. 1278-1292. 1971

[22] **McCue, K**. "Promotions and Wage Growth," *Journal of Labor Economics*, Vol 14(2) pp. 175-209, April, 1996.

[23] **Neal D. and S. Rosen** "Theories of the Distribution of Earnings," in Anthony Atkinson and Francois Bourguignon, eds., *Handbook of Income Distribution.* New York: Elsevier Science, North Holland, 2000, pp. 379-427.

[24] **Nekipelov D.** (2007) "Empirical Content of a Continuous-Time Principal-Agent Model," Mimeo Duke University.

[25] **Prendergast, Canice.** 1999. "The Provision of Incentives in Firms," *Journal of Economic Literature* XXXVII pp. 7-63 (1999).

Table 1: Executives Characteristics by Sector and Firm Size
Compensation and Salary are measured in Thousand of 2006US$

Variable	Service	Primary	Consumer	Asset Small	Asset Large	Employee Small	Employee Large
Rank 1	0.04	0.05	0.07	0.04	0.06	0.04	0.06
Rank 2	0.21	0.27	0.26	0.28	0.26	0.28	0.26
Rank 3	0.07	0.06	0.09	0.05	0.08	0.05	0.08
Rank 4	0.22	0.20	0.22	0.18	0.22	0.18	0.22
Rank 5	0.20	0.17	0.18	0.15	0.18	0.15	0.18
Rank 6	0.18	0.18	0.14	0.21	0.15	0.22	0.15
Rank 7	0.08	0.06	0.04	0.09	0.05	0.08	0.06
Age	52.7	54.8	53.6	53.9	53.7	53.7	53.8
	(9.5)	(9.2)	(9.4)	(10.3)	(9.3)	(11.2)	(9.3)
Female	0.056	0.03	0.06	0.06	0.04	0.05	0.04
No Degree	0.20	0.18	0.26	0.23	0.21	0.21	0.21
Bachelor	0.82	0.81	0.73	0.77	0.79	0.78	0.78
MBA	0.23	0.24	0.22	0.19	0.23	0.18	0.23
MS/MA	0.22	0.19	0.15	0.24	0.18	0.23	0.19
Ph.D.	0.18	0.20	0.15	0.18	0.18	0.21	0.17
Prof. Certification	0.21	0.24	0.21	0.26	0.21	0.27	0.21
Executive Experience	18.28	18.7	17.9	20.6	17.1	19.4	17.2
	(53.3)	(49.8)	(18.7)	(12.3)	(11.3)	(12.1)	(11.3)
Tenure	13.62	15.0	14.28	16.2	14.1	15.7	14.1
	(10.93)	(11.5)	(11.5)	(12.07)	(11.4)	(12.1)	(11.4)
# of past moves	2.11	2.02	2.00	2.5	2.0	2.3	2.0
	(1.98)	(2.01)	(2.00)	(2.2)	(2.0)	(2.1)	(2.0)
# of executive moves	0.82	0.82	0.846	0.93	0.81	0.86	0.82
	(1.32)	(1.34)	(1.39)	(1.5)	(1.3)	(1.4)	(1.33)
Promotion	0.085	0.34	0.34	0.33	0.36	0.34	0.36
	(0.28)	(0.47)	(0.475)	(0.47)	(0.47)	(0.47)	(0.47)
Salary	442	496	584	327	544	361	546
	(271)	(296)	(392)	(185)	(334)	(233)	(334)
Total Compensation	3,270	1,841	2,041	1,350	3,022	1,538	3,056
	(14,435)	(8461)	(12,153)	(10,188)	(13,858)	(11,311)	(13,753)

*Standard Deviation in Parenthesis

Table 2: Executives Characteristics

Compensation and Salary are measured in Thousand of 2006 US$

Variable	Rank1	Rank2	Rank3	Rank4	Rank5	Rank6	Rank7
Age	59.6	55.7	52.4	52.0	52.8	52.4	52.2
	(9.8)	(7.6)	(8.0)	(8.8)	(10)	(10.3)	(11.2)
Female	0.02	0.02	0.03	0.05	0.06	0.06	0.05
	(0.13)	(0.12)	(0.16)	(0.23)	(0.24)	(0.24)	(0.21)
No Degree	0.25	0.21	0.25	0.21	0.21	0.17	0.21
	(0.43)	(0.41)	(0.43)	(0.40)	(0.41)	(0.37)	(0.41)
MBA	0.24	0.26	0.23	0.27	0.19	0.18	0.22
	(0.42)	(0.44)	(0.42)	(0.44)	(0.39)	(0.39)	(0.41)
MS/MA	0.16	0.17	0.17	0.19	0.21	0.21	0.21
	(0.37)	(0.37)	(0.37)	(0.39)	(0.41)	(0.40)	(0.40)
Ph.D.	0.15	0.15	0.14	0.13	0.21	0.27	0.17
	(0.37)	(0.35)	(0.34)	(0.33)	(0.41)	(0.44)	(0.38)
Prof. Certification	0.15	0.14	0.15	0.22	0.24	0.37	0.30
	(0.36)	(0.34)	(0.35)	(0.42)	(0.43)	(0.47)	(0.45)
Executive Experience	22.3	19.8	16.1	15.9	16.6	16.5	16.9
	(13.0)	(10.5)	(10.7)	(11.0)	(12)	(11.7)	(11.7)
Tenure	17.1	15.1	13.7	13.8	14.1	13.7	14.2
	(13.5)	(11.7)	(11.4)	(11.2)	(12)	(11.0)	(10.8)
# of past moves	1.9	1.9	1.7	1.9	2.2	2.3	2.3
	(2.0)	(1.9)	(1.9)	(1.9)	(2.0)	(2.1)	(2.1)
# of Executive	0.9	0.93	0.73	0.76	0.77	0.80	0.84
Moves	(1.4)	(1.38)	(1.3)	(0.13)	(1.32)	(1.3)	(1.4)
Salary	640	767	591	438	408	323	340
	(375)	(398)	(320)	(197)	(190)	(141)	(217)
Total	2682	4199	4055	2587	2311	1598	1867
Compensation	(18229)	(20198)	(14892)	(8536)	(7319)	(5539)	(6634)

Table 3: Compensation Regressions

Level	OLS	LAD	Slope	OLS	LAD
Constant	964.053	1,222	Excess Return	11,636.76	8,478.87
	(1,417)	(191.9)**		(967.506)**	(129.384)**
			Excess Return Square	-908.68	-238.373
				(27.210)**	(3.649)**
Consumer	-4.737	83.106	Excess Return×Consumer	2,246.78	334.718
	(161.543)	(21.863)**		(353.561)**	(47.699)**
Service	965.097	519.103	Excess Return×Service	2,694.64	1,427.43
	(149.900)**	(20.291)**		(288.870)**	(39.047)**
Assets	0.029	0.03	Excess Return×Asset	0.115	0.086
	(0.001)**	(0.000)**		(0.006)**	(0.001)**
Employees	16.82	16.613	Excess Return×Employees	34.181	32.124
	(1.346)**	(0.182)**		(4.481)**	(0.606)**
Rank 2	2,090.11	1,388.09	Excess Return×Rank 2	-388.042	1,423.73
	(289.289)**	(39.143)**		(655.597)	(88.196)**
Rank 3	896.515	65.889	Excess Return×Rank 3	-7,142.15	-5,254.64
	(352.374)*	-47.683		(745.473)**	(100.422)**
Rank 4	-197.024	-767.392	Excess Return×Rank 4	-12,219.21	-8,068.44
	(302.908)	(40.986)**		(665.071)**	(89.477)**
Rank 5	-484.074	-932.005	Excess Return×Rank 5	-14,409.11	-8,921.51
	(308.492)	(41.736)**		(675.818)**	(90.755)**
Rank 6	-998.282	-1,139.54	Excess Return×Rank 6	-14,047.82	-9,188.51
	(313.464)**	(42.411)**		(670.508)**	(90.146)**
Rank 7	-783.61	-1,109.86	Excess Return×Rank 7	-13,148.96	-9,227.35
	(379.645)*	(51.357)**		(748.188)**	(100.593)**

Table 3(cont.): Compensation Regressions

Level	OLS	LAD	Slope	OLS	LAD
Age	75.732	20.155	Excess Return× Age	136.767	29.214
	(47.603)	(6.444)**		(12.835)**	(1.711)**
Age Square	-0.879	-0.155			
	(0.411)*	(0.056)**			
Female	355.209	91.731	Excess Return×Female	-377.221	-286.293
	(339.929)	(45.917)*		(607.244)	(75.045)**
No. Degree	136.194	12.363	Excess Return×No. Degree	-622.6	-68.224
	(189.753)	(25.679)		(328.146)	(44.118)
MBA	367.872	130.474	Excess Return×MBA	-249.712	234.566
	(162.991)*	(22.060)**		(314.901)	(42.495)**
MS/MA	-79.861	-74.731	Excess Return×MS/MA	-64.16	-355.654
	(165.083)	(22.344)**		(299.351)	(40.481)**
Ph.D.	309.473	32.827	Excess Return×Ph.D.	-22.42	100.848
	(172.953)	(23.409)		(312.742)	(42.259)*
Prof. Cert.	-385.793	-101.85	Excess Return×Prof. Cert.	-1,478.81	-199.566
	(160.076)*	(21.665)**			
Exec. Experience	-0.977	-0.078	Excess Return×Exec. Experience	-2.464	-1.086
	(1.582)	(0.203)		(1.891)	(0.151)**
Tenure	-17.339	-4.573	Excess Return×Tenure	15.764	9.271
	(6.709)**	(0.906)**		(11.078)	(1.469)**
# of past moves	-32.503	-31.781	Excess Return×# of past moves	-392.886	-80.655
	(48.569)	(6.574)**		(84.423)**	(11.360)**
# of Executive Moves	52.739	21.603	Excess Return×# of Exec. moves	153.524	10.868
	(65.354)	(8.839)*		(114.343)	(15.297)
First Year with firm	994.989	551.859	Excess Return× first year in firm	-579.266	-513.588
	(464.134)*	(62.789)**		(854.534)	(115.601)**

Table 4: Multinominal Logit of Firm Choice
(Staying with your Current Firm in the Based)

Variables	1	2	3	4	5	6	Retirement
MBA	-0.026	0.205	0.146	0.167	0.413	0.353	-0.049
	(0.200)	(0.181)	(0.140)	(0.230)	(0.280)	(0.161)*	(0.036)
MS/MA	-0.467	-0.727	-0.335	-0.145	-0.107	-0.207	-0.014
	(0.225)*	(0.238)**	(0.164)*	(0.240)	(0.314)	(0.192)	(0.035)
PhD	-0.787	-0.338	-0.316	-0.281	-0.371	-0.151	-0.080
	(0.248)**	(0.217)	(0.168)	(0.270)	(0.363)	(0.205)	(0.037)*
No Degree	-0.319	-0.436	-0.298	0.435	0.184	0.113	-0.118
	(0.246)	(0.242)	(0.184)	(0.254)	(0.332)	(0.204)	(0.041)**
Moves befere Exec.	-0.141	-0.265	-0.202	-0.046	-0.315	-0.377	0.045
	(0.063)*	(0.075)**	(0.055)**	(0.066)	(0.107)**	(0.073)**	(0.010)**
Female	0.198	0.127	-0.242	-0.173	-1.410	-0.226	0.342
	(0.365)	(0.349)	(0.328)	(0.482)	(1.021)	(0.344)	(0.073)**
Tenure	-32.248	-32.149	-32.277	-31.894	-32.262	-31.935	0.010
	(1.09e+6)	(9.9e+5)	(7.8e+5)	(9.3e+5)	(1.4e+5)	(6.8e+5)	(0.002)**
Moves after Exec.	-0.024	-0.021	0.061	-0.108	-0.123	0.003	0.062
	(0.052)	(0.050)	(0.035)	(0.067)	(0.086)	(0.044)	(0.010)**
Age	0.340	0.165	0.360	0.270	0.340	0.321	0.039
	(0.105)**	(0.075)*	(0.083)**	(0.130)*	(0.173)*	(0.101)**	(0.009)**
Age square	-0.003	-0.001	-0.003	-0.002	-0.003	-0.003	-0.000
	(0.001)**	(0.001)	(0.001)**	(0.001)	(0.002)	(0.001)**	(0.000)*
Firm Type : 2	-0.197	0.650	0.463	-0.781	-0.303	-1.182	0.291
	(0.219)	(0.230)**	(0.200)*	(0.457)	(0.473)	(0.474)*	(0.044)**
Firm Type : 3	-0.932	0.049	0.640	-1.097	-1.378	-0.262	0.232
	(0.210)**	(0.223)	(0.175)**	(0.407)**	(0.516)**	(0.298)	(0.038)**
Firm Type : 4	-1.500	-1.058	-1.096	2.048	1.587	1.452	0.673
	(0.476)**	(0.538)*	(0.441)*	(0.293)**	(0.388)**	(0.304)**	(0.048)**
Firm Type : 5	-1.954	-1.316	-2.072	0.859	1.286	1.317	0.440
	(0.603)**	(0.613)*	(0.728)**	(0.383)*	(0.426)**	(0.319)**	(0.060)**
Firm Type : 6	-1.743	-1.323	-0.729	0.846	0.573	1.828	0.339
	(0.340)**	(0.370)**	(0.254)**	(0.304)**	(0.379)	(0.254)**	(0.044)**
Previous Rank :2	-1.064	0.083	0.059	-0.176	0.239	-0.277	-1.060
	(0.422)*	(0.455)	(0.277)	(0.649)	(0.768)	(0.278)	(0.054)**
Previous Rank :3	0.186	0.810	0.535	1.170	1.478	0.065	-0.560
	(0.454)	(0.503)	(0.308)	(0.662)	(0.802)	(0.331)	(0.069)**
Previous Rank :4	0.677	1.382	0.633	1.310	1.426	0.293	-0.340
	(0.373)	(0.435)**	(0.267)*	(0.606)*	(0.742)	(0.265)	(0.048)**
Previous Rank : 5	0.857	1.134	0.391	1.746	1.329	-0.255	-0.340
	(0.391)*	(0.460)*	(0.295)	(0.611)**	(0.765)	(0.313)	(0.052)**
Constant	-12.389	-8.882	-12.618	-11.794	-14.162	-11.705	-2.918
	(2.794)**	(2.086)**	(2.208)**	(3.325)**	(4.471)**	(2.603)**	(0.281)**
Observations	59066	59066	59066	59066	59066	59066	35019

Table 5 : Multinominal Logit of Rank Choice (Rank 4 is excluded)

variables	1	2	3	5
MBA	0.232	0.232	0.011	-0.021
	(0.082)**	(0.067)**	(0.069)	(0.062)
MS/MA	-0.011	-0.131	-0.117	0.014
	(0.089)	(0.073)	(0.075)	(0.061)
PhD	-0.117	-0.094	-0.147	0.187
	(0.094)	(0.076)	(0.079)	(0.060)**
No Degree	0.198	0.142	0.144	-0.086
	(0.091)*	(0.075)	(0.075)	(0.070)
Moves befere Exec.	-0.144	-0.169	-0.117	0.038
	(0.028)**	(0.023)**	(0.023)**	(0.017)*
Female	-0.749	-0.608	-0.435	0.220
	(0.214)**	(0.162)**	(0.152)**	(0.106)*
Tenure	-0.002	-0.008	-0.006	0.001
	(0.004)	(0.003)**	(0.003)*	(0.003)
Moves after Exec.	-0.008	-0.019	-0.048	0.013
	(0.026)	(0.022)	(0.023)*	(0.019)
Age	0.156	0.226	0.060	-0.009
	(0.025)**	(0.024)**	(0.022)**	(0.015)
Age square	-0.001	-0.002	-0.001	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)
Firm Type : 2	0.077	0.193	0.084	-0.224
	(0.104)	(0.086)*	(0.088)	(0.073)**
Firm Type : 3	0.283	0.352	0.216	-0.374
	(0.089)**	(0.075)**	(0.076)**	(0.067)**
Firm Type : 4	-0.585	-0.388	-0.324	0.020
	(0.133)**	(0.104)**	(0.110)**	(0.079)
Firm Type : 5	-0.262	-0.115	0.013	-0.152
	(0.148)	(0.118)	(0.118)	(0.099)
Firm Type : 6	0.239	0.195	0.191	-0.262
	(0.103)*	(0.086)*	(0.087)*	(0.077)**
Previous Rank :2	-2.196	3.745	-0.413	0.209
	(0.132)**	(0.144)**	(0.177)*	(0.296)
Previous Rank :3	-3.544	0.652	3.031	0.265
	(0.159)**	(0.154)**	(0.162)**	(0.309)
Previous Rank :4	-7.890	-4.656	-3.662	-1.951
	(0.124)**	(0.134)**	(0.145)**	(0.255)**
Previous Rank : 5	-7.181	-3.512	-2.402	3.922
	(0.232)**	(0.170)**	(0.168)**	(0.253)**

Table 6: Structural Estimates and Simulations
τ_2, τ_3 and τ_4 are measured in US100,000 of dollars
τ_1 is measured in percentage per year

Measure	Rank	Estimates	Standard Deviation.
ρ		0.45	
	1	5.2	3.4
	2	10.9	14
	3	8.3	2.9
τ_1	4	4.2	2.7
	5	1.6	1.2
	1	4.0	0.2
	2	9.0	0.5
τ_2^H	3	11.8	0.9
	4	16.4	1.3
	5	18.8	2.2
	1	18.6	34.7
	2	24.8	56.6
τ_2^{PM}	3	8.3	14.2
	4	2.5	8.6
	5	.9	1.2
	1	17.3	34.0
τ_3	2	32.5	45.6
	3	16.03	24.8
	4	1.2	2.5
	5	0.8	1.3
	1	0.5	1.4
	2	2.6	3.9
	3	12.0	14.3
τ_4	4	14.0	18.9
	5	18.2	22.7