Promotion, Turnover and Compensation in the Executive Labor Market*

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

October 14, 2011

Abstract

This paper develops an equilibrium model of the executive labor market, analyzes identification issues, and estimates it with a matched panel of firms and executives to quantify the roles of moral hazard, human capital, career concerns, and bargaining between executives and employers. We find that the investment value of acquiring more human capital reduces compensation in all ranks. This effect increases with job turnover, but declines with age and experience. We show that career concerns ameliorate conflicting shareholder and executive goals, and we explain why it is most effective at the middle ranks. Expected compensation increases with firm size; similarly, executives with more human capital are assigned to higher ranks and are paid more. Yet, these differences are mostly attributable to the risk premium, which also increases with firm size and rank. Our study pays special attention to well-networked executives, who we find receive more nonpecuniary benefits but have a lower certainty-equivalent wage; to women, who are distinguished mainly by their higher exit rate; and to educational background. We estimate MBAs have more career incentives and occupationally specific human capital than PhD graduates, but the latter group has a higher marginal product and better outside options.

Keywords: moral hazard, executive compensation, networking, promotion, turnover, human capital, career concerns, reputation, sequential equilibrium, compensating differential, certainty-equivalent wages, risk premium, structural estimation, gender, educational background differences.

1 Introduction

This paper develops and estimates an equilibrium model of the market for top executives. The model incorporates moral hazard, human capital, career concerns, bargaining, and market competition. We derive the conditions under which the model is identified, develop a multistep estimation technique, and estimate the model on a matched data set of firms and executive characteristics. We then interpret the estimated parameters within the economic framework developed in the paper.

There is a large literature on executive compensation. A number of papers have studied the importance of moral hazard in managerial compensation. (See Jensen and Murphy 1990; Garen 1994;

*We thank the participants at the 2008 World Congress on National Accounts and Economic Performance Measures for Nations 2008, Society of Labor Economists 2009, Econometric Society Summer meeting 2009, Cowles Microeconometrics Conference 2009, Society of Economic Dynamics 2009, NBER 2009 Summer Institute, CRES Applied Economics conference 2009, Conference in honor of James Heckman at the Milton Friedman Institute 2010, as well as seminar participants at Carnegie Mellon, Columbia, Duke, Harvard, NYU, Minnesota, Stony Brook, Wisconsin Madison, Queens, Rochester, UCLA, UNC Chapel Hill 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.

Hall and Liebman 1998; Haubrich 1994; Margiotta and Miller 2000; Bertrand and Mullainathan 2001; Gayle and Miller 2009 among others.) Other papers have shown that many organizational incentives for executives are not explicit but implicit through career concerns (see Gibbons and Murphy 1992; Chevalier and Ellison 1999; Dewatripont, Jewitt, and Tirole 1999; Holmstrom 2003). There is also a small but growing body of literature that argues that executive compensation is more likely the outcome of powerful executives capturing the pay process and extracting rents (see Bebchuk and Fried 2004; Kuhnen and Zwiebel 2009; Acharya and Volpin 2010; Dicks 2010). An alternative view is that executive pay is more likely the outcome of a competitive market for scarce executives (see Rose and Shepard 1997; Gabaix and Landier 2008; Terviö 2008).

None of these papers addresses the dynamic choices managers make as they balance long-term career advancement through the ranks against shorter term benefit such as higher current compensation within an equilibrium framework where rational shareholders are at an informational disadvantage. In our model, executives make sequential job and effort choices taking into account the compensation, nonpecuniary benefits from working and the future value of human capital accumulated. Their effort choices are private information, and ultimately the source of moral hazard. We incorporate career concerns by allowing human capital accumulated on the job to depend on the effort. Effort on the job can also be interpreted as a cost of investment in human capital, in addition to the monetary cost of taking a job with lower compensation which varies over the life cycle. At the beginning of every period, the equity returns of firms from decisions made in the previous period are revealed to everyone, the executives' human capital state variables are updated, and each executive is compensated following the schedule of the previous period's employment contract. Firms assess their demand for executives in the current period and post one-period contracts for positions within their firms. The one-period equilibrium spot contracts are sequentially optimal and are designed to align the goals of the executives and the shareholders. The compensation required to align the goals of the shareholders and managers in each position depends on the manager's characteristics; both the nonpecuniary benefits and the amelioration of the compensation due to career concerns varies with executives' characteristics and changes over the executives' life cycle.

Our framework incorporates several prominent models of wage determination, allowing us to analyze identification of these models and quantify the effect of each on compensation. These include the Roy (1951) model, which accounts for sorting of managers with heterogeneous productivity to firms and positions within the firms. (See Heckman and Sedlacek 1985; Heckman and Honoré 1990; Heckman, Lochner, and Taber 1998 for treatment of the empirical content of the Roy model.) We draw on standard human capital theory (Becker 1964; Ben-Porath 1967), which accounts for the pay patterns associated with investment and accumulation of human capital over the life cycle. (For estimates of human capital models driven by job choices, see Miller 1984; Sicherman and Galor 1990, among others.) Our model also incorporates the compensating-differentials model (Rosen 1974), to account for the differences in pay across firms and positions associates with the executives' nonpecuniary benefits from the job, and the heterogeneity of executives' tastes for positions and firms. (See Ekeland, Heckman, and Nesheim 2004 for identification and estimation of hedonic models.) Finally our model incorporates explicit incentives provided by formal contracts in the presence of moral hazard and rent extraction by closing the model using an ultimatum game between shareholders and the executive. In equilibrium, firms compete over executives who extract rents.

Several papers quantify the effect of several theories of wage determination mentioned above using worker employment data in an equilibrium framework (see Altug and Miller 1998; Lee and Wolpin 2006; Taber and Vejlin 2010). Our paper uses matched firm–executive data, which allows us to incorporate firm and worker heterogeneity as in Postel-Vinay and Robin (2002). This set of papers does not incorporate bargaining and agency. Using matched employee-employer data, Cahuc, Postel-Vinay, and Robin (2006) separates the effects of productivity, search frictions, bargaining power, and competition on wages in a general equilibrium model. Gayle and Golan (forthcoming) performs a decomposition exercise in a dynamic general equilibrium model with adverse selection using workers' employment data. Thus, our paper contributes to this literature by providing a unified framework for assessing the effect of moral hazard and career concerns in a framework with human capital, sorting, and compensating differentials.

The semiparametric structural econometric model we estimate comes from two equations that hold in the sequential equilibrium we analyze. The first equation applies to a manager who is indifferent between taking a given job match and exiting in equilibrium. It equates the systematic portion of his expected utility (the sum of current utility, the certainty equivalent of compensation, and the investment value of human capital), conditional on human capital and job-match choice, with the net value of the disturbance he would receive from exiting. The net value of the marginal disturbance and the value of human capital can be written as functions of the conditional choice probabilities. The results we show follow from Hotz and Miller's (1993) inversion theorem, and these probabilities have sample analogues in the data.

The second equation is derived from the wage schedule for the optimal contract. Drawing from Gayle and Miller (2011), we show that, up to a factor of proportionality, the slope of the contract identifies the likelihood ratio of abnormal returns for different effort choices. This fact provides the means for estimating the remaining parameters in the model. We also show the extent to which our model is nonparametrically identified. We prove an observational equivalence holds between long-term optimal contracts when human capital is public information and equilibrium spot contracts when information is private. This result indicates difficulties researchers encounter when trying to assess the efficiency of labor market contracts when making inferences from longitudinal data on compensation and job matches.

The conclusion to this study, Section 8, provides a detailed summary of our empirical results. Here, we briefly mention our findings on human capital, agency, and governance. Human capital acquired through experience on the job, and especially from new job matches, is an important motivator in the executive labor market. The investment value of human capital accumulation significantly reduces the equilibrium level of compensation in all ranks. It declines with age and all types of job experience. Finally, female executives place a lower value on human capital investment. The estimated risk premium increases with firm size and is highest in the service sector. Executives who deviate from shareholder goals have the biggest negative impact on abnormal returns if their employer is a small firm; however, this measure of moral hazard increases less than proportionately with firm size. The main reason why executive compensation increases with firm size is the higher cost of agency, not a greater nonpecuniary cost of working in a large firm, nor greater demand for more qualified executives in large firms. We also find the risk premium is increasing with rank, so differences across ranks in the certainty-equivalent wage are much smaller than differences in expected compensation. Career concerns mitigate the moral hazard problem at all ranks, but have the least impact at the top rank, which has the highest exit probability, and at the bottom rank, which has the lowest probability of ever attaining a higher rank. Finally, interlocked executives, and also executives in firms with more insiders on their board of directors, are less likely to exit the market for top executives, enjoy higher nonpecuniary benefits from their positions, but receive a lower certainty-equivalent wage.

The rest of the paper is organized as follows. In the next section, we describe our large longitudinal data set on executive–firm job matches, compensation, firm returns, and background characteristics. Section 3 presents the unrestricted or reduced-form empirical results on the conditional-choice probabilities characterizing turnover, promotion, and exit behavior, as well as the compensation regression. Then we present our theoretical model and its equilibrium in Section 4. The fifth section describes our estimation approach based on the equilibrium sorting conditions and presents the estimated compensation equation, decomposing it into measures of human capital investment, nonpecuniary benefits, the risk premium, and remaining factors representing net market demand. Section 6 exploits the first-order condition of the optimal contract to estimate the extent of agency including an executive's the span of control, defined as the gross loss to shareholders from an executive ignoring his contract to unilaterally pursue private interests; how much he would gain from pursuing those interests if the contract was not enforced; and the extent to which career concerns that affect his unobserved accumulation of human capital mitigate the agency problem. In Section 7, we discuss the extent to which our model is robust to alternative parametric specifications.

2 Data

The data for our empirical study was compiled from three sources. The main data source is Standard & Poor's ExecuComp database, which contains annual records on 30,614 individual executives, itemizing their compensation and describing their title. Each executive worked for one of the 2,818 firms comprising Standard & Poor's (composite) 500, Midcap, and Smallcap indices for at least one year spanning the period 1992 to 2006, which covers about 85 percent of the U.S. equities market; in the years for which we have observations, the executive was one of up to the top eight paid in the firm whose compensation was reported to the Securities and Exchange Commission. Data on the 2,818 firms for the ExecuComp database were supplemented by Standard & Poor's COMPUSTAT North America database and monthly stock-price data from the Center for Securities Research database. We also gathered background history for a subsample of 16,300 executives, recovered by matching the 30,614 executives from our COMPUSTAT database, using their full name, year of birth, and gender, with the records in *Who's Who*, which contains biographies of about 350,000 executives. The matched data 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 sociodemographic characteristics—including age, gender, education, and a comprehensive description of their career path sequence described by their annual transitions through the possible positions and firms.

2.1 Construction of Variables

Ranks In this paper, executive management is defined as an occupation of general managers in publicly traded firms whose compensation and financial assets in their employer firm are reported to the SEC. Although each firm is only required to report on its top five executives, the SEC accepts and publishes data from firms which provide the records on a greater number of employees; most firms do. Using Standard & Poor's ExecuComp database, we coded the position of each executive in any given year with one of 35 abbreviated titles, forming the basis of the hierarchy from which the ranks are constructed. We define a career hierarchy as a rational (complete and transitive) ordering over a set of job titles based on transitions independent of compensation. (See Gayle, Golan, and Miller 2011 for a detailed description of the titles and the construction of the hierarchy.) Applying this procedure, we consolidate the data into five ranks, Table 1 lists the ranks and the corresponding titles. It is clear that Rank 1 consists roughly of chairman of the board of the company or chairman of a subsidiary who does not have any other executive positions in the firm. Rank 2 consists of the CEO of the company, while Rank 3 consists mainly of chairman of board of the company who holds some executive position in the company other than CEO. Rank 4 and Rank 5 consist of other lower level executives. The first observation is that CEOs are not in Rank 1, but instead in Rank 2; since this hierarchy is based on transitions, this reflects life-cycle consideration more than control.

Total Compensation, Abnormal Return, and Bond Price We followed Antle and Smith (1985), Hall and Liebman (1998), Margiotta and Miller (2000), and Gayle and Miller (2009) 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.[1] Hence, the change in wealth from holding their firms' stock is the value of the stock at the beginning of the period multiplied by the abnormal return, defined as the residual component of returns that cannot be priced by aggregate factors the manager does not control. Our bond price series comes from the Federal Reserve's Economic Research Database and is based on Treasury bills with maturities of 1, 2, 3, 5, 7, 10, 20, and 30 years. We assumed the marginal annuitized yield rate for any bond maturing over 30 years is the same as the 30-year rate. For each date, τ, we imputed a yield curve using the data on newly issued bonds for various maturities, using a cubic spline for each date–maturity combination in the data. Using these imputed yields, we constructed the bonds for each date. (See Gayle and Miller 2009 for more details on the construction of this series.)

Sector and Firm Size Most of the executives' and firms' characteristics in the subsample of matched data require no (further) explanation, but the construction of several variables merit some remarks. The sample of firms was initially partitioned into three industrial sectors by GICS code. The first is primary and includes firms in energy (GICS:1010), materials (1510), industrials (2010, 2020, and 2030), and utilities (5510). The next, consumer goods, comprises firms from consumer discretionary (2510, 2520, 2530, 2540, and 2550) and consumer staples (3010, 3020, and 3030). Firms in health care (3510, 3520), financial services (4010, 4020, 4030, and 4040), and information technology and telecommunication services (410, 4520, 4030, 4040, and 5010) make up the services sector.

We classified firms into three sizes, large, medium, and small, based on the value of their assets and number of employees over the sample period. A firm is classified as large if both its asset value and its number of employees are above the median for its sector over the sample period and as small if both its asset value and number employees are below the median for its sector over the sample. All other firms are classified as medium.

Interlock and Large Insider Board Following the literature on corporate governance, we construct two measures of good governance and executive power. The first measure is at the executive level and is called interlock. A executive is classified as being interlocked if at least one of the following are true.

a) The executive serves on the board committee that makes his compensation decisions.

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

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

[1] Changes in wealth from holding firm stock and options reflect the costs a manager incurs from not being able to fully diversify his wealth portfolio because of restrictions on stock and option sales. When forming their portfolio of real and financial assets, managers recognize that part of the return from their firm-denominated securities should be attributed to aggregate factors, so they reduce their holdings of other stocks to neutralize those factors.

The second is at the company level and is the number of its own executives that serves on its board of directors. This measure is constructed from the variables reported in Standard & Poor's ExecuComp database indicating whether or not a given exectutive is a member of the board of directors. From this variable we created a variable for the number of insiders on the board of directors, and we classified a company as has having a large insider board if the number of insiders on its board is above the median for its sector and firm size over the sample.

Definition of the Outside Option For the purposes of this study, we define executive management as an occupation of general managers in publicly traded firms whose compensation and financial assets in their employer firm are reported to the Securities and Exchange Commission. Recall that although each firm is only required to report on its top five executives, the SEC accepts and publishes data from firms which provide the records on a greater number of employees, and most firms do. For all such firms, the SEC requirement is not a binding constraint, but a device to help the firms establish and maintain credibility with their shareholders and bondholders. Like any tightly defined occupation, executive management is porous. People become executive managers through promotion within the firm or from another publicly traded company, transfer from a privately held company or a nonprofit organization, or coming out of retirement. They exit from executive management by retiring, by accepting less prestigious and less well-paid positions within management (having been overtaken by other executives within the company and sidelined without a title change or summarily demoted), by transferring to an organization not listed on an exchange (such as starting a sole proprietorship), or entering another occupation (that makes more use of previously acquired professional qualifications, for example).

We construct a sample measure of this population-exit variable that captures the above type of exit from executive management. As such, we define our outside option called exit as an absorbing state, so executives who leave all our data sets and do not return for four years are classified as exited. Note that the following are not classified as exited by our measure: executives disappearing because the firm becomes a nonpublicly traded company, the firm is dropped from the COMPUSTAT data sets, the company merges with another company and does not exist any more, or the firm goes completely out of business, as well as executives who exit the sample in the last four years of the sample. Less than one percent of those who leave for more than three years ever show up again in our data sets. As such, we are confident we do not have a right-censuring problem.

2.2 Characteristics of Data

Table 2 presents the main characteristics of our sample by firm type. Although we report on between five and eight executives per firm, many have more than one manager in some ranks and none in others. For example a typical small firm has one person in Rank 2 (who is both CEO and President), and four others at one other rank, 4 or 5, which explains why only a tiny proportion, 0.03, are in Rank 1. Rank 4 is the mode followed by Rank 2. The exit rate is between 12 percent and 15 percent per year, but the turnover rate is much lower, about 2 percent to 3 percent per year. Executives average between 51 and 54 years old and on average have about 13 to 14 years' firm tenure. They average about 17 years of executive experience. Female executives comprise about 4 percent of the sample and are more concentrated in the consumer goods sector. Just under half of all executives are on the board of directors, but only 3 percent are interlocked. About 80 percent of executives graduated from college and about 20 percent have an MBA. The firm size differences are noteworthy. On average large firms in our sample have 50 times more assets than small firms, 19 times the equity value, 13 times as many employees, and the variation in the size of large firms relative to small firms is even greater. Total compensation averages between $1.5 and $4.5 million, with executives in the service sector at the upper end of that range, and salary comprising only about 20 percent of the total. Compensation increases substantially with firm size, as does its variability. The difference in average compensation of executives across the number of insiders on the firm's board is not significant, although executives working for firms with a small number of board members receive a lower proportion of their total compensation in salary.

Table 3 presents the main characteristics of our sample by executive rank. Rank 1 has the highest exit rate while Rank 2 has the lowest exit rate and the highest turnover rate. Average age, tenure, and executive experience increase with rank. Female executives are disproportionately represented in the lower ranks. Rank 2 executives have the most experience in other firms since becoming an executive, but the least experience with other firms before becoming an executive. Those with no college are more likely to fill the upper ranks, while those with a PhD are most likely to be found in Ranks 4 and 5. Thus, Rank 5 is the most educated by every measure except MBA while a Rank 2 executive is more likely to have an MBA than an executive in any other rank. Salary, total compensation, and the likelihood of being on the board rise with advancing rank, peak at Rank 2, and then decline at Rank 1.

3 Job Transitions and Wages

Our economic model is embedded in a dynamic system that tracks the manager's employer, rank within the firm, and compensation, given personal background. The state space for the dynamic system is the Cartesian product of the manager's age, t, and personal background, $h_t \in \{1, \ldots, H\}$, at the beginning of each period, a vector which includes his employer firm last period, $j_{t-1} \in \{1, \ldots, 36\}$, management rank last period, $k_{t-1} \in \{0, 1, \ldots, 5\}$, and fixed components (such as cohort, gender, and education) and other variable components (such as measures of executive experience). To capture aggregate conditions, we also include bond prices in our framework, but where possible suppress them in the notation. We assume throughout that, given the manager's job-match selection, (j, k), at age t, human capital is updated deterministically, denoted by $h_{t+1} \equiv H_{jk}(h_t)$. Job matches follow a stochastic law of motion: We denote by $p_{jkt}(h_t)$ the probability of choosing job match (j, k) at age t, conditional on human capital at the beginning of the period, h_t. The conditional exit probability, $p_{0t}(h_t)$, is defined the same way.

Taken together, the conditional choice probabilities for job matches, $p_{jkt}(h)$, the law of motion for human capital, $H_{jk}(h)$, and the compensation regressions described below constitute the reduced form of our structural econometric model because they are the inputs for estimating the economic model we develop in Section 4. Therefore, the reduced-form estimates are informative about patterns in the data our economic model seeks to explain. We estimated a multinomial logit model of firm type and position transitions with some (but not all) interactions to show exit, promotions, and turnover before conducting linear regressions to summarize the compensation schedule. In estimation, we exploited Bayes' rule: Given background h, the (joint) probability, $p_{jkt}(h_t)$, is the product of the probability of choosing the j^{th} firm conditional on choosing the k^{th} rank, and the (marginal) probability of choosing Rank k.

3.1 Exit

Table 4 presents the estimated coefficients and elasticity from the logit regression of the probability of exit. The regression included variables in the managers' state space. These variables are age, age squared, tenure, tenure squared, executive experience, executive experience squared, number of employers before becoming an executive, number of employers after becoming an executive, current bond price, and next-period bond price. Included indicators are Rank 1 lagged, Rank 2 lagged, Rank

3 lagged, Rank 4 lagged, board membership, interlocked, no college degree, MBA, MS/MA, PhD, and gender. The table reports the coefficients of all variables that are significant at the five percent level, and one that is marginally significant.

Table 4 shows that Rank 1 has the highest probability of exit and Rank 2 has the lowest, preserving the ordering reported in Table 3 for the (unconditional) relative frequencies. The estimated exit probability is increasing in age, tenure with the firm, years of executive experience, and the number of firms an executive has worked for. These patterns are consistent with life-cycle behavior that predicts the investment value of human capital and the scope for finding better job matches declines with the accumulation of work experience with one's current employer and in other jobs, and eventually declines with age as death approaches. The table shows that female executives are 17 percent more likely to exit than men, while those who do not have college degrees and MBA graduates are less likely to exit. Being an interlocked executive or on the board of directors also reduces the probability of exit by 55 and 65 percent respectively. Exit probabilities do not significantly differ across firm size and sector, confirming results from Table 2 that show only minor differences in the relative frequencies. Finally, exit is inversely related to the bond price. Since stocks and bond prices typically move in opposite directions, we infer executives are more likely to exit when stock prices increase.

3.2 Promotion and Demotion

Table 5 presents the estimated coefficients and elasticities from the multinomial logit regression of the probability of promotion and demotion. As with Table 4, all variables in the state space are included and we report all variables that are significant at the five percent level. Promotion is much more likely than demotion, but most executives remain in their current position. Also, after controlling for rank last period, executive experience and tenure are associated with the lower ranks this period. However, controlling for rank last period, older executives are more likely to be the CEO (Rank 2). The number of moves after being an executive does not have any effect on the probability of choosing ranks. However, the pattern for the number of moves before becoming an executive is similar to tenure. Being a board member increases the probability of being/becoming CEO.

3.3 Turnover

Table 6 presents the estimated coefficients and elasticity from the logit regression of the conditional probability of choosing a new employer. In addition to state-space variables in Tables 4 and 5, the conditional probability of choosing a new employer is conditional on the other choice variables: rank, industrial sector, and insider of board firm type. We also allowed for full interactions between the choice variables and the state-space variables. We report all variables that are significant at the five percent level.

The ordering of the top three ranks taken individually and the bottom two taken together exactly match the corresponding ordering of relative frequencies reported in Table 3. A CEO is more likely to be a new hire than executives in other ranks, and an executive in Rank 4 or 5 is less likely to be a new hire than anyone else. The probability that a Rank-2 executive just joined the firm is lower if the person is female, or (not surprisingly) was in the same rank last period. The table shows that previous turnover reduces the probability of turnover in the future. Age, tenure, and executive experience reduce turnover, the quadratic term dominating the linear term. As we found in the probability of exit, being a board member, and being interlocked reduces turnover. Also, executives in firms that have a large insider board are less likely to change firms. When the bond prices rise (and stock prices fall) turnover falls. Finally, we note that after controlling for current rank, the probability of joining a new firm is not significantly affected by last period's rank; we later appeal to this finding when constructing instruments to conduct structural estimation.

3.4 Compensation

Table 7 presents ordinary least squares estimates of the compensation schedule. In addition to the choice and state-space variables, the compensation schedule is also a function of abnormal return. We included both linear and a quadratic terms to capture the effect of abnormal returns. We also allowed for full interactions between three classes of variables, namely the state-space variables, the choice variables, and the abnormal return variables. Table 7 reports those variables that are significant at the five percent level.

The table shows that the ordering in total compensation by rank, size, and sector displayed in Tables 2 and 3 is robust to controlling for background variables. That is, average executive compensation increases up to Rank 2 and then declines, Rank 1 executives receiving a little less than Rank 3. It is increasing in firm size and executives in the service sector receive more lucrative compensation, while on average those in the primary sector receive the least. Rank 1 is most affected by excess returns, which is a little surprising given the titles of executives holding this rank (Table 1). Executives on the board are paid a premium of about $845,000, but are also more affected by firm abnormal returns.

Only in Ranks 2 and 3 is an executive in his first year at the firm paid significantly higher compensation, but expected compensation of new hires in all ranks is not as closely tied to firm performance. Compensation is more closely tied to firm's performance in larger firms, firms with more insider board members, and for interlocked executives. Similarly, being highly ranked last period, and having a lot of executive experience ties compensation firm's performance more closely. Increasing tenure reduces compensation, age has a concave profile; both trends are commonly found in other labor markets. Turning to the aggregate economy, Table 7 shows that lower bond prices increase dependence of pay on excess return, possibly reflecting a greater divergence between shareholders' interest and executives' goals when stock prices are higher.

3.5 Summarizing the Reduced Form

With the notable exception that there is less mobility between firms in the primary sector (Table 6), which could well be due to technological considerations and specialized training, firm size and sector differences only affect compensation (Table 7), not promotion, turnover, or exit (Tables 4 through 6), suggesting that a static model of compensating differentials might account for them. Exit is convex increasing in age (Table 4), older executives are more likely to be found in the highest paid ranks (Table 6), and moreover are paid a premium for any rank they hold (Table 7). This is suggestive of a nonstationary dynamic model in which aging executives become increasingly valuable to the firm, but as death approaches become less willing (and ultimately unable) to remain employed.

Job turnover complicates the model, because newly hired executives at Ranks 2 and 3 receive a substantial sign-on bonus, reinforced by declining compensation with increased tenure. Similarly, Table 7 shows that newly hired executives at all ranks are not subject to the same performance pay criteria as executives with more tenure. This result could be construed as evidence for an orientation phase in which new hires are initially given less responsibility in order to familiarize themselves with their working environment, and consequently are not held as accountable for firm performance in their first year.

Aside from age, our study focuses on two background demographics, gender, and education. Table 4 shows that women are more likely to exit than men, and Table 5 shows they are more likely to be found in the lower ranks, while Table 6 shows that they are less likely to have been promoted to Rank 2 from outside the firm than men. Yet, from Table 7, Rank 2 women are paid about $2.7 million more than men, and there are no other significant gender differences in compensation. This suggests to us that lower nonpecuniary benefits and better outside options might explain gender differences in executive behavior, possibilities taken up in our structural analysis. Either way, the argument suggests that female executives would value investment in this form of human capital less than their male counterparts, a hypothesis we formalize, quantify, and test below.

With regard to education, from Table 4 those holding an MBA degree, or no degree at all, are significantly less likely to exit. From Tables 5 through 7, educational background does not have a significant effect on job turnover or promotion and is only marginally significant in explaining compensation. This suggests that, conditional on becoming an executive, the main effect of educational background is on the value of outside options. Perhaps MBA degree holders, and those without a degree have less valuable outside options than other executives in the other education groups. Presumably fewer outside options would engender greater career concerns and make executives less prone to moral hazard, a hypothesis we develop in our structural framework.

On the surface, our findings about the mobility of interlocked executives, and those employed in companies with large numbers of insiders on the board, lend support to a theory of entrenched management. Both groups are less likely to exit (Table 4) and less likely to leave the firm for another (Table 6). But the results on compensation in Table 7 should give pause. We cannot reject the hypothesis at the five percent level that both groups are paid at the same level as other executives, yet their compensation is more sensitive to firm performance, exposing them to more risk. In the remainder of the paper, we seek an economics explanation for the stylized facts characterizing the mobility and compensation of well-networked executives, subject to the additional challenge of utilizing the same estimated framework to explain the effects of age, tenure, gender, educational attainment, and firm size.

4 The Model

Our model analyzes promotion, turnover, and executive compensation, where expected-value-maximizing shareholders are subject to moral hazard from choices made by their risk-averse managers who have private information about their own effort levels. The model is dynamic; managers accumulate both firm-specific and general occupational human capital through experience on the job. We consider two cases: when human capital depends only on information observed by everyone and is therefore public information and when human capital depends on the manager's unobserved effort and is consequently private information. In both cases, managers sequentially choose employment, bargain with firms about their compensation, and choose their effort levels, which determines the probability distribution of the returns to the firms. In the one-period sequential-equilibrium contracts we analyze, managers extract all the rent from their job matches. When human capital is public information, a risk premium is paid to align managerial goals with those of shareholders, but when human capital is private information, career concerns attenuate, and maybe even eliminate, the need for compensation to vary with firm returns.

4.1 Preferences, Choices and Human Capital

There is a finite number of firms in the executive market indexed by $j \in \{1, ..., J\}$, with $j = 0$ representing retirement. There are K different positions within each firm j, indexed by $k \in \{1, ..., K\}$ and ranked in hierarchical order. Let $t \in \{0, 1, ...\}$ denote the executive's age, and that retirement occurs upon reaching or before age $T < \infty$. Mainly to simplify the notation, we assume that executives are infinitely lived. The background of the manager is defined by age t and a vector h_t, which includes fixed demographic characteristics such as gender and education and indexes of work experience.

At the beginning of period t, managers also choose consumption, c_t, and employment. They negotiate their compensation, sign an employment contract which determines how they will be paid, and then choose their effort, which is unobserved by the shareholders. Let $d_{jkt} \in \{0,1\}$ indicate the manager's job rank k in firm j at age t, and let d_{0t} denote the indicator variable for retirement. The $JK+1$ choices are mutually exclusive, implying

$$d_{0t} + \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} d_{jkt} = 1. \tag{1}$$

Summarizing, $d_t \equiv (d_{0t}, d_{11t}, \ldots, d_{JKt})$ denotes the vector of job matches from which an executive chooses at any age t preceding retirement. There are two effort levels, called working and shirking and denoted by $l_t \in \{0,1\}$, where $l_t \equiv 0$ means the manager shirks at age t and $l_t \equiv 1$ means the manager works. Although only the manager observes his own effort, it does affect the distribution of the firm returns, the nonpecuniary utility he receives in the current period, and, in one version of our model, the human capital he accumulates.

Human capital is produced through different types of work experience. Our specification encompasses three dimensions of how human capital is accumulated. The first relates to where it can be acquired, for example in lower ranks versus higher, or in large versus small firms. The second dimension relates to where it might apply, such as to all firms versus only firms belonging to the same industry or only the firm he currently works in. For example, firm-specific experience in the j^{th} firm at rank k might increase productivity in that firm at that rank more than elsewhere.

The third dimension is who observes the manager's human capital. If human capital is public information, accumulating irrespective of whether the manager works or shirks, it follows the law of motion

$$h_{t+1} = \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} d_{jkt} \overline{H}_{jk}(h_t) \tag{2}$$

for some mapping $\overline{H}_{jk}(h_t)$. Alternatively, suppose that if the manager works in rank k in the j^{th} firm, his human capital is augmented according to the function $\overline{H}_{jk}(h_t)$, but if he shirks his human capital evolves according to the function $\underline{H}_{jk}(h_t)$. Thus, if human capital is private information it evolves following the law of motion

$$h_{t+1} \equiv \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} d_{jkt} \left[l_t \overline{H}_{jk}(h_t) + (1-l_t)\, \underline{H}_{jk}(h_t) \right]. \tag{3}$$

Note that for the specialization defined by $\overline{H}_{jk}(h) = \underline{H}_{jk}(h)$ for all (j,k,h), the effort choice l_t drops out of Equation (3), establishing that the model where human capital is private information nests the model where it is public.

A manager's preferences depend on his employer and rank, effort level, and background characteristics. Work is scaled by the factor $\alpha_{jkt}(h)$ and shirking by $\beta_{jkt}(h)$. We assume there is more disutility from working than shirking; noting that exponential utility is negative, this means $\alpha_{jkt}(h_t) > \beta_{jkt}(h_t)$. At the beginning of each period t, the manager privately observes a vector of idiosyncratic taste shocks, denoted by $\varepsilon_{0t} \equiv (\varepsilon_{0t}, \varepsilon_{11t}, \ldots, \varepsilon_{JKt})$, that are independently distributed over time. If he chooses employment in (j,k), he receives nonpecuniary taste shock of ε_{jkt} and if he retires at that age he receives ε_{0t}. After retiring, the manager receives neither income nor taste shocks. Let δ denote the subjective discount factor, ρ is the constant absolute risk-aversion parameter. The manager's lifetime utility can be summarized as

$$-\sum_{t=1}^{\infty} \delta^t \exp(-\rho c_t) \left[1 + d_{0t} \exp(-\varepsilon_{0t}) - \sum_{j=1}^{J} \sum_{k=1}^{K} d_{jkt} \left[\alpha_{jkt}(h_t)\, l_t + \beta_{jkt}(h_t)(1-l_t) \right] \exp(-\varepsilon_{jkt}) \right]. \tag{4}$$

Thus, preferences 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.

We assume there exists a complete set of markets for all publicly disclosed events relating to commodities with price measure Λ_τ and derivative λ_τ at date τ. For convenience, we assume the manager at age t knows the future sequence of bond prices, $\{b_t, \ldots, b_T\}$, an assumption that can be relaxed in theory but in empirical application would require us to estimate the bond price process. The manager's wealth is endogenously determined by his compensation, which he cannot insure against. Let ξ_t denote his endowment at age t. We also measure $w_{jk,t+1}$, the manager's compensation for employment in position k at firm type j at the beginning of year $\tau+1$, 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 $E_t[\bullet \mid l_t, d_t, h_t]$ as the expectations operator conditional on work and effort level choices at age t, the subscript on the operator indicating shocks in the commodities market. The budget constraint at age t can then be expressed as

$$E_t\left[\lambda_{\tau(t)+1}\xi_{t+1}|l_t, d_{jkt}, h_t\right] + \lambda_{\tau(t)}c_t \leq \lambda_{\tau(t)}\xi_t + E_t\left[\lambda_{\tau(t)+1}w_{jk,t+1}|l_t, d_{jkt}, h_t\right], \tag{5}$$

where $\tau(t)$ is the (calendar) date when the executive reaches age t.

4.2 Firm Technology

In our model, executives add to firm output and value in two ways. Collectively, their efforts scale up its equity value by a random variable whose distribution depends on whether they all work or at least one shirks. The other factor is an additively separable individual component that depends on the executive's characteristics and past choices, h_t. When human capital is public information, the individual deterministic factor has a level effect on the manager's marginal product, while the collective stochastic factor is the sole source of moral hazard. When human capital is private information, both factors are potential sources of moral hazard because shareholders do not observe the additively separable individual component, which is a mapping of h_t. In this case, we denote the amount of human capital shareholders attribute to him by h_t': This is the manager's reputation.

Letting $e_{j\tau}$ denote the net or equity value of firm j at the beginning of calendar time τ, the collective factor increases its value from $e_{j\tau}$ to $e_{j\tau}\pi_{j,\tau+1}$. When all the executives work $\pi_{j,\tau+1}$ is drawn from a probability density function denoted by $f_j(\pi_{j,\tau+1})$. When everyone except the k^{th} ranked executive works, $\pi_{j,\tau+1}$ is drawn from $f_j(\pi_{j,\tau+1})\,g_{jk,t(\tau)}(\pi_{j,\tau+1}|h)$, where $t(\tau)$ denotes the age of an executive at calendar time τ. Thus, $g_{jkt}(\pi|h)$, a strictly positive continuous function with $E[g_{jkt}(\pi|h)] = 1$, represents the likelihood of executive k with characteristics (t, h) shirking versus working in position (j, k) when everyone else at j works and the return is π . We follow Gayle and Miller (2009) by imposing a regularity condition on the likelihood ratio: Shareholders are certain that all the executives have worked during the period if firm performance at the end of the period is truly outstanding. Formally, we assume the likelihood ratio $g_{jkt}(\pi|h)$ converges to zero as π diverges to infinity, or that, for all (j, k, h, t),

$$\lim_{\pi\to\infty} g_{jkt}(\pi|h) = 0. \tag{6}$$

Expected returns to firms are higher if everybody works, meaning that, for all (t, h, j, k),

$$\int \pi f_j(\pi)\,d\pi > \int \pi f_j(\pi)\,g_{jkt}(\pi|h)\,d\pi. \tag{7}$$

One conflict of interest between owners and managers arises regardless of whether information about human capital is observed. Managers prefer to shirk rather than work, meaning $\alpha_{jkt}(h) >$

$\beta_{jkt}(h)$, yet from (7), shareholders prefer everyone to work. Moral hazard arises in this model because the firm only observes $\pi_{j,\tau+1}$ and its probability distribution depends on the effort of all the executives.

The other factor, denoted by $F_{jkt}(h)$, does not directly depend on the number of executives in the firm, but when human capital is private information, does depend on whether the executive has ever shirked in the past. We assume that shirking taints an executive, reducing his individual contribution from that time forward. We label a manager tainted if $l_s = 0$ for some $s < t$ and assume in that case that $F_{jkt}(h) \equiv \underline{F}$, placing an upper bound on $\underline{F}$, determined by the primitives in the model, that ensures a firm does not benefit from employing a tainted executive. This creates a second conflict of interest between shirking managers and firms.

Aside from the contribution of the firm's executives, an additively separable aggregate factor, $\pi_{\tau+1}$, multiplicatively scales $e_{j\tau}$ to affect firm output and the next period's equity value. Let the subscript $t(\tau)$ refer to the age of the manager in a given job match (j,k) at calendar time τ, and $h_{jk\tau}$ his human capital. Denoting by $Div_{j,\tau+1}$ dividends paid to shareholders at $\tau+1$, the production technology for firm j can be summarized by the accounting identity:

$$e_{j,\tau+1} + Div_{j,\tau+1} + \sum_{k=1}^{K} w_{jk,t(\tau)+1} \equiv e_{j\tau}\left(\pi_{\tau+1} + \pi_{j,\tau+1}\right) + \sum_{k=1}^{K} F_{jk,t(\tau)}\left(h_{jk\tau}\right).$$

On the right side is the gross value of the firm at the beginning of period $\tau+1$ from production in period τ, the sum of an aggregate factor, $e_{j\tau}\pi_{\tau+1}$, a firm-specific component, $e_{j\tau}\pi_{j,\tau+1}$, and an individual component associated with each position, $F_{jk,t(\tau)}(h_{jk\tau})$, that depends on the demographics of its occupant, $(t(\tau), h_{jk\tau})$. On the left side is the sum of retained firm assets, $e_{j,\tau+1}$, dividends, $Div_{j,\tau+1}$, and managerial compensation to each of its managers, $w_{jk,t(\tau)+1}$.[2]

4.3 Employment Choices

Job matches differ in their pecuniary compensation, systematic and idiosyncratic nonpecuniary benefits, and investment in human capital. It follows that the characteristics of potential jobs in the future determine the current value of human capital conditional on choices made right now. To derive a formula for human capital in our model, let

$$\upsilon_{jk,t+1} \equiv \exp\left(-\rho w_{jk,t+1}(h_t, \pi_t)/b_{t+1}\right) \tag{8}$$

denote the risk-adjusted utility weight of a manager t years old for receiving compensation $w_{jk,t+1}$ at the beginning of next period for working in position (j,k). Also, let $p_{jkt}(h)$ denote the probability of choosing (j,k) at age t conditional on h: that is, everything except the idiosyncratic shock vector ε_t. Similarly, we denote the retirement probability by $p_{0t}(h)$. Finally, denote by ε^*_{jkt} the value of ε_{jkt} conditional on choosing (j,k) at t. Thus, $\varepsilon^*_{jkt} = \varepsilon_{jkt}$ if $d_{jkt} = 1$ and is not defined if $d_{jkt} = 0$.

We now define an index of human capital for a t-year-old executive with characteristics h who always works as

$$\begin{aligned} A_t(h) = {} & p_{0t}(h)\, E\left[\exp\left(\frac{-\varepsilon^*_{0t}}{b_t}\right)\right] \\ & + \sum_{j=1}^{J}\sum_{k=1}^{K} p_{jkt}(h)\left[\alpha_{jkt}(h)\right]^{\frac{1}{b_t}} E\left[\exp\left(\frac{-\varepsilon^*_{jkt}}{b_t}\right)\right]\left\{A_{t+1}\left[\overline{H}_{jk}(h)\right] E_t[\upsilon_{jk,t+1}]\right\}^{1-\frac{1}{b_t}}. \end{aligned} \tag{9}$$

[2] We can account for the fact that not all positions in the firm are filled at any one time by setting $t(\tau) = \infty$ if match (j,k) is unfilled at calendar time τ, and writing $F_{jk,\infty}(h_{jk\tau}) = w_{jk\infty} = 0$.

Note that $A_t(h)$ is a choice-probability-weighted average of expected outcomes from making different (j,k) choices, including retirement. By inspection, the index is strictly positive, and lower values of $A_t(h)$ are associated with a higher value of human capital. Thus, increasing expected compensation reduces $E_t[\upsilon_{jk,t+1}]$ and $A_t(h)$. Similarly, $A_t(h)$ is monotone increasing in $\alpha_{jkt}(h)$, the utility-weighted nonpecuniary losses of job characteristics. The rationale for defining a human-capital index in this way is provided by Lemma 1.

Lemma 1 *Let a_t denote the price of a security paying the (random) dividend $(\ln \lambda_{t+s} - (t+s)\ln\delta)$. Let $V_t(h, \xi_t, a_t)$, denote the discounted sum of expected utility from age $t < R$ onwards, for an executive with characteristics h and wealth ξ_t who has not yet observed ε_t and will make optimal consumption and job-match choices thereafter, subject to the constraint that he will never shirk:*

$$V_t(h, \xi_t, a_t) = -b_t \exp\left(-\frac{a_t + \rho\xi_t}{b_t}\right) A_t(h). \tag{10}$$

We remark that $-b_t \exp[-(a_t + \rho\xi_t)/b_t]$ is the well-known value function for an infinitely lived retiree with exponential utility, risk-aversion parameter, ρ, and wealth, ξ_t, who optimally smooths consumption over his lifetime. Thus, Equation (10) shows that the optimized lifetime expected utility is the product of utility from financial wealth and human capital.

Upon observing ε_t, the idiosyncratic features of the set of job options, the executive makes his employment decision mindful of the expected utility from wages, $E_t[\upsilon_{jk,t+1}]$, the effect on his human capital as of next period, $A_{t+1}\left[\overline{H}_{jk}(h)\right]$, and the nonpecuniary features of the job, both systematic, $\alpha_{jkt}(h)$, and idiosyncratic, ε_{jkt}. Lemma 2 formulates the trade-off, demonstrating that each of these components is additively separable in logarithmic form.

Lemma 2 *If $t \leq R$ and $l_s = 1$ for all $s \in \{t, \ldots, R\}$, then job choices d_t are picked to sequentially maximize*

$$d_{0t}\varepsilon_{0t} + \sum_{j=1}^{J}\sum_{k=1}^{K} d_{jkt}\left\{\varepsilon_{jkt} - \ln\alpha_{jkt}(h) - (b_t - 1)\ln A_{t+1}\left[\overline{H}_{jk}(h)\right] - (b_t - 1)\ln E_t[\upsilon_{jk,t+1}]\right\}. \tag{11}$$

In equilibrium, the vector of conditional choice-probability functions, $p_t(h) \equiv (p_{11t}(h), \ldots, p_{JKt}(h))$, that the manager uses to compute $A_t(h)$ in Equation (9) are precisely the probability functions that characterize his choices when solving the optimization function described by (11). We appeal to Proposition 1 of Hotz and Miller (1993), which states that there exists a mapping $q(p) \equiv (q_{11}[p_t(h)], \ldots, q_{JK}[p_t(h)])$ from the simplex to R^{JK} such that

$$q_{jk}[p_t(h)] = \ln\alpha_{jkt}(h) + (b_t - 1)\ln A_{t+1}\left[\overline{H}_{jk}(h)\right] + (b_t - 1)\ln E_t[\upsilon_{jk,t+1}]. \tag{12}$$

Given h, the solution to the optimization problem in Equation (11) only depends on the vector of differences $(\varepsilon_{11t} - \varepsilon_{0t}, \ldots, \varepsilon_{JKt} - \varepsilon_{0t})$ rather than their levels ε_t. This becomes apparent from substituting out

$$d_{0t} = 1 - \sum\nolimits_{j=1}^{J}\sum\nolimits_{k=1}^{K} d_{jkt}$$

in Equation (11), collecting terms involving d_{jkt}, and noting the additive constant ε_{0t} has no effect on the optimal choices. Substituting Equation (12) into (11), we see that if position (j,k) is the optimal employment choice, then $\varepsilon_{jkt} - \varepsilon_{0t} > q_{jk}[p_t(h)]$ and

$$(j,k) = \underset{(j',k')}{\arg\max}\left\{\varepsilon_{j'k't} - q_{j'k'}[p_t(h)]\right\}. \tag{13}$$

Given (t, h), the manager is indifferent between all positions if ε_t satisfies the condition

$$(\varepsilon_{11t} - \varepsilon_{0t}, \ldots, \varepsilon_{JKt} - \varepsilon_{0t}) \equiv q\left[p_t(h)\right] \equiv (q_{11t}, \ldots, q_{JKt}). \tag{14}$$

It now follows that $(\varepsilon_{0t}, q_{11t} + \varepsilon_{0t}, \ldots, q_{JKt} + \varepsilon_{0t})$ defines, for all ε_{0t}, the set of idiosyncratic shocks ε_t for a manager who would marginally accept any of the JK positions or retire.

Figure 1 illustrates the inversion result for a choice set of two jobs plus retirement, $JK = 2$. For convenience, we plot ε_{0t} on the vertical axis and drop the time subscript. Focusing initially on the top-left frame, we see that $(q_1, q_2) = (0, 70)$. If $\varepsilon_0 > \max\{\varepsilon_1, \varepsilon_2 - 70\}$, the worker exits; similarly, if $\varepsilon_2 > \max\{\varepsilon_0 - 70, \varepsilon_1 - 70\}$, the worker chooses the second job. Thus, the three regions of choice are displayed for the entire $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$ space by the half planes defined by $\varepsilon_0 = \max\{\varepsilon_1, \varepsilon_2 - 70\}$, $\varepsilon_2 = \max\{\varepsilon_0 - 70, \varepsilon_1 - 70\}$ and $\varepsilon_2 = \max\{\varepsilon_0, \varepsilon_1 - 70\}$. The union of the half planes is an object resembling an infinitely large paper jet plane with cantilevered wings joined at the top of the fuselage or keel along the (gutter like) line $\varepsilon_0 = \varepsilon_1 = \varepsilon_2 - 70$. Superimposed over the three regions is an isosphere with density 5.0E $-$ 4 for a trivariate normal distribution of $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$ formed from three identically and independent normal distributions, each with mean zero and standard deviation 100. Since the normal is unimodel about its mean, isospheres with lower densities form larger balls enclosing the one depicted. Integrating $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$ within each region, we obtain the conditional choice probabilities, which in the example are $(p_0, p_1, p_2) = (0.415, 0.415, 0.17)$. The bottom-left frame depicts exactly the same situation as the top left, but from a different perspective. Now suppose the first job becomes less desirable, q_1 increasing from 0 to 50. The new choice regions are illustrated in the right two panels; integrating over the probability density function, the conditional choice probabilities become $(0.52, 0.26, 0.22)$, some workers exiting from the second job and others switching from the second to the third job.

4.4 Effort Choices

In our model, shirking by just one manager is disguised, because every firm return outcome that might occur when one manager shirks could also occur when every manager works; technically, the likelihood ratio, $g_{jkt}(\pi \mid h)$, is bounded. Similarly in the private-information model of human capital, firms cannot definitively recognize job-match profiles of tainted managers: In the equilibrium we analyze below, every job history has strictly positive mass even though no shirking occurs along the equilibrium path. Underlying this result is our assumption that ε_{jkt} has full support and is private information.

The allure of shirking depends on whether it affects human capital or not. Extending the choice set to include shirking is more straightforward when information about human capital is public. In that case, the transition of human capital only depends on d_t, which is directly observed by shareholders, but not l_t. Analogously to the definition of $A_t(h)$, we define the recursion

$$A_t^o(h) = p_{0t}(h) E_t\left[\exp\left(\frac{-\varepsilon_{0t}^*}{b_t}\right)\right] + \sum_{j=1}^{J}\sum_{k=1}^{K}\left\{p_{jkt}(h) E_t\left[\exp\left(\frac{-\varepsilon_{jkt}^*}{b_t}\right)\right] A_{t+1}^o\left[\overline{H}_{jk}(h)\right] V_{jkt}^o(h)\right\}, \tag{15}$$

where

$$V_{jkt}^o(h) \equiv \min\left\{\alpha_{jkt}(h)^{\frac{1}{b_t}} E_t\left[\upsilon_{jk,t+1}\right]^{1-\frac{1}{b_t}}, \beta_{jkt}(h)^{\frac{1}{b_t}} E_t\left[\upsilon_{jk,t+1} g_{jkt}(\pi, h)\right]^{1-\frac{1}{b_t}}\right\}. \tag{16}$$

Comparing (9) with (15), the only differences arise because the manager has choice of shirking. This does not affect the state variables determining the next period's human capital, but it does give him another JK combinations of nonpecuniary and financial packages to choose from. The increased choice set implies $A_t^o(h) \leq A_t(h)$. Following the same steps as the proof to Lemma 2, we can show that the optimization problem with public human capital simplifies to sequentially choosing job matches that maximize

$$\varepsilon_{0t}d_{0t} + \sum_{j=1}^{J}\sum_{k=1}^{K} d_{jkt}\left[\varepsilon_{jkt} - \ln V_{jkt}^{o}(h) - (b_t - 1)\ln A_{t+1}^{o}\left[\overline{H}_{jk}(h)\right]\right]. \tag{17}$$

In the case of private human capital, the contract is based on the manager's reputation, h_t', not the manager's actual human capital, h_t, but if the executive shirks, firm returns are based on $g(\pi, h) f(\pi)$, not $g(\pi, h') f(\pi)$. Consequently, the conditional choice probabilities depend on both the manager's actual human capital, h_t, and the manager's reputation, h_t'. Our discussion in the previous paragraph implies h_t' follows the law of motion $h_{t+1}' = \overline{H}_{jk}(h_t')$. In truth, if a manager deviates and shirks at age t, next period his human capital is $h_{t+1} = \underline{H}_{jk}(h_t)$. To complete the description of the manager's choice problem, we now formulate the value of job matches to the manager when $h_t' \neq h_t$.

Denote the manager's choice probabilities over positions in firms by $p_{jkt}(h, h')$. Since compensation payments are based on human capital attributed by the firm to the manager, in place of $\upsilon_{jk,t+1}$, the risk-adjusted utility from compensation is

$$\upsilon_{jk,t+1}' \equiv \exp\left(-\rho w_{jk,t+1}(h_t', \pi_t)/b_{t+1}\right), \tag{18}$$

Analogously to the definition of $A_t(h)$, we define the recursion

$$B_t(h, h') = p_{0t}(h, h') E_t\left[\exp\left(\frac{-\varepsilon_{0t}^*}{b_t}\right)\right] + \sum_{j=1}^{J}\sum_{k=1}^{K}\left\{p_{jkt}(h, h') E_t\left[\exp\left(\frac{-\varepsilon_{jkt}^*}{b_t}\right)\right] V_{jkt}'(h, h')\right\}, \tag{19}$$

where

$$V_{jkt}'(h, h') \equiv \min\left[\begin{array}{c} \alpha_{jkt}(h)^{\frac{1}{b_t}}\left\{B_{t+1}\left[\overline{H}_{jk}(h), \overline{H}_{jk}(h')\right] E_t\left[\upsilon_{jk,t+1}'\right]\right\}^{1-\frac{1}{b_t}} \\ \beta_{jkt}(h)^{\frac{1}{b_t}}\left\{B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h')\right] E_t\left[\upsilon_{jk,t+1}' g_{jkt}(\pi, h)\right]\right\}^{1-\frac{1}{b_t}} \end{array}\right]. \tag{20}$$

The difference between $A_t(h)$ and $B_t(h, h')$ stems from the minimization undertaken to define $V_{jkt}'(h, h')$, the conditional valuation function of match (j, k) for a manager with demographics (t, h) and reputation h'. The top entry in the minimization operator of (20) is the conditional valuation function, net of lifetime utility conferred by endowment wealth, of a manager at age t in position (j, k) with human capital h and reputation h' from choosing to work. If the manager never had a shirking option, then human capital and reputation would always be equated and $B_t(h, h')$ would simplify to $A_t(h)$. Thus, the bottom element is a conditional valuation function for a similarly placed manager from choosing to shirk; he reaps the immediate benefit from shirking since $\beta_{jkt}(h) < \alpha_{jkt}(h)$, but firm returns are drawn from $g_{jkt}(\pi, h) f(\pi)$ rather than $f(\pi)$, affecting the probability distribution of his compensation; his reputation subsequently diverges further from his true human capital.

Lemma 3 now extends the job-match problem in Equation (11) to include the choice of effort.

Lemma 3 *If $h_{t+1}' \equiv \overline{H}_{jk}(h_t')$, then job matches d_t and effort levels l_t are picked to sequentially maximize*

$$\varepsilon_{0t}d_{0t} + \sum_{j=1}^{J}\sum_{k=1}^{K} d_{jkt}\left[\varepsilon_{jkt} - \ln V_{jkt}'(h, h')\right] \tag{21}$$

Again an (omitted) induction could be used to prove that $B_t(h, h') \leq A_t(h)$ for any given compensation schedule, essentially because adding the option to shirk unambiguously increases the opportunity set. Consequently the value from solving (21) exceeds the value from solving (11). Recalling, from our discussion below Equation (3), the fact that the private-information human-capital case nests the public-information case, it immediately follows that if $\underline{H}_{jk}(h) = \overline{H}_{jk}(h)$ for all (j, k, t, h), then $B_t(h, h) = A_t^{o}(h)$ for all (t, h).

4.5 Contracts

We define the certainty-equivalent wage that solves Equation (12) as

$$w_{jk,t+1}^{*}(h)=\frac{b_{t+1}}{\rho}\left\{\frac{1}{(b_t-1)}\ln\alpha_{jkt}(h)+\ln A_{t+1}\left[\overline{H}_{jk}(h)\right]-\frac{1}{(b_t-1)}q_{jk}\left[p_t(h)\right]\right\}. \tag{22}$$

Supposing effort could be costlessly monitored and working was demanded, it follows from Equations (8) and (12), that if a cohort of managers aged t all with human capital h were confronted with job opportunities across K ranks in J firms offering $w_{jk,t+1}^{*}(h)$, they would sort into the jobs following the probability distribution $p_t(h)$. The equilibrium compensation schedule must respect the participation constraint implied by the certainty-equivalent wage to attract the mix of managers dictated by the conditional choice probabilities.

There is little reason to presume that a contract respecting the participation constraint induces working over shirking. Given our assumptions about the production technology, it is unprofitable for a firm to employ a manager to shirk, and, in the case of private information about human capital, to employ a manager who has ever shirked. Because firms do not observe effort expended, they resort to voluntarily inducing working through the use of incentives in their employment contracts. The trade-off between working and shirking depends on whether human capital is public information or not.

When h is public information, shirking has no ramifications for the manager beyond the current period. Inspecting the optimization problem given by (17), and in particular the definition of $V_{jkt}^{o}(h)$ given by (16), the firm can deter shirking in a one-period contract by offering a compensation schedule that satisfies the incentive-compatibility constraint:

$$\left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)}\right]^{1/(b_t-1)}\leq\frac{E_t\left[\upsilon_{jk,t+1}g_{jkt}(\pi\mid h)\right]}{E_t\left[\upsilon_{jk,t+1}\right]}. \tag{23}$$

Suppose the manager were paid a constant wage, such as the certainty-equivalent $w_{jk,t+1}^{*}(h)$. The right side of the (23) then simplifies to

$$\frac{\exp\left(-\rho w_{jk,t+1}^{*}(h)/b_{t+1}\right)E_t\left[g_{jkt}(\pi\mid h)\right]}{\exp\left(-\rho w_{jk,t+1}^{*}(h)/b_{t+1}\right)}=1, \tag{24}$$

and since $\alpha_{jkt}(h)>\beta_{jkt}(h)$, the inequality given by (23) is violated. In other words, paying a constant wage guarantees shirking when human capital is public information.

In contrast, career concerns may ameliorate the divergence of incentives between managers and shareholders when human capital is private information. Consider shareholders offering employment to a manager who they believe has never shirked, and accepting a one-period contract that convinces the manager to work. From the definition of $V_{jkt}^{\prime}(h,h')$ given in (20) the compensation schedule must satisfy the incentive-compatibility constraint:

$$\left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)}\right]^{1/(b_t-1)}\leq\frac{E_t\left[\upsilon_{jk,t+1}g_{jkt}(\pi\mid h)\right]B_{t+1}\left[\underline{H}_{jk}(h),\overline{H}_{jk}(h)\right]}{E_t\left[\upsilon_{jk,t+1}\right]A_{t+1}\left[\overline{H}_{jk}(h)\right]}. \tag{25}$$

Thus, whenever $A_{t+1}\left[\overline{H}_{jk}(h)\right]<B_{t+1}\left[\underline{H}_{jk}(h),\overline{H}_{jk}(h)\right]$, career concerns ameliorate the agency problem. Setting compensation to any constant wage, we see that (25) is satisfied if and only if

$$\ln\alpha_{jkt}(h)+(b_t-1)\ln A_{t+1}\left[\overline{H}_{jk}(h)\right]\leq\ln\beta_{jkt}(h)+(b_t-1)\ln B_{t+1}\left[\underline{H}_{jk}(h),\overline{H}_{jk}(h)\right]. \tag{26}$$

When the ratio of investment value of human capital is large enough relative to the increase in disutility from working, the incentive compatibility does not bind, obviating the need to tie remuneration to the abnormal returns of the firm and pay a risk premium.

Because the formulation of the incentive-compatibility constraint depends on whether human capital is public or private information, so does the compensation schedule that minimizes expected wage payments from employment subject to the participation and incentive-compatibility constraints. To highlight the differences, let $1\{private\}$ denote an indicator function taking a value of one if human capital is private and zero if not. We now define the variable part of compensation by

$$r_{jk,t+1}(h,\pi) \equiv \frac{b_{t+1}}{\rho} \ln \left[1 - \eta g_{jkt}(\pi \mid h) + \eta \left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)} \right]^{1/(b_t - 1)} \left[\frac{A_{t+1}\left[\overline{H}_{jk}(h)\right]}{B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]} \right]^{1\{private\}} \right], \quad (27)$$

where η is the unique positive root to

$$\int \left\{ \frac{f_j(\pi)}{\eta^{-1} + \left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)}\right]^{1/(b_t-1)} \left[\frac{A_{t+1}\left[\overline{H}_{jk}(h)\right]}{B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]} \right]^{1\{private\}} - g_{jkt}(\pi \mid h)} \right\} \mathrm{d}\pi = 1. \quad (28)$$

It is evident from (27) that a greater $g_{jkt}(\pi \mid h)$ leads to a lower $r_{jk,t+1}(h,\pi)$. Contracting to pay less in states that are relatively more likely to occur when there is shirking encourages the manager to work. Since $g_{jkt}(\pi \mid h) \to 0$ as $\pi \to \infty$, it follows that $r_{jk,t+1}(h,\pi)$ has a finite upper bound of

$$\overline{r}_{jk,t+1}(h) \equiv \frac{b_{t+1}}{\rho} \ln \left[1 + \eta \left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)} \right]^{1/(b_t-1)} \left[\frac{A_{t+1}\left[\overline{H}_{jk}(h)\right]}{B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]} \right]^{1\{private\}} \right]. \quad (29)$$

The higher the firm's returns, the less likely they could have been generated by shirking, and hence the lower the slope of the variable component to compensation.

Lemma 4 states that the optimal contract is the sum of the compensating-equivalent wage and the variable component defined in the optimal contract.

Lemma 4 *If $h' = h$, then the cost-minimizing one-period contract that attracts a manager of age t with experience h to select the k^{th} position in the j^{th} firm with probability $p_t(h)$ and work is*

$$w_{jk,t+1}(h,\pi) = w^*_{jk,t+1}(h) + r_{jk,t+1}(h,\pi). \quad (30)$$

In this model, the optimal long-term contract can be implemented by a sequence of the one-period contracts defined in (30) if human capital is public information. Intuitively, if the firm is not serving a banking function for wealth the manager has already accumulated, and if the firm does not receive any further information about a shirking deviation after the period it occurs, then any punishment the firm might wish to administer for poor performance can be administered immediately (See Malcomson and Spinnewyn 1988; Fudenberg, Holmstrom, and Milgrom 1990; Rey and Salanie 1990).

However, these arguments do not apply when human capital is private information. In this case, tainted executives affect the future returns to the firm, both directly through $\underline{F}$, and also, since $h \neq h'$ for tainted executives, indirectly through the cost of achieving incentive compatibility. Thus, a long-term contract that promises to punish managers for poor firm performance several periods from now has a current deterrent effect, and when used in conjunction with immediate punishment is therefore potentially cheaper to implement because more than one signal is used to achieve incentive compatibility in any given period. We interpret the optimal one-period contract in the model with private information about human capital as an economically meaningful departure from the null hypothesis that the data can be rationalized by a sequence of short-term contracts replicating an optimal long-term contract.

4.6 Sequential Equilibrium

An ultimatum game is at the heart of negotiations between managers and shareholders over compensation in our model. After some preliminary discussions, managers have one final opportunity to propose a plan, which shareholders can either approve in its entirety or reject. If the final proposal is rejected, the manager retires and shareholders neither lose nor receive anything. Otherwise, the contract is implemented in accordance with the final proposal. Preceding this final negotiation round could be all manner of bargaining, but since preceding rounds are assumed to be neither costly nor binding, there is no reason to impose any further structure on that part of the model.

We establish by construction the existence of a sequential equilibrium in which managers sequentially expropriate all the rent that can be extracted from one-period contracts. Along the equilibrium path, managers work every period, so $h = h'$ for all t. Furthermore, in the model where human capital is public information, whether they have shirked in the past, selected poor job matches, or previously agreed to nonoptimal contracts, at age t with human capital h in position (j, k) having agreed to $w_{jk,t+1}(h, \pi)$, the best response of a manager is to work. While poor decisions may have led them to deviate from the equilibrium path of the original game, the criterion of sequential rationality requires managers and firms to play an equilibrium to the (effective) subgame that begins with the set of state variables they now confront.

When human capital is private information, $h \neq h'$ if the manager is tainted, and the variable pay components, designed for reputation h', do not necessarily align the incentives of shareholders with the manager off the equilibrium path. Having deviated from the equilibrium path by shirking once, it may be optimal for a manager to shirk at some future time, as (19) indicates.

One possibility that the construction of $B_t(h, h')$ does not accommodate is a tainted manager attempting to confess during his negotiations with shareholders. What happens if he offers a contract in the ultimatum game that differs from $w_{jk,t+1}(h', \pi)$, such as $w_{jk,t+1}(h, \pi)$? In the equilibrium we construct, shareholders interpret any deviation from $w_{jk,t+1}(h', \pi)$ as proof the manager is tainted, and therefore a liability to the firm. In particular, it is straightforward (but not instructive) to write down an upper bound for $\underline{F}$ in terms of the model's primitives that ensures no manager has accumulated sufficient wealth to compensate the firm for expected losses that the firm will incur by retaining a tainted manager. This assumption effectively truncates behavior off the equilibrium path because, given those beliefs by shareholders (which we show in our proof of Lemma 5 are consistent), it is a best response of the manager who has optimally selected (j, k) to demand $w_{jk,t+1}(h', \pi)$ and follow the continuation path implied by $B_t(h, h')$.

Lemma 5 *A sequential equilibrium with one-period contracts exists where expected compensation equals the worker's marginal productivity:*

$$E_t[w_{jk,t+1}(h, \pi)] = F_{jkt}(h). \tag{31}$$

5 Estimating the Sorting Equation

Equation (22) implies the certainty-equivalent wage factors into the additive decomposition,

$$w^*_{jk,t+1}(h) = \Delta^{\alpha}_{jkt}(h) + \Delta^{A}_{jkt}(h) + \Delta^{q}_{jkt}(h), \tag{32}$$

where

$$\Delta^{\alpha}_{jkt}(h) \equiv \rho^{-1}\left(b_t - 1\right)^{-1} b_{t+1} \ln \alpha_{jkt}\left(h\right)$$

is the nonpecuniary utility gain or loss incurred by working in (j, k) relative to the outside option that would arise in a Roy model where there is consumption over the life cycle but no human capital accumulation, where

$$\Delta^{A}_{jkt}(h) \equiv \rho^{-1} b_{t+1} \ln\left\{A_{t+1}\left[\overline{H}_{jk}\left(h\right)\right]\right\}$$

is the investment value of (j, k) from accumulating executive experience of the type analyzed in Becker (1964) and Ben-Porath (1967), and where

$$\Delta^{q}_{jkt}(h) \equiv -\rho^{-1}\left(b_t - 1\right)^{-1} b_{t+1} q_{jk}\left[p_t\left(h\right)\right]$$

is a compensating differential of the form analyzed in Rosen (1974). Recalling $q_{jk}\left[p_t\left(h, b_\tau\right)\right]$ is the value of the disturbance $\varepsilon_{jkt} - \varepsilon_{0t}$ that makes the marginal manager in (j, k) indifferent between that position and his outside option, this third term arises in a model of job choice with private values, making the executive in $p_{jkt}\left(h, b_\tau\right)$ fractal indifferent between (j, k) and the outside option. For a given stock of managers aged t with capital h, and a technology determining the nonpecuniary benefits of (j, k), namely $\alpha_{jkt}\left(h\right)$, this compensating differential moves inversely with demand; to raise the equilibrium probability, $p_{jkt}\left(h, b_\tau\right)$, the certainty-equivalent wage must rise to attract managers with lower values of $\varepsilon_{jkt} - \varepsilon_{0t}$ into (j, k).

The difference between the expected compensation and its certainty equivalent is the risk premium

$$\Delta^{r}_{jkt}(h) \equiv E\left[w_{jk,t+1}(h, \pi)\right] - w^*_{jk,t+1}(h) = E\left[r_{jk,t+1}(h, \pi)\right] \tag{33}$$

The risk premium is an agency cost that arises because compensation to risk-adverse executives is subject to uncertainty that depends on the abnormal returns of the firm.

The equilibrium sorting condition (12) exploits the notion that when risk-averse managers make rational choices between different uncertain outcomes or lotteries, they are simultaneously revealing their attitude towards risk, the value of the human capital component of the job match, and its nonpecuniary characteristics (both systematic and temporary). This section describes how we estimated the sorting equation and reports our estimates of the four factors $\Delta^{\alpha}_{jkt}(h)$, $\Delta^{A}_{jkt}(h)$, $\Delta^{q}_{jkt}(h)$, and $\Delta^{r}_{jkt}(h)$.

5.1 Assumptions and Implementation

In our structural estimation, we assume throughout that ε_t is distributed as a Type 1 extreme value. The computational advantages of parameterizing $\varphi\left(\varepsilon\right)$ this way are most evident from the lemmas which follow, where we provide formulas for $A_t\left(h\right)$ and $B_t\left(h\right)$, the value of human capital on and off the equilibrium path, and also an expression for marginal disturbances, $q_{jk}\left[p_t\left(h\right)\right]$, all of which play a central role in the participation constraint (12). However, our estimation approach does not rely on this particular parametric assumption. In principle, alternatives, such as a multivariate normal distribution coupled with some parametric restrictions on $\alpha_{jkt}\left(h\right)$, are available.

Lemma 6 *If ε_{jkt} is independently and identically distributed as an Type I extreme value with location and scale parameters* $(0, 1)$*, then*

$$q_{jk}\left[p_t\left(h\right)\right] = \ln p_{0t}\left(h\right) - \ln p_{jkt}\left(h\right), \tag{34}$$

where $p_{0t}(h)$ is the probability that the optimal choice is retirement and

$$A_t(h) = p_{0t}(h)^{\frac{1}{b_t}} \Gamma\left[(b_t + 1)/b_t\right]. \tag{35}$$

Interpreting the lemma, the IIA property of Type 1 extreme values implies that the marginal idiosyncratic shocks for a manager who is indifferent between his best job match (j,k) and retiring is the log odds ratio of the probability that a manager with characteristics (t,h) who accepts employment in (j,k) versus retire, and does not depend on the other components to the conditional choice-probability vector. The greater the probability of retirement observed in equilibrium, the less important is the human capital component, and the higher is the unobserved shock for the marginal person.

Substituting Equations (34) and (35) into (12) yields, upon simplification,

$$\begin{aligned} z_{jkt}(h) &\equiv \Gamma\left[\frac{b_{t+1}+1}{b_{t+1}}\right]^{-1} p_{0,t+1}\left(\overline{H}_{jk}(h)\right)^{\frac{-1}{b_{t+1}}} \left[\frac{p_{0t}(h)}{p_{jkt}(h)}\right]^{\frac{1}{(b_t-1)}} \\ &= \alpha_{jkt}(h)^{\frac{1}{(b_t-1)}} E_t[\upsilon_{jk,t+1}] \end{aligned} \tag{36}$$

The right side of Equation (36), denoted by $z_{jkt}(h)$, can be approximated by substituting first round estimates of the conditional choice probabilities $p_{0t}(h)$, $p_{jkt}(h)$ and $p_{0,t+1}\left(\overline{H}_{jk}(h)\right)$. Defining the unconditional density of π for the entire sample as $f(\pi)$, and letting x denote a vector of instruments constructed from (h,j,k,t) for each observation,

$$E\left[z_{jkt}(h)x\right] = E\left[\alpha_{jkt}(h)^{\frac{1}{1-b_t}} \exp\left[\frac{-\rho w_{jk,t+1}(\pi,h)}{b_{t+1}}\right] \frac{x f_j(\pi)}{f(\pi)}\right]. \tag{37}$$

Sample analogs were constructed for the conditional choice probabilities, compensation schedule, and conditional and unconditional densities of the abnormal return. The parameter vector formed from ρ and the coefficients defining the certainty equivalent of $\alpha_{jkt}(h)$ is identified from exclusion restrictions. We used bond prices as instruments because b_t affects $z_{jkt}(h)$ and $w_{jkt}(\pi,h)$ but not $\alpha_{jkt}(h)$. The nonpecuniary costs in job match (k,j) depend on whether the executive changed firms, capturing the nonpecuniary cost of switching firms, but we assume the cost is not a function of previous rank and thus use previous rank as an instrument.

We specified $\alpha_{jkt}(h)$ as a log-linear function of age, age squared, tenure, tenure squared, executive's experience, executive's experience squared, number of employers before becoming an executive, number of employers after becoming an executive, and indicators for board membership, interlocked, no college degree, MBA, MS/MA, PhD, and gender. We interacted these 16 variables with rank and firm type to form $\alpha_{jkt}(h)$. We also permitted the risk-aversion parameter to vary by the 36 firm types, but not rank. In total, there are $(16 \times 5 + 1) \times 36 = 2{,}916$ parameters to be estimated. Equation (37) yields an orthogonal condition for each rank and firm combination giving $5 \times 36 = 180$ moment conditions. In addition, to the variables affecting $\alpha_{jkt}(h)$, we used bond prices and the lag of Ranks 1 through 4 as instruments, adding another $5 \times 20 \times 36 = 3{,}600$ moment conditions. After correcting for the preestimation of the conditional choice probabilities, the p-value of the J-statistic for testing the $3{,}780 - 2{,}916 = 864$ overidentifying restrictions is 0.002, implying the model is not rejected at the one percent level of significance.

For presentation purposes only, the results we report in this section impose other restrictions (as indicated in the tables) using a minimum distance estimator, but the unrestricted estimates are carried forward in all subsequent estimations. The results from this estimation are presented in Table 8 to Table 11. The first row in each table embodies the normalization and is measured relative to the benefit of exiting the occupation. The constant is the relative cost for a 50-year-old Rank-5 executive who is choosing to remain with his/her current company, a small firm in the consumer goods sector with few inside executives on its board of directors. The columns are the state variables of the executive at the end of last period. The other rows are measured relative to the first row.

5.2 Compensating Differential for Observed Factors

Table 8 presents the our estimates of $\Delta^{\alpha}_{jkt}(h)$, the compensating differential for working versus retiring.[3] It shows that a 50-year-old, Rank-5 male executive in a small consumer goods company receives an extra $1.6 million compensation for nonpecuniary costs, $263,000 more in Rank 2, $241,000 less in the primary sector, $400,000 more in the service sector, and $553,000 less in large firms. Overall, higher ranked executives receive a higher compensating differential from nonpecuniary cost of working than do lower ranked executives.

The most striking result of Table 8 is that executives prefer large firms to small ones. An executive is willing to accept $373,000 less to work in a medium-size firm compared to a small firm, and $553,000 less to work in a large firm. Thus, the compensating differential declines from $1.63 million for a small firm to $1.07 million for a large firm. The fact that larger firms pay more than small firms is well documented and clearly illustrated by our sample in Tables 2 and 7: this result refutes one contending explanation for the differential, taste.

The results on sector explain most of differences reported in average compensation reported in Table 7, and the sector ordering of compensation is also the same as in Table 2. The model mostly attributes compensating differentials in sectors to working conditions, although the large differences in average compensation between the service sector and the other two depicted in Table 2 are not fully accounted for in either Table 7 or Table 8.

We also find that executives prefer firms with many inside directors; for example, a Rank-5 executive working for a small firm in the consumer sector with a large number of inside directors receives $1.39 million from nonpecuniary cost of working compared to $1.63 million for a similarly placed executive in a firm with a small number of inside directors. At the higher ranks, they give up compensation to be board members; a Rank-5 executive receives an additional $333,000 compensation for being on the board, but the top three ranked executives with at least a year's experience with their firm are willing to forego more than $200,000 to become a board member. Similarly, interlocked executives generally receive a lower compensating differential compared to those who are not; only the lowest ranked executives in medium or large firms demand a (small) positive premium to be interlocked. These three measures of networking opportunities reduce the nonpecuniary costs of a job match, and hence its equilibrium compensation.

Female executives receive a higher differential than men to accept Rank-1 and -2 jobs in the consumer sector, $176,000 and $304,000 respectively, plus an additional $100,000 for primary- and service-sector jobs. At the average age, tenure, and executive experience, female executives receive $1.6 million overall, as compared to $1.5 million for men, to offset nonpecuniary utility losses from continuing to work one more year. Of all the education groups, executives with a PhD receive the highest compensating differential for nonpecuniary losses from working versus retiring, $1.52 million averaged overall, while those with MBA degrees receive the lowest, $1.41 million. The pattern we observe for education and gender may reflect superior outside options, in other labor markets and retirement, for female executives and executives with a PhD.

The differential increases with age, tenure, number of moves after becoming an executive, and board membership. Executives moving to a new employer receive $380,000 less compensation for nonpecuniary losses, but one third of this is wiped out if they are placed on the board in their first

[3] The standard errors were obtained using the multistep procedure given by Newey and McFadden (1994).

year. On average, an executive in his/her first year with the firm receives \$1.16 million as compared to \$1.54 million in his/her second year to offset nonpecuniary losses. This suggests that part of the reason why managers turn over is to take job matches with more attractive nonpecuniary benefits.

5.3 Compensating Differential for Unobserved Factors

Table 9 reports our estimates of $\Delta^{q}_{jkt}(h)$, which measures how much extra managers matching with a given personal background (t, h) must be paid to attract the proportion we observed in the data selecting job match (j, k). It shows a marginal Rank-5 executive in a small consumer goods company gives up \$569,000 for the unobserved idiosyncratic component difference relative to exit. Rank 4 has the highest net demand, whereas Rank 1 has the lowest net demand, with employers offering a negative differential of \$151,000 to Rank-1 executives and a positive differential of \$181,000 to Rank-4 executives. Firms in the primary sector pay an additional \$48,000, while large firms pay an extra \$170,000 to meet demand. Comparing the ordering by rank in the top entries of the first column, with the ordering of the estimated unconditional exit probabilities by rank implied by the first column of Table 2, we see that the effect of conditioning on the observed variables affects the ordering of the conditional choice variables by rank. For example, compensation in Rank 4 is most boosted by the unobservables factors; in our model, it attracts managers who receive relatively low values of $\varepsilon_{14t} - \varepsilon_{0t}$ (as reflected in the low cutoff value $q_{j3t}(h)$); yet Rank 2 is more likely to be selected than any other rank, including exiting. This demonstrates that the observed variables in our model explain why managers are deterred from taking the post of CEO.

Table 9 shows that larger firms have higher demand for executives, which partially explains the positive relationship between pay and firm size. It also shows that when there are a large number of inside executives on the board of directors, employers are able to pay \$117,000 less in compensation in order to meet demand, and restrict those positions to executives who have drawn more favorable idiosyncratic shocks for those matches. There is greater net demand for high-ranked executives to be on the board of directors. Low-ranked executives sacrifice \$320,000 to be on the board (even more if they have just joined the firm), but higher ranked executive board members command a premium of over \$100,000. Only at low ranks in the consumer and service sectors is there greater net demand for women relative to men. Finally our finding, that the marginal executive takes a discount of \$85,000 in compensation to switch firms, shows that executives only switch firms when they receive a relatively favorable idiosyncratic shock from their new employer job match.

5.4 Compensating Differential from Human Capital

Leaving aside career concerns, the value of human capital to managers offsets their equilibrium compensation by $\Delta^{A}_{jkt}(h)$. With reference to Table 10, we find that human capital investment is important for executives at all ranks. They would demand an extra \$200,000 to \$300,000 in compensation per annum, but for the benefits of experience on the job. As fraction of their certainty-equivalent wage, $w^{*}_{jk,t+1}(h)$, the value of human capital is bracketed between approximately one quarter and one half of total compensation, remarkably high given the distribution of ages, positions, and the lengths of future careers.

The value of human capital investment is concave in rank, peaking in Rank 2 and then dropping off sharply, its value in Rank 1 falling below that in Rank 5. Within our model, the formula for $A_t(h)$ in Equation (35) shows that the investment value of human capital is inversely related to the probability of exit. So it is not surprising to see that the relationship between human capital and rank shown in Table 10 is exactly the opposite to the plot of relative frequencies of exit by rank implied by Table 3.

Predictably, the value of human capital investment declines with age and all types of experience;

similarly moving to a new firm increases the value of human capital investment. Reflecting their higher exit rate, female executives place a lower value on human capital investment. A female executive is willing to give up $200,000 because of the human capital investment, whereas men are willing to forego $300,000.

5.5 Certainty-Equivalent Wage

The certainty-equivalent wage is obtained by appealing to Equation (32) and summing the estimates of the components reported in Tables 8, 9, and 10. We find that the certainty-equivalent wage is concave over rank, lowest in Rank 5, $570,000, increasing monotonically to $900,000 in Rank 2, before declining to $690,000 in Rank 1. It is instructive to note that Rank-3 executives have a higher certainty-equivalent compensation, $730,000, than Rank-1 executives, but that Rank-1 executives have a slightly higher certainty-equivalent compensation than Rank-4 executives, $660,000. This ordering follows that of average total compensation by executive rank reported in Table 3, which ranges from $1,269,000 (for Rank 5) to $4,794,000 (for Rank 2). However, the compression due to the risk premium, which we discuss in more detail below, is noteworthy.

Another striking result is that the certainty-equivalent wage decreases with firm size. The average certainty-equivalent wage of an executive in a small firm is $780,000, falling to $430,000 for a medium-size firm, and to $390,000 for a large firm. This is because the overall negative effect of firm size on the nonpecuniary loss from working versus retiring outweighs the positive effect of firm size on net demand. This result contrasts with the positive relationship between expected compensation and firm size as documented in the literature and illustrated in Table 2. However, Table 2 also shows a positive relationship between firm size and the variance of compensation. In principle, the higher variability of compensation in large firms could be due to volatility in abnormal returns that figure into compensation packages and are accounted for by the risk premium, or to other forms of heterogeneity, both observed and unobserved. We find that the risk premium is key to explaining the sign reversal in the slope of expected compensation and the slope of certainty-equivalent wage with respect to firm size.

The average certainty-equivalent wage of executives working for firms with a large number of inside executives on the board of directors is only $390,000, versus $730,000 for those working in firms with only a small number. An executive who is interlocked receives a certainty-equivalent wage of $560,000, less than those who are not, $710,000. Again, in both cases, the negative effect of nonpecuniary losses from working versus retiring outweighs the positive effect of net demand.

Turning to different demographic groups, female executives on average receive a higher certainty-equivalent wage than men, although they have a lower net demand (which is more than offset by compensation for nonpecuniary losses on the job). An executive with a PhD receives a certainty-equivalent wage of $720,000, an MBA degree holder $600,000, and those without either $680,000. One possible explanation is that demographic groups with education specifically geared to business (such as those with MBA degrees) and less experience in the nonmarket sector (such as men) cannot extract as much rent from firms because their options outside the executive market are more limited. Finally, those in their first year only receive on average $180,000 in certainty-equivalent compensation, compared to a certainty-equivalent wage of $770,000 for those who have (exactly) one year's tenure, starkly illustrating the gains from survival beyond one year. With respect to this last result, all three factors in the decomposition point in the same direction; investment in human capital plays a bigger role, the nonpecuniary losses are lower and unobserved factors affecting the job are more beneficial for a manager in his first year.

5.6 Risk Aversion and the Risk Premium

We initially specified the risk-aversion parameter as function of gender, but at the one percent level could not reject the null hypothesis that male and female executives have the same coefficient of risk aversion. Our estimate of the risk-aversion parameter (for all groups) is 0.534 with a standard error of 0.152, for compensation measured in millions of 2006 US$. For example, a manager with risk-aversion parameter of 0.534 would be willing to pay $255,199 to avoid a gamble that has an equal probability of losing or winning one million dollars. This is similar to results in Gayle and Miller (2009), who found a risk-aversion parameter of 0.501 using data on 37 firms for the period 1944–1978 and 0.519 using data on 151 firms for the period 1993–2004. They estimated a fully parametric model without job choices or human capital using a much smaller data set; they also deployed a nested, fixed-point, full-solution estimation technique instead of only exploiting the equilibrium sorting conditions to identify and estimate the risk attitude. Yet the results are strikingly similar.

Table 11 displays our estimates of $\Delta^{r}_{jkt}(h)$, showing that at Ranks 4 and 5 the cost is small and insignificant in small firms, but adjusts to $1.5 million, $3.3 million and $1 million for Ranks 3, 2, and 1. Roughly 82 percent of the compensation of a CEO (Rank 2), versus 72 percent for Rank 1, 76 percent for Rank 3, 65 percent for Rank 4, and 69 percent for Rank 5, is due to the risk premium. The service sector pays a higher risk premium than the other two, a factor which helps close the gap between the considerably higher levels of average compensation paid in that sector and those reported in Table 2. Our finding of a higher risk premium in the service sector is anticipated in the summary statistics and reduced-form analysis. Table 2 shows that there is greater variability in compensation from all sources in that sector (including heterogeneity amongst executives, which is not a source of risk), and Table 7 shows that the specific source of variability determining the risk premium in our model, namely abnormal returns, is also highest in that sector.

The risk premium increases significantly with firm size. On average an executive in a small firm receives $1.6 million in risk premium (56 percent of expected compensation), in a medium-size firm $2.8 million (85 percent of expected compensation), and in a large firm $4.8 million (90 percent of expected compensation). These results are a further demonstration that the positive relationship between expected compensation and firm size is fully accounted for by the positive relationship between the size of the risk premium paid to managers and the size of their employer firms.

Firms with a many executives on their board of directors pay a higher risk premium, which was foreshadowed in Table 2 (from the standard deviation on compensation) and Table 7 (from the dependence of compensation on abnormal returns). An interlocked executive, however, receives a lower risk premium, $1.9 million, but a higher percentage of expected compensation, 77 percent, than an executive who is not interlocked. This is because interlocked executives receive a lower certainty-equivalent wage, $560,000, than a noninterlocked executive, $710,000.

Female executives receive a lower risk premium, $2.1 million, than men, $2.2 million, equalizing expected compensation, $2.9 million across gender. In our framework, expected compensation is the executive's marginal product: Thus, we find female and male executives are equally productive. By the same token, executives with a PhD, who receive an average expected compensation of $3.0 million are more productive than those with an MBA, $2.7 million, and without either, $2.8 million. An executive with a PhD receives a higher risk premium, $2.3 million, than one with an MBA, $2.1 million, but an executive with an MBA has a higher fraction of expected compensation, 78 percent, than one with PhD, 76 percent, as risk premium. There is a $362,000 spike in the risk premium for new executives, but it declines $65,000 with each extra year of tenure and age. Consequently the lower certainty-equivalent wage offered to first-year executives is partially hidden by data on their average compensation.

6 Estimating the Costs and Benefits of Shirking

The risk premium arises whenever career concerns are insufficient to compensate the manager for shirking when he is paid his certainty-equivalent wage. In this case, we can estimate the gross loss to the firm from shirking, and also the net gain from shirking to the manager. The difference between expected return to the firm from a manager working versus shirking when all other executives work is

$$\Delta^{g}_{jkt}(h) \equiv E\left[\pi - \pi g_{jk}(\pi, h)\right]. \tag{38}$$

The estimated gross expected loss is obtained from first estimating $g_{jkt}(\pi, h)$, the likelihood ratio of working versus shirking, the importance of which can be measured by the gross output loss to the firm from switching from $f_j(\pi)$, the density of abnormal returns obtained from working, to its shirking counterpart, $f_j(\pi)\, g_{jkt}(\pi|h)$. We estimate this loss from the slope of the compensation schedule and the volatility of compensation under different aggregate conditions captured by bond prices, and do not exploit assumptions about the agency problem that motivates the contract form, optimal or not.

The net benefit from shirking to the manager is given by two factors. One is the compensating differential for current utility when the executive weighs shirking against working:

$$\Delta^{\beta}_{jkt}(h) \equiv \frac{b_{t+1}}{\rho\,(b_t - 1)}\left[\ln \alpha_{jkt}(h) - \ln \beta_{jkt}(h)\right]. \tag{39}$$

This factor measures the misalignment of incentives from the manager's perspective. The other is the difference in the conditional valuation functions (continuation values), from working in the current period t, versus shirking (where in both cases optimal decisions are taken thereafter):

$$\Delta^{B}_{jkt}(h) \equiv \frac{b_{t+1}}{\rho}\left\{\ln A_{t+1}\left[\overline{H}_{jk}(h)\right] - \ln B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]\right\}. \tag{40}$$

This measures the amelioration of the agency problem due to the executive's career concerns.

6.1 Gross Loss to Firms from Shirking

To estimate the decline in gross returns firms that would incur from a manager shirking for one period, we first estimate $g_{jkt}(\pi|h)$, by appealing to the following lemma in Gayle and Miller (2011), who show that up to two normalizing constants, $g_{jkt}(\pi|h)$ is identified from the relative slope of the wage compensation schedule. The two normalizing constants come from the fact the expected value of $g_{jkt}(\pi|h)$ is one, and our assumption that as π increases without bound, the limit of $g_{jkt}(\pi|h)$ is zero. The assumption implies compensation is bound by a limit defined from Equations (29) and (30):

$$\overline{w}_{jk,t+1}(h) \equiv \lim_{\pi\to\infty} w_{jk,t+1}(h, \pi) = w^{*}_{jk,t+1}(h) + \overline{r}_{jk,t+1}(h).$$

The lemma is proved by manipulating the first-order conditions characterizing the minimum-cost contract.

Lemma 7 *Setting* $\overline{w}_{jk,t+1}(h) \equiv w^{*}_{jk,t+1}(h) + \overline{r}_{jk,t+1}(h)$*, for all* (j, k, t, h)*,*

$$g_{jkt}(\pi|h) = \frac{\exp(\rho\overline{w}_{jk,t+1}(h)/b_{t+1}) - \exp(\rho w_{jk,t+1}(\pi, h)/b_{t+1})}{\exp(\rho\overline{w}_{jk,t+1}(h)/b_{t+1}) - E_t[\exp(\rho w_{jk,t+1}(\pi, h)/b_{t+1})]}. \tag{41}$$

We formed $\widehat{w}(h_t, \pi)$, the nonparametric estimates of the compensation schedule as a polynomial expansion, using them in conjunction with our estimate of the risk-aversion parameter obtain earlier. We approximated the conditional expectation $E_t[\exp(-\widehat{\rho}\widehat{w}(h_t, \pi)/b_{\tau(t)+1}]$ by integration using the nonparametrically estimated density of π for a given j, and computed $\overline{w}_{jk,t+1}(h)$ using the maximum $\widehat{w}(h_t, \pi)$ for each value of (j, k, t, h). Finally our estimate of $g_{jkt}(\pi|h)$ was obtained by substituting our estimates of $\overline{w}_{jk,t+1}(h)$, ρ, and $E_t[v_{jk,t+1}(\widetilde{\rho}, \pi)]$ into Equation (41).

Table 12 reports our estimates of $\Delta^g_{jkt}(h)$, showing that small consumer sector firms lose 33.6 percent of their equity value when a Rank-5 executive shirks, but large firms lose much less, eight percent. Intuitively, shirking executives in small firms cause significantly more damage than they would in large firms, because an executive in a smaller firm has a greater marginal impact on each unit of equity than any one executive working for a large firm. However, there is a positive relationship between firm size and the expected gross loss in equity from shirking. Multiplying our estimates in Table 12 by the average equity value in Table 2 gives gross equity losses of \$102 million for a small firm, \$203 million for a medium one, and \$393 million for a large one. These numbers are comparable to those found in Gayle and Miller (2009), whose estimates range between \$160 million and \$230 million for a similar time period, smaller firm size, and more restricted sample. Since the gross loss in equity value from shirking would be higher in large firms, it is not surprising they pay a greater risk premium to motivate their management, as documented in Table 11.

Gross loss monotonically declines in rank; thus, when a Rank-1 executive in a large firm shirks, only a small proportion of equity value is lost. Similarly the extent of destruction is lower for higher lagged ranks. These findings contradict conventional wisdom that shareholders risk more from Chairmen and CEOs who shirk than lower ranked officers; our results are consistent with the view that executives closer to the firm's operations can wreak the most havoc. The losses are greatest in the service sector, least in the primary sector. Executive directors and interlocked executives would be less destructive if they were not motivated (perhaps because these extra duties are associated with greater monitoring), and the losses are smaller if there are many insiders on the board, possibly for similar reasons.

There is no significant difference in gross loss by gender. It is, however, noteworthy that shirking executives with doctorates would have a greater negative impact than all other educational groups; we speculate specialized knowledge could be brought to bear when pursuing their own objectives if their incentives are not aligned with shareholder wealth. Similarly, an executive in his first year with the company (perhaps with more contacts outside the firm) can cause significantly more damage than more tenured executives. On average, the gross loss to the firm from not motivating an executive in his/her first year is 34 percent, falling to 25 percent by the second year.

6.2 Net Benefit of Shirking to Executives

To estimate the net benefits of shirking to an executive, we now define a parameter $\beta^*_{jkt}(h)$ that captures the virtual preference for shirking:

$$\beta^*_{jkt}(h) \equiv \beta_{jkt}(h) \left\{ \frac{B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]}{A_{t+1}\left[\overline{H}_{jk}(h)\right]} \right\}^{1\{private\}(b_t - 1)}. \tag{42}$$

Thus, when human capital is public information $\beta^*_{jkt}(h) \equiv \beta_{jkt}(h)$, but when human capital is private information $\beta^*_{jkt}(h) \neq \beta_{jkt}(h)$ because $A_{t+1}\left[\overline{H}_{jk}(h)\right] \neq B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]$. The net benefit of shirking to the manager, the sum of (39) and (40), can now be simply expressed as

$$\Delta^{\beta^*}_{jkt}(h) \equiv \Delta^{\beta}_{jkt}(h) + \Delta^{B}_{jkt}(h) = \frac{b_{t+1}}{\rho(b_t - 1)}\left[\ln \alpha_{jkt}(h) - \ln \beta^*_{jkt}(h)\right]. \tag{43}$$

The following lemma implies estimates of this value can be obtained from the estimates of the sorting equation.

Lemma 8 *For all (j, k, t, h),*

$$\beta^*_{jkt}(h) \equiv \frac{p_{0t}(h)}{p_{jkt}(h)} A_{t+1} \left[\overline{H}_{jk}(h)\right]^{1-b_t} \left\{ \frac{E_t[v_{jk,t+1}] - \overline{v}_{jk,t+1}}{1 - \overline{v}_{jk,t+1} E_t[v^{-1}_{jk,t+1}]} \right\}^{1-b_t}. \tag{44}$$

Table 13 reports our estimates of $\Delta^{\beta^*}_{jkt}(h)$, the value an executive would place on shirking if he were paid the certainty-equivalent wage. It is about \$10 million for a 50-year-old Rank-5 executive in a small firm in the consumer goods sector. The differentials across rank are not significant, suggesting that the benefits of shirking do not depend on rank.

However, the differential declines in firm size, by \$3.1 and \$4.5 million for medium and large firms, respectively, and differs across sectors, \$3.8 million higher in the service sector than the consumer sector, and \$2.6 lower in the primary sector. The differential also declines with measures of networking. For interlocked executives, it falls by \$930,000 in small firms in the consumer sector, a further \$616,000 in the service sector, although these differences are less pronounced in other firm types. Likewise when the board is dominated by insiders, it falls by \$2.2 million.

The differential declines by a further \$4.8 million for executives just joining the firm. Evidently new hires are less familiar with opportunities to exploit the firm's resources for personal gain. Curiously, the female differential is negative for new firms, although experienced female executives have a higher differential at all ranks except Rank 3. We speculate their minority gender status initially serves as a barrier to benefiting from shirking as much as their male colleagues. Arguing along these lines, we are, therefore, not surprised to find that the differential increases with age, tenure, and executive experience, although admittedly these effects are mitigated by firm size.

6.3 Differentiating Shirking Preferences from the Value of Human Capital

When human capital is private information, its law of motion is given by Equation (3) and the incentive-compatibility constraint (25) applies.

Lemma 9 *When human capital is private information, for all (j, k, t, h).*

$$\beta_{jkt}(h) \equiv \frac{p_{0t}(h)}{p_{jkt}(h)} B_{t+1} \left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]^{1-b_t} \left\{ \frac{E_t[v_{jk,t+1}] - \overline{v}^{-1}_{jk,t+1}}{1 - \overline{v}_{jk,t+1} E_t[v^{-1}_{jk,t+1}]} \right\}^{1-b_t}, \tag{45}$$

where

$$B_t(h, h') = \frac{\Gamma\left(\frac{b_t+1}{b_t}\right)\left[1 + \sum_{j=1}^{J}\sum_{k=1}^{K} V'_{jkt}(h, h')^{-\frac{1}{b_t}}\right]}{\left\{1 + \sum_{j=1}^{J}\sum_{k=1}^{K}\left[V'_{jkt}(h, h')\right]^{-1}\right\}^{1+\frac{1}{b_t}}}. \tag{46}$$

Up until this point, we have not placed any restrictions on $\underline{H}_{jk}(h)$. To separate the effects of career concerns from the current benefits of shirking, we must make an assumption on the functional form of $\underline{H}_{jk}(h)$, how is experience updated when the manager shirks. In our empirical specification, we assume that the first time a manager shirks, an indicator variable called tainted is activated, remaining that way forever, and that no experience variable other than age changes in periods the manager shirks. We further assume that the indicator variable for being tainted only affects $F_{jkt}(h)$, while the other experience variables only affect $g_{jkt}(\pi|h)$.

Estimates of $\beta_{jkt}(h)$ and $B_t(h,h')$ can then be obtained recursively. Noting that $B_{T+1}(h,h') \equiv 1$ and substituting our estimated risk-aversion parameter and conditional choice probabilities into Equation (45) yields $\beta_{jkT}(h)$. Substituting $\beta_{jkT}(h)$ into (20) yields $V'_{jkT}(h,h')$ and hence $B_T(h,h')$, using (46). More generally, given $B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]$, the parameter $\beta_{jkt}(h)$ is obtained from (45), and hence estimates of $V'_{jkt}(h,h')$ and $B_t(h,h')$ are produced from (20) and (46), respectively.

Table 14 reports our estimates of how much the agency problem is mitigated by career concerns given our assumption on $\underline{H}_{jk}(h_t)$.[4] The table shows that the data is consistent with a hidden-information model of human capital in which there are significant career concerns at all ranks. Career concerns reduce the differential for diligent work versus shirking under perfect monitoring by between 15 percent and 22 percent; as a percentage of the gross compensating differential, it is lowest in Rank 1 and highest in Rank 3. The lower percentage in Rank 1 reflects its position at the end of the life cycle, while the higher percentage in Rank 3 reflects the eminent possibility of promotion to CEO. There are significant career concerns at the CEO rank, 19 percent of the gross compensating differential from working versus shirking, the same as in Rank 4 and higher than in Rank 5 (17 percent). We conclude that in order to maintain the hard charging life demanded of executives in our population, pursuing the goal of value maximization teaches executives to sell themselves more effectively to shareholders.

The role of career concerns declines with age, tenure, executive experience, and experience in different firms. Interlocked executives places lower value on career concerns, executive directors higher. Generally, female executives also place less weight on career concerns than men, with the notable exceptions of those in Rank 2 and those joining new firms, where they have more. Executives with no college degree, an MBA, or a PhD have greater career concerns than executives with only a bachelor degree, who in turn have greater career concerns than those holding an MS/MA.

7 Nonparametric Identification

Finally, we briefly investigated how sensitive inference about our agency model is to underlying parametric assumptions. The purpose of this exercise is not only to test our quantitative findings, but also to provide a more general commentary on the limitations of inference from large longitudinal panel data sets, such as ours, on (only) compensation, firm returns, and executive careers. We demonstrate that the decomposition of gross benefit from shirking into current nonpecuniary utility and career concerns is not identified unless $\underline{H}_{jk}(h)$ is known. Given $\underline{H}_{jk}(h)$, we then prove that our framework is identified for data on job matches, compensation, and firm returns (alone), up to the probability density, $\varphi(\varepsilon)$, and the risk parameter, ρ. This implies that at least one exclusion restriction is required to estimate ρ, and additional restrictions would be necessary to estimate $\varphi(\varepsilon)$ as a function of an unknown parameter vector.

Imagine the data is generated by a model where human capital is private information. Substituting the virtual parameter $\beta^*_{jkt}(h)$ defined in (42) into (70), the incentive-compatibility constraint for the private information model, gives the incentive-compatibility constraint for the public-information model, Inequality (66) with $\beta^*_{jkt}(h)$ replacing $\beta_{jkt}(h)$. Moreover the participation constraint is not affected by these alterations to the model because the manager always works in the equilibrium, solving (11) and implying (12) does not depend on $\beta_{jkt}(h)$ or the information structure. Consequently the solution to the optimal contract problem given by Equations (27), (28), and (30) for the private-information model is obtained by replacing $\beta_{jkt}(h)$ with $\beta^*_{jkt}(h)$. Therefore, a model with private

[4] Subtracting the estimates in Table 14 from those in Table 13, we obtain the gross compensating differential for diligent work versus shirking under perfect monitoring. The estimates in Tables 13 are mainly of a higher order of magnitude than those in Table 14. Therefore, the qualitative patterns of the gross compensating differential for diligent versus shirking is similar to the net differential.

information about human capital with shirking preference $\beta_{jkt}(h)$ is observationally equivalent to a model with public information about human capital in which the shirking parameter is $\beta^*_{jkt}(h)$ but is the same in all other respects. More generally $\beta^*_{jkt}(h)$ indexes observationally equivalent models that differ only in their specification of $\underline{H}_{jk}(h)$ and $\beta_{jkt}(h)$.

This identification result indicates the scope for inference off the equilibrium path in the optimal spot-contract models we investigate and illustrates the difficulty in differentiating long-term optimal contracts from nonoptimal but sequentially rational spot contracts. Pairing the former with a model of public information about human capital is observationally equivalent to pairing the latter with a model of private information of human capital. To summarize, testing for the optimality of contract type is confounded by simultaneously trying to identify the information technology.

Given the first result on $\underline{H}_{jk}(h)$, no further generality is lost by restricting the analysis to the public-information case with a virtual shirking parameter $\beta^*_{jkt}(h)$. Accordingly, let Θ denote the class of models under consideration, consisting of elements

$$\theta \equiv \left(\alpha_{jkt}(h), \beta^*_{jkt}(h), \rho, f(\pi), g_{jkt}(\pi | h), \varphi(\varepsilon)\right).$$

To preserve the notational conventions already established, we prove this result for the case where bond prices are constant over time. A more general result is available when b_τ varies over time providing the other parameters are also allowed to vary with calendar time. We assume $b_\tau = b$ for all τ and suppose the data, comprising histories of job matches, abnormal returns, and compensation, summarized by (π_t, d_t, w, h) is generated by $\widetilde{\theta}$. Our final lemma states that for any parameterization of the density, $\varphi(\varepsilon)$, and any risk-aversion parameter, ρ, there is a unique model with public information that has the same generating process, a result which resonates with previous work by Magnac and Thesmar (2002) on the identification of discrete-choice models and by Gayle and Miller (2009) on the identification of moral-hazard models.

Lemma 10 *Suppose $b_t = b$ for all t and (π_t, d_t, w, h) is generated by $\widetilde{\theta}$. For every $\widehat{\rho} > 0$ and all proper probability density functions $\widehat{\varphi}(\varepsilon)$ defined on the same support as $\widetilde{\varphi}(\varepsilon)$, there exists a unique $\widehat{\theta}$ solving the equations in Lemmas 1 through 5 plus 7 and 8 that is observationally equivalent to $\widetilde{\theta}$.*

8 Conclusion

To paraphrase Rosen's (1990) eloquent summary statement, the executive labor market performs three functions: assigning control amongst members of the management team; providing performance incentives that align the interests of individual managers with their employer firms; and attracting, nurturing and retiring executives to facilitate the transfer of control from one generation of managers to the next. Our paper makes an empirical contribution to the research agenda of understanding these three functions. We embed a simple dynamic model of career choices within a principal–agent contracting framework, where the primitives define managerial preferences, production technology and information, and estimate its parameters from equilibrium conditions.

The assignment function has both internal and an external aspects. The external assignment function emphasized by Lucas (1978) determines the size and distribution of firms. In support of Aron's (1988) model, and confirming results in Gayle and Miller (2009), we find that higher agency costs associated with large firms is the main reason why executive compensation increases with firm size. However, our model cannot explain why further amalgamation does not take place given observed prices. Our results suggest that the agency costs are concave increasing in firm size; for example, the estimated risk premium is \$1.6 million for small firms, \$2.6 million for medium-size firms, and \$4.9 million for large firms, while from Table 2 the average equity value is \$322 million for small firms, \$1071 million for medium-size firms, and \$6,022 million for large firms. Moreover, the value of investment in human capital does not vary appreciably with firm size, and the larger the firm, the greater the nonpecuniary benefits (Table 9). Thus, human capital acquisition, the nonpecuniary costs of work, and agency costs are not the only factors determining the distribution of firm size.

An important issue relating to the internal assignment function, first analyzed in Williamson (1967), Alchian and Demsetz (1972), and Mirrlees (1976), is whether a firm is more like a decentralized contractual organization, or a centralized command dictatorship. How to distribute influence within an organization depends on the interplay between human capital and the power to control. Supposing these two factor inputs are complements, it is efficient to assign greater control to executives with more human capital, and in equilibrium the executives with the most human capital would occupy the top positions in the largest firms, where the marginal productivity of their actions is magnified by the greater resources they manage. In this spirit, Rosen (1982), Gabaix and Landier (2008), and Terviö (2008) develop models to explain why executive pay is positively correlated with firm size, and why top executives are so well paid.

Our construction of the managerial hierarchy, defined axiomatically and constructed in Gayle, Golan, and Miller (2011), is based on life-cycle transitions, not pay. Nonetheless, we find that higher ranked positions command higher expected compensation, the value of the marginal product of executive labor in our framework. From Table 7, we also find support for the hypothesis that larger firms have a higher demand for more qualified executives. Note though, that the firm-size effects on compensation are quantitatively smaller than the agency effects (in Table 11), largely because executives prefer working in large firms to small (Table 8), all else equal. Even more poignant, executive power or their span of control, measured in our model by the gross loss to shareholders from shirking, declines significantly with rank (Table 12), whereas productivity, as measured by expected compensation (and displayed in Table 7), increases up to Rank 2. It seems to us that firms more closely resemble multilateral contractual obligations between self-interested parties, rather than chains of command bound by loyalty, coercion, and firm-specific capital. Other features of our estimates support this contractarian interpretation: compensation falls with tenure (Table 7) and nonpecuniary costs rise with tenure (Table 8).

The second function concerns incentives. Our results support the prediction of agency theory that pay structure should strike a balance between incentives and insurance. Confirming previous empirical findings by Margiotta and Miller (2000) and Gayle and Miller (2009), we find that the risk premium accounts for more than half of total expected compensation in our population of top executives. As Rosen (1990) predicted, we also find that career incentives serve as a better substitute for current performance incentives at earlier stages in a career because the diminishing horizon reduces the value of human capital investment; in Table 13, the divergence between the shareholder and executive goals increases with all measures of experience.

Whether the promotional structure and compensation incentives make for good governance has recently drawn the attention of several empirical researchers, namely Rose and Shepard (1997), Hallock (1997), Bebchuk and Fried (2004), Kuhnen and Zwiebel (2009), Acharya and Volpin (2010), and Dicks (2010). Our study makes a theoretical and an empirical contribution on this topic. Section 7 demonstrates that for our framework and data-generating process, an optimal long-term equilibrium contract with public information about human capital cannot be identified from a sequence of inefficient spot contracts with private information about human capital. Typically, empiricists cannot tell whether human capital is private or public information; our negative result limits the scope for testing whether contracts (and more informal arrangements within the firm) are optimal or not.

Empirically, we cannot reject the hypothesis that executives in companies with a large number of insiders on the board receive the same expected compensation as other executives (Table 7). In our model, every executive has an incentive to work. Placing more of them on the board to monitor each other mitigates gross losses to the firm should any one of them shirk (Table 12), reduces the net benefits from shirking (Table 13), and increases the gross value of the firm from greater coordination (reflected in the firm's equity value and thus impounded into its financial returns). But greater executive representation on the board does more than create a more challenging signalling problem to solve, thereby raising the risk premium (Table 11); giving more votes to executives fosters better executive working conditions (Table 8), which in turn is offset by a lower certainty-equivalent wage in equilibrium (as reported in Section 5.5). Thus, our estimates undergird a plausible explanation of how large shareholders determine the number of insiders on the board to maximize the expected value of their equity.

The third function, propagating managerial talent, also has several facets, namely how executives are selected, human capital they accumulate on the job, the role of turnover, and the role of exit from the occupation. The literature on the returns to experience and job turnover has emphasized the trade-off between losing firm-specific human capital when leaving a job match and the gains of improving a job match that evolve with the worker's experience (see Miller 1984; Dustman and Meghir 2001; Altonji and Shakotko 1987; Topel 1991; Topel and Ward 1992; Neal 1995 among others). Three measures of human capital emerge from our estimates of the conditional choice probabilities and compensation regressions: years of tenure with the current employer firm; years accumulated in the market for executives; and the number of employers an executive has had. From Table 7, we find that age and executive experience increase productivity but that expected compensation falls with tenure. Likewise, switching firms increases productivity in Ranks 2 and 3, and productivity is increasing in the number of firms the manager has worked after joining the executive ranks. We also find that the nonpecuniary costs of work are significantly but only temporarily lower when an executive joins a new firm (see Table 8). Job matching theory predicts that the young and inexperienced should experiment with different types of work; we find that even late in the career cycle, variety in job experience adds to human capital.

Becker's (1964) and Ben-Porath's (1967) life-cycle theory of human capital predicts that as executives age, human capital investment becomes less important. In support of the theory, Table 3 shows higher ranks are held by older executives with more executive experience, and from Tables 10 and 14, the value of human capital investment decreases with all measures of experience. However, Table 10 also shows that executives give up more compensation for human capital investment as they progress through the ranks right up until Rank 1, where the trend falls off. This pattern reflects the exit probability, which from Table 4 is lowest in Rank 2 and highest in Rank 1. Intuitively, the effective discount factor used to compute the value of human capital, in terms of summed future increased earnings within the occupation, must account for the probability of exit. Consequently, standard models of human capital where everybody retires at the same rank would overpredict human capital investment in the lower ranks and underpredict the level of investment in higher ranks.

Our data comes from the truncated population of those workers who reach upper management in a publicly listed company, so we cannot infer much about the lengthy incubation phase that characterizes executive selection. However, we do investigate how the career profiles of executives vary with backgrounds, paying special attention to gender and educational attainment. Our empirical results show that, after controlling for other observed characteristics including rank, women are paid the same expected compensation as their male counterparts. Table 8 shows that women are more likely to quit because of greater opportunities from exiting relative to the nonpecuniary characteristics of work. They value investment in human capital less than men, there is lower net demand for their services, they receive higher certainty-equivalent compensation, and would reap smaller net benefits from shirking. These results confirm and expand upon findings in Bertrand and Hallock (2001), Bell (2005), Albanesi and Olivetti (2008), Selody (2010), and Gayle, Golan, and Miller (2011). Our framework shows that the gender differential in the nonpecuniary benefit ratio of executive work to exit creates its own dynamic, reflected in human capital accumulation and career movement within the executive sector: The small minority of women in executive management are behaving like discouraged workers, even though we cannot reject the joint hypothesis that there is no gender discrimination within this employment sector and women have better outside options than men.

As Arcidiacono, Cooley, and Hussey (2008) note, the return to an MBA degree is usually contaminated by the benefits of previous work experience, a requirement of many MBA programs. Our study does, however, shed light on the long-term benefits of a general business education versus a more specialized degree. We find MBA degree holders have a lower marginal productivity than graduates with a PhD or another specialized degree. An executive with a PhD has a higher nonexecutive market outside option relative to the nonpecuniary benefits of executive work (Table 8) and a higher certainty-equivalent compensation than an executive with an MBA. This implies that the MBA graduate has more implicit incentives and hence requires less explicit incentives and current compensation, which translates to a higher value of human capital investment (Table 10) and greater career concerns (Table 14). There seems to be a higher net demand for executives with an MBA (Table 9), while PhD graduates would destroy significantly more of the firm's value if they shirked (Table 12), which we attribute to their specialized knowledge and intellectual prowess.

The three functions of the executive labor market are intertwined: the assignment functions are inherently related to the nurturing functions through working experience in management; the incentive issues are task specific because of role information plays in their determination. We dissect the complexities of this problem by imposing structure only as dictated by identification and sample-size considerations. First, we display the conditional choice probabilities and executive-compensation regressions. Then we estimate an equilibrium sorting equation from a dynamic individual optimization model facing a compensation schedule to obtain a decomposition of the certainty-equivalent wage and the risk premium. Only at this point do we impose further structure, derived from the optimality conditions for the equilibrium contract, to obtain the virtual preferences for shirking off the equilibrium path. Finally, lacking identification, we appeal to introspection and the estimated parameters of the model to calibrate the extent of career concerns. In this way, our estimation framework exploits the hallmarks of structural estimation, internal rigor and clarity of interpretation, while simultaneously mining the rich veins of systematic variation exhibited in the large longitudinal data set we have compiled from multiple sources on job matches between executives and their employer firms.

9 Appendix

Proof of Lemma 1. We proceed by induction, first showing that the expression for the value function is true for age T, and then for all $t \in \{1, \ldots, T-1\}$.

1. From Proposition 1 of Margiotta and Miller (2000, 678), the value function solving the consumption savings problem at retirement date $T+1$ is

$$V_{T+1}\left(h, \xi_{T+1}, a_{t+1}\right) \equiv -b_{T+1} \exp\left[-\left(a_{T+1} + \rho\xi_{T+1}\right)/b_{T+1}\right].$$

Suppose a manager works in firm and rank coordinate pair (j,k) at age T for one period and then retires. After selecting job match (j,k), he chooses consumption and next period's endowment

(c_T, e_{T+1}) optimally to maximize

$$
\begin{aligned}
&- \alpha_{jkT}(h) \exp\left(-\varepsilon_{jkT}^{*}\right) \exp\left(-\rho c_T\right) - E_T\left[b_{T+1} \exp\left(-\frac{a_{T+1} + \rho\xi_{T+1}}{b_{T+1}}\right) \upsilon_{jk,T+1}\right] \\
&\qquad\qquad \equiv -\alpha_{jkT}(h) \exp\left(-\varepsilon_{jkT}^{*}\right) \exp\left(-\rho c_T\right) \\
&\qquad\qquad - E_T\left[\upsilon_{jk,T+1} A_{T+1}\left(\overline{H}_{jk}(h)\right) b_{T+1} \exp\left(-\frac{a_{t+1} + \rho\xi_{T+1}}{b_{T+1}}\right)\right] \qquad (47)
\end{aligned}
$$

subject of his budget constraint. Differentiating with respect to c_T and solving, with reference to Equation (15) of Margiotta and Miller (2000, 680), the value function for this problem is

$$
V_{jkT}(h, \xi_T, a_t) \equiv -b_T \alpha_{jkT}(h)^{1/b_T} \exp\left(-\varepsilon_{jkT}^{*}/b_T\right) E_T[\upsilon_{jk,T+1}]^{1-\frac{1}{b_T}} \exp\left(-\frac{a_T + \rho\xi_T}{b_T}\right).
$$

$V_{0T}(h, \xi_T, a_T)$ is similarly defined. Integrating over ε_T, the idiosyncratic disturbance vector that is revealed at the beginning of the period, and averaging over job matches (j,k) using the choice probabilities yields

$$
\begin{aligned}
V_T(h, \xi_T, a_T) &\equiv -b_T\left[p_{0T}(h) V_{0T}(h) + \sum_{j=1}^{J}\sum_{k=1}^{K} p_{jkT}(h) V_{jkT}(h)\right] \\
&= -b_T \exp\left(-\frac{a_t + \rho\xi_T}{b_T}\right) \\
&\quad \times \left(\begin{array}{l} p_{0T}(h)\alpha_0(h)^{\frac{1}{b_T}} E\left[\exp\left(\frac{-\varepsilon_{0T}^{*}}{b_T}\right)\right] \\ + \sum_{j=1}^{J}\sum_{k=1}^{K}\left\{p_{jkT}(h)\alpha_{jkT}(h)^{\frac{1}{b_T}} E\left[\exp\left(\frac{-\varepsilon_{jkT}^{*}}{b_T}\right)\right] E_T[\upsilon_{jk,T+1}]^{1-\frac{1}{b_T}}\right\} \end{array}\right) \\
&= -b_T \exp\left(-\frac{a_t + \rho\xi_T}{b_T}\right) A_T(h).
\end{aligned}
$$

2. The proof is completed with an induction showing that for all ages $t \in \{1, \ldots, T-1\}$,

$$
\begin{aligned}
V_{jkt}\left(h, \xi_t, a_t, \varepsilon_{jkt}^{*}\right) &\equiv -\alpha_{jkt}(h)^{1/b_t} \exp\left(-\varepsilon_{jkt}^{*}/b_t\right) \\
&\qquad \times E_t\left[\upsilon_{jk,t+1} A_{t+1}\left(\overline{H}_{jk}(h)\right)\right]^{1-\frac{1}{b_t}} b_t \exp\left(-\frac{a_t + \rho\xi_t}{b_t}\right) \qquad (48)
\end{aligned}
$$

and

$$
V_t(h, \xi_t, a_t) = -b_t \exp\left[-\left(a_t + \rho\xi_t\right)/b_t\right] A_t(h)
$$

Suppose both equations are true for all ages $s \in \{t+1, \ldots, T\}$. Given job selection (j,k), the solution to the consumption savings decision at age t is found by maximizing

$$
\begin{aligned}
&- \alpha_{jkt}(h) \exp\left(-\varepsilon_{jkt}^{*}\right) \exp\left(-\rho c_t\right) - E_t[V_{t+1}\left(\overline{H}_{jk}(h), \xi_{t+1}, a_{t+1}\right)] \\
&= -\alpha_{jkt}(h) \exp\left(-\varepsilon_{jkt}^{*}\right) \exp\left(-\rho c_t\right) - E_t\left\{b_{t+1} \exp\left(-\frac{a_{t+1} + \rho\xi_{t+1}}{b_{t+1}}\right) \upsilon_{jk,t+1} A_{t+1}\left[\overline{H}_{jk}(h)\right]\right\}
\end{aligned}
$$

with respect to $\left(c_t, \xi_{t+1}\right)$. Substituting t for T and $\upsilon_{jk,t+1} A_{t+1}\left[\overline{H}_{jk}(h)\right]$ for $\upsilon_{jk,T+1}$ in Expression (47) above, Expression (48) follows directly. Integrating over ε_t, the idiosyncratic disturbance vector revealed at the beginning of the period, and averaging over the JK job matches yields:

$$\begin{aligned}
V_t(h,\xi_t,a_t) &= p_{0t}(h)\,E_t\left[V_{0t}\left(h,\xi_t,a_t,\varepsilon_{0t}^*\right)\right]+\sum_{j=1}^{J}\sum_{k=1}^{K}p_{jkt}(h)\,E_t\left[V_{jkt}\left(h,\xi_t,a_t,\varepsilon_{jkt}^*\right)\right] \\
&= -b_t\exp\left(\frac{a_t+\rho\xi_t}{b_t}\right)\left\{\begin{array}{l} p_{0t}(h)E\left[\exp\left(\frac{-\varepsilon_{0t}^*}{b_t}\right)\right] \\ +\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\begin{array}{l} p_{jkt}(h)\alpha_{jkt}(h)^{\frac{1}{b_t}}E\left[\exp\left(\frac{-\varepsilon_{jkt}^*}{b_t}\right)\right] \\ \times E_t\left[\upsilon_{jk,t+1}A_{t+1}\left(\overline{H}_{jk}(h)\right)\right]^{1-\frac{1}{b_t}}\end{array}\right)\end{array}\right\} \\
&= -b_t\exp\left(\frac{a_t+\rho\xi_t}{b_t}\right)A_t(h)
\end{aligned}$$

the third equality following from the recursive definition of $A_t(h)$. Substituting the expression for $V_t(h,\xi_t,a_t)$ back into the expression for $V_{jkt}\left(h,e_t,\xi_t,\varepsilon_{jkt}^*\right)$ then completes the induction.

■

Proof of Lemma 2. The manager optimizes his expected lifetime utility at age t by choosing the highest valued conditional valuation function, given by Expression (48), of the JK job matches and retirement. The solution can be found by taking logarithms (a monotone transformation) and maximizing

$$\frac{\varepsilon_{jkt}}{b_t}-\ln b_t-\frac{1}{b_t}\ln\alpha_{jkt}(h)-\left(1-\frac{1}{b_t}\right)\ln E_t\left[\upsilon_{jk,t+1}A_{t+1}\left(\overline{H}_{jk}(h)\right)\right]-\frac{a_t+\rho e_t}{b_t}$$

with respect to potential job matches and retirement, where

$$-\ln\left[-V_{0t}(h)\right]=\frac{\varepsilon_{0t}}{b_t}-\ln b_t-\frac{a_t+\rho e_t}{b_t}.$$

Subtracting $\left[\ln b_t-(a_t+\rho e_t)/b_t\right]$ from each conditional valuation function and multiplying by b_t completes the proof. ■

Proof of Lemma 3. The proof of this lemma essentially follows from Lemmas 1 and 2 by extending the choice set to effort levels as well, and substituting $B_t(h,h')$ for $A_t(h)$ in the proofs of those lemmas.

■

Proof of Lemma 4. We first prove the lemma for the case when human capital is private information. Throughout this proof, we fix (j,k,t,h) and consolidate the notation by defining

$$\gamma_1\equiv\exp\left\{q_{jk}\left[p_t(h)\right]\right\}^{1/(1-b_t)}\alpha_{jkt}(h)^{1/(b_t-1)}A_{t+1}\left[\overline{H}_{jk}(h)\right], \tag{49}$$

$$\gamma_2\equiv\alpha_{jkt}(h)^{1/(1-b_t)}A_{t+1}\left[\overline{H}_{jk}(h)\right], \tag{50}$$

and

$$\gamma_3\equiv\beta_{jkt}(h)^{1/(b_t-1)}B_{t+1}\left[\underline{H}_{jk}(h),\overline{H}_{jk}(h)\right], \tag{51}$$

where, for convenience, we have suppressed the dependence of $(\gamma_1,\gamma_2,\gamma_3)$ on (j,k,t,h) to reduce the notational clutter. Thus, the participation and incentive-compatibility constraints can be expressed in terms of the new notation as

$$\gamma_1E_t\left[\upsilon_{jk,t+1}\right]=1$$

and

$$\gamma_2E_t\left[\upsilon_{jk,t+1}\right]\leq\gamma_3E_t\left[\upsilon_{jk,t+1}g_{jkt}(\pi,h)\right].$$

Since the expectation operator preserves linearity, both the participation constraint (12) and the incentive-compatibility constraint (25) are rendered linear in $\upsilon_{jk,t+1}$, after multiplying both sides of the latter by $A_{t+1}\left[\overline{H}_{jk}(h)\right]E_t\left[\upsilon_{jk,t+1}\right]$. The objective function, the expected wage bill $E_t(w_{jk,t+1})$, can be expressed as a concave function of $\upsilon_{jk,t+1}$, namely $E_t(\ln\upsilon_{jk,t+1})$. Therefore, the Kuhn Tucker Theorem applies, and the Lagrangian for the problem in which the j^{th} firm elicits diligent work from the k^{th} rank can be written as

$$E_t[\ln(\upsilon_{jk,t+1})] + \eta_0 E_t\left[1 - \upsilon_{jk,t+1}\gamma_1\right] + \eta_1 E_t\left[\upsilon_{jk,t+1} g_{jkt}(\pi, h)\gamma_3 - \upsilon_{jk,t+1}\gamma_2\right],$$

where, for convenience, we have also suppressed the dependence of η_0 and η_1 on (j, k, h).

The proof now follows directly from Proposition 3 of Margiotta and Miller (2000, 713–14). Briefly, the first-order condition is

$$\upsilon_{j,k,t+1}^{-1} = \eta_0\gamma_1 - \eta_1 g_{jkt}(\pi, h)\gamma_3 + \eta_1\gamma_2. \tag{52}$$

Multiplying this equation by $\upsilon_{j,k,t+1}$ and taking expectations, we obtain

$$1 = \eta_0\gamma_1 E_t\left[\upsilon_{jk,t+1}\right] + \eta_1 E_t\left[\upsilon_{jk,t+1} g_{jkt}(\pi, h)\gamma_3 - \upsilon_{jk,t+1}\gamma_2\right] = \eta_0.$$

Substituting for η_0 back into the first-order condition yields

$$\upsilon_{j,k,t+1}^{-1} = \gamma_1 - \eta_1 g_{jkt}(\pi, h)\gamma_3 + \eta_1\gamma_2, \tag{53}$$

where η_1 solves the incentive-compatibility condition

$$0 = E_t\left[\frac{\gamma_3 g_{jkt}(\pi, h) - \gamma_2}{\gamma_1 - \eta_1 g_{jkt}(\pi, h)\gamma_3 + \eta_1\gamma_2}\right]. \tag{54}$$

Using Equations (49), (50), and (51) to replace $(\gamma_1, \gamma_2, \gamma_3)$ with $q_{jk}\left[p_t(h)\right]$, $\alpha_{jkt}(h)$, $A_t(h)$, and $B_{t+1}(h, h')$ in Equations (53) and (54) yields the equations in the lemma upon rearrangement. This proves the case where human capital is private information. To prove the public-information case, we reset $\gamma_2 \equiv \alpha_{jkt}(h)^{1/(1-b_t)}$ and $\gamma_3 \equiv \beta_{jkt}(h)^{1/(b_t-1)}$ and follow the same steps. ■

Proof of Lemma 5. Appealing to the optimization problems in Lemmas 2 and 3, we recursively define, for each (h, π), the probability vector $(p_{0t}^e(h), \ldots, p_{Jt}^e(h))$ and the human capital functions $A_t^e(h)$, and $B_t^e(h, h')$ by successively substituting the compensation function,

$$w_{jk,t+1}^e(\pi, h) \equiv F_{jkt}(h) + \frac{b_{t+1}}{\rho}\left\{r_{jk,t+1}^e(\pi, h) - E\left[r_{jk,t+1}^e(\pi, h)\right]\right\},$$

for $w_{jk}(\pi, h)$ into the respective recursions, where $r_{jk,t+1}^e(\pi, h)$ is defined using (27) of Lemma 4. Next, we define the strategies of the managers and firms, as well as the firms' beliefs about their managers.

1. Managers who have never shirked before solve the problem of Lemma 2, and managers who have shirked at least once solve the choice problem of Lemma 3.

2. In the ultimatum game, all managers demand $w_{jkt}^e(\pi, h)$ from the firm prescribed by Item 1.

3. Firms accept contracts of the form $w_{jkt}^e(\pi, h)$ and reject everything else.

4. Firms believe their managers have never shirked unless they are confronted with a wage schedule in the ultimatum game that deviates from $w_{jkt}^e(\pi, h)$ and is not incentive compatible, in which case, they revert to thinking that the manager has shirked at least once.

By construction, $w^e_{jkt}(\pi, h)$ does not depend on future returns to the firm and therefore satisfies the no-commitment property. By inspection, $E_t[w^e_{jkt}(\pi, h)] \equiv F_{jkt}(h)$, implying the expected rents to each firm are zero. To establish that the strategies and beliefs together constitute a sequential equilibrium, as defined in Kreps and Wilson (1982), we now prove the strategies given in the first three items are sequentially rational for the beliefs ascribed to firms in the fourth, and that those beliefs are consistent.

1. From the recursive definition of $w^e_{jkt}(\pi, h)$, $A^e_t(h)$, and $B^e_t(h, h')$, it follows from Lemma 2 that when $h = h'$ the manager's job-match choices are sequentially rational. Now suppose $h \neq h'$. If the manager is constrained to offer $w^e_{jkt}(\pi, h')$ or to exit, then Lemma 3 implies his value of human capital is $B^e_t(h, h')$. it then follows from the recursive definition of $B^e_t(h, h')$ that the manager's match choices are sequentially rational. If a manager offers any contract aside from $w^e_{jkt}(\pi, h')$, it is rejected. Hence, the expected utility is the same as exiting, and no value is gained from managers departing from the strategies prescribed for them.

2. When firms are offered $w^e_{jkt}(\pi, h)$, they believe $h = h'$ and hence will break even by accepting the contract. If they offered a different contract, they believe $h \neq h'$ and the manager has shirked at least once in the past. It is optimal to accept the contract in the first case and to reject it in the second.

3. To demonstrate these beliefs are consistent, we suppose with probability $1/i$ a firm accepts a contract not of the form $w^e_{jkt}(\pi, h')$ and with probability $1/i$ a firm rejects a contract of the form $w^e_{jkt}(\pi, h')$. With probability $1/i$, a manager who has not shirked before deviates from diligent work, and with probability $1/i$ a manager who has shirked before deviates from his prescribed strategy of diligent work. Managers deviate from their optimal job-match choice to one of the other choices with probability $1/i$, giving each of the other choices equal weight. Managers who have shirked before demand contracts of the form $w_{jkt}(\pi, h') \neq w^e_{jkt}(\pi, h')$ with probability $1/i$, but managers who have not shirked before demand contracts of the form $w_{jkt}(\pi, h') \neq w^e_{jkt}(\pi, h')$ with probability $(1/i)^3$. The support of the distribution of non-$w^e_{jkt}(\pi, h')$ contracts covers the entire space of such contracts. This perturbation from the conjectured equilibrium strategy is completely mixed, so the Bayes rule applies for computing the probabilities of nodes within any given information set. In particular, the probability of a firm being confronted with a non-$w^e_{jkt}(\pi, h')$ contract from a manager who has shirked before is greater than $(1/i)^2$. So, when a firm is confronted with a non-$w^e_{jkt}(\pi, h')$ contract, it places a probability of less than $(1/i)^3 \left[(1/i)^3 + (1/i)^2\right] = (1+i)^{-1}$ that the manager has never shirked. In the limit of $i \to \infty$, this probability converges to zero, the firm's assessment.

■

Proof of Lemma 6. The formula for $q_{jk}[p_t(h)]$ given by (34) is well known; see Hotz and Miller (1993) for an example. Denoting the probability density function of $\varepsilon^*_{jkt} \equiv d_{jk}\varepsilon_{jkt}$ by $\overline{\varphi}(\varepsilon^*_{jkt})$, we first derive an expression for $E[\exp(-\varepsilon^*_{jkt}/b_t)]$ and then use it in our derivation of the formula for $A_t(h_t)$:

1. For each (j, k, t), denote the deterministic part of utility by

$$W_{jkt} \equiv \ln \alpha_{jkt} + (b_t - 1) \ln A_{t+1}\left[\overline{H}_{jk}(h)\right] + (b_t - 1) \log\left\{E_t\left[\upsilon_{jk,t+1}\right]\right\}. \tag{55}$$

Then, (j, k) is chosen at t if $\varepsilon_{jkt} + W_{jkt}$ is maximal for all (j', k'). 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,

$$G_{jk}(\varepsilon_{11t}, \ldots, \varepsilon_{JKt}) = \exp(-\varepsilon_{jkt}) \prod_{(j',k')} \exp\left[-\exp\left(-\varepsilon_{j'k't}\right)\right]. \tag{56}$$

Using the well-known fact that

$$W_{jkt} - W_{j'k't} = \log p_{jkt} - \log p_{j'k't}, \tag{57}$$

it now follows from (56) and (57) that

$$G_{jk}(\varepsilon_{jkt} + W_{jkt} - W_{11t}, \ldots, \varepsilon_{jkt} + W_{jkt} + W_{JKt}) = \exp[-\varepsilon_{jkt} - \exp(-\varepsilon_{jkt} - \log p_{jkt})]. \tag{58}$$

From Equation (55) and Lemma 3, the conditional choice probability for (j,k) can be expressed as

$$p_{jkt} = \int_{-\infty}^{\infty} G_{jk}(\varepsilon_{jkt} + W_{jkt} - W_{11t}, \ldots, \varepsilon_{jkt} + W_{jkt} + W_{JKt})\, d\varepsilon_{jkt}. \tag{59}$$

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

$$\begin{aligned}
\overline{\varphi}\left(\varepsilon^*_{jkt}\right) &= p_{jkt}^{-1} \frac{\partial \int_{-\infty}^{\varepsilon^*_{jkt}} G_{jk}(\varepsilon_{jkt} + W_{jkt} - W_{11t}, \ldots, \varepsilon_{jkt} + W_{jkt} + W_{JKt})\, \mathrm{d}\varepsilon_{jkt}}{\partial \varepsilon^*_{jkt}} \\
&= \exp\left[-\varepsilon^*_{jkt} - \log p_{jkt} - \exp\left(-\varepsilon^*_{jkt} - \log p_{jkt}\right)\right]
\end{aligned}$$

To derive $E[\exp(-\varepsilon^*_{jkt}/b_t)]$, we draw from Equations (15) and (17) of Chapter 21 of Johnston and Kotz (1970, 277–78), who prove that the moment generating function for ε^*_{jkt} is

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

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

$$E\left[\exp\left(\varepsilon^*_{jkt}/b_t\right)\right] = \exp\left(\log p_{jkt}(h)^{1/b_t}\right) \Gamma\left[(b_t+1)/b_t\right] = p_{jkt}(h)^{1/b_t} \Gamma\left[(b_t+1)/b_t\right]. \tag{60}$$

2. Rearranging the participation constraint (12) and substituting for $q_{jk}[p_t(h)]$ from (34), we obtain

$$\alpha_{jkt}(h)^{1/b_t} E_t\left[\upsilon_{jk,t+1}\right]^{(b_t-1)/b_t} A_{t+1}\left[\overline{H}_{jk}(h)\right]^{(b_t-1)/b_t} = \left[p_{jkt}(h)/p_{0t}(h)\right]^{1/b_t}. \tag{61}$$

In the recursion for $A_{t+1}(h, b_t)$ given in (9), we now substitute for

$$\alpha_{jkt}(h)^{1/b_t} E_t\left[\upsilon_{jk,t+1}\right]^{(b_t-1)/b_t} A_{t+1}\left[\overline{H}_{jk}(h)\right]^{(b_t-1)/b_t}$$

using (61), and also for $E[\exp(\varepsilon^*_{jkt}/b_t)]$ using Equation (60), to obtain

$$\begin{aligned}
A_t(h) &= p_{0t}(h)^{1+\frac{1}{b_t}} \Gamma\left[\frac{b_t+1}{b_t}\right] + \sum_{j=1}^{J}\sum_{k=1}^{K} \left\{ p_{jkt}(h)^{1+\frac{1}{b_t}} \Gamma\left[\frac{b_t+1}{b_t}\right] \left[\frac{p_{0t}(h)}{p_{jkt}(h)}\right]^{1/b_t} \right\} \\
&= p_{0t}(h)^{\frac{1}{b_t}} \Gamma\left[\frac{b_t+1}{b_t}\right]
\end{aligned}$$

as required.

■
Proof of Lemma 7. Following Gayle and Miller (2009), take the expectation of (53) to obtain

$$E_t\left[\upsilon_{jk,t+1}^{-1}\right] = \gamma_1 - \eta_1\gamma_3 + \eta_1\gamma_2, \tag{62}$$

and also take the limit of (53) as $\pi \to \infty$ to get

$$\overline{\upsilon}_{jk,t+1}^{-1} = \gamma_1 + \eta_1\gamma_2. \tag{63}$$

Differencing (53) and (62) gives

$$\overline{\upsilon}_{jk,t+1}^{-1} - E_t\left[\upsilon_{jk,t+1}^{-1}\right] = \eta_1\gamma_3. \tag{64}$$

Subtracting (53) from (63) gives

$$\overline{\upsilon}_{jk,t+1}^{-1} - \upsilon_{jk,t+1}^{-1} = \eta_1 g_{jkt}(\pi,h)\gamma_3. \tag{65}$$

The proof of the lemma is completed by taking the quotient of (64) and (65), which yields (41). ■
Proof of Lemma 8. It is straightforward to show that the incentive-compatibility constraint (25) is satisfied with equality in the proof of Lemma 4:

1. When human capital is public information, incentive compatibility can be expressed as

$$\beta_{jkt}(h)\left\{E_t\left[\upsilon_{jk,t+1} g_{jkt}(\pi,h)\right]\right\}^{(b_t-1)} = \alpha_{jkt}(h)\left\{E_t\left[\upsilon_{jk,t+1}\right]\right\}^{(b_t-1)}. \tag{66}$$

   Exponentiating Equation (12) in the text gives

$$\exp\left\{q_{jk}\left[p_t(h)\right]\right\} A_{t+1}\left[\overline{H}_{jk}(h)\right]^{(1-b_t)} = \alpha_{jkt}(h) E_t\left[\upsilon_{jk,t+1}\right]^{(b_t-1)}. \tag{67}$$

   Differencing (67) and (66) we obtain

$$\beta_{jkt}(h)\left\{E_t\left[\upsilon_{jk,t+1} g_{jkt}(\pi,h)\right]\right\}^{(b_t-1)} = \exp\left\{q_{jk}\left[p_t(h)\right]\right\} A_{t+1}\left[\overline{H}_{jk}(h)\right]^{(1-b_t)}. \tag{68}$$

   Making $\beta_{jkt}(h)$ the subject of the equation, and substituting for $g_{jkt}(\pi,h)$ from (41) and $q_{jk}\left[p_t(h)\right]$ from (34) yields

$$\begin{aligned}
\beta_{jkt}(h) &= \exp\left\{q_{jk}\left[p_t(h)\right]\right\} A_{t+1}\left[\overline{H}_{jk}(h)\right]^{(1-b_t)}\left\{E_t\left[\upsilon_{jk,t+1}\frac{\overline{\upsilon}_{jk,t+1}^{-1} - \upsilon_{jk,t+1}^{-1}}{\overline{\upsilon}_{jk,t+1}^{-1} - E_t\left[\upsilon_{jk,t+1}^{-1}\right]}\right]\right\}^{1-b_t} \\
&= \frac{p_{0t}(h)}{p_{jkt}(h)} A_{t+1}\left[\overline{H}_{jk}(h)\right]^{(1-b_t)}\left\{\frac{E_t[\upsilon_{jk,t+1}] - \overline{\upsilon}_{jk,t+1}}{1 - \overline{\upsilon}_{jk,t+1} E_t[\upsilon_{jk,t+1}^{-1}]}\right\}^{1-b_t} \\
&= \beta_{jkt}^{*}(h).
\end{aligned} \tag{69}$$

2. When human capital is hidden information, the incentive-compatibility constraint becomes

$$\begin{aligned}
&\beta_{jkt}(h)\left\{E_t\left[\upsilon_{jk,t+1} g_{jkt}(\pi,h)\right] B_{t+1}\left[\underline{H}_{jk}(h), \overline{H}_{jk}(h)\right]\right\}^{(b_t-1)} \\
&\qquad = \alpha_{jkt}(h)\left\{E_t\left[\upsilon_{jk,t+1}\right] A_{t+1}\left[\overline{H}_{jk}(h)\right]\right\}^{(b_t-1)}
\end{aligned} \tag{70}$$

and the exponentiated inversion equation can be written as

$$\exp\left\{q_{jk}\left[p_t\left(h\right)\right]\right\} = \alpha_{jkt}\left(h\right)\left\{E_t\left[\upsilon_{jk,t+1}\right]A_{t+1}\left[\overline{H}_{jk}\left(h\right)\right]\right\}^{(b_t-1)}. \tag{71}$$

Differencing (71) and (70) yields

$$\beta_{jkt}\left(h\right)\left\{E_t\left[\upsilon_{jk,t+1}g_{jkt}\left(\pi,h\right)\right]B_{t+1}\left[\underline{H}_{jk}\left(h\right),\overline{H}_{jk}\left(h\right)\right]\right\}^{(b_t-1)} = \exp\left\{q_{jk}\left[p_t\left(h\right)\right]\right\} \tag{72}$$

or

$$\beta_{jkt}\left(h\right)B_{t+1}\left[\underline{H}_{jk}\left(h\right),\overline{H}_{jk}\left(h\right)\right]^{(b_t-1)} = \exp\left\{q_{jk}\left[p_t\left(h\right)\right]\right\}E_t\left[\upsilon_{jk,t+1}g_{jkt}\left(\pi,h\right)\right]^{(1-b_t)}.$$

Appealing to the definition of $\beta^*_{jkt}\left(h\right)$, we now obtain

$$\begin{aligned}
\beta^*_{jkt}\left(h\right) &= \beta_{jkt}\left(h\right)\left\{\frac{B_{t+1}\left[\underline{H}_{jk}\left(h\right),\overline{H}_{jk}\left(h\right)\right]}{A_{t+1}\left[\overline{H}_{jk}\left(h\right)\right]}\right\}^{(b_t-1)} \\
&= \exp\left\{q_{jk}\left[p_t\left(h\right)\right]\right\}E_t\left[\upsilon_{jk,t+1}g_{jkt}\left(\pi,h\right)\right]^{(1-b_t)}A_{t+1}\left[\overline{H}_{jk}\left(h\right)\right]^{(1-b_t)} \\
&= \exp\left\{q_{jk}\left[p_t\left(h\right)\right]\right\}A_{t+1}\left[\overline{H}_{jk}\left(h\right)\right]^{(1-b_t)}\left\{\frac{E_t[\upsilon_{jk,t+1}]-\overline{\upsilon}_{jk,t+1}}{1-\overline{\upsilon}_{jk,t+1}E_t[\upsilon^{-1}_{jk,t+1}]}\right\}^{1-b_t}
\end{aligned}$$

upon simplification, using the formulas for $q_{jk}\left[p_t\left(h\right)\right]$ and $E_t\left[\upsilon_{jk,t+1}g_{jkt}\left(\pi,h\right)\right]$ derived above.

■

Proof of Lemma 9. The formula for $\beta_{jkt}\left(h\right)$ is obtained by substituting (42) into (45). To prove (46), the formula for $B_t\left(h,h'\right)$, we first note that if ε_{jkt} is independently and identically distributed as a Type I extreme value with location and scale parameters $(0,1)$, then from (34) and (21)

$$V'_{jkt}(h,h') = \frac{p_{0t}(h,h')}{p_{jkt}(h,h')}. \tag{73}$$

Summing over (j,k) and rearranging we obtain

$$p_{0t}(h,h') = \left\{1+\sum_{j=1}^{J}\sum_{k=1}^{K}\left[V'_{jkt}(h,h')\right]^{-1}\right\}^{-1}. \tag{74}$$

Following the same logic used to derive (60), we can show when shirking is an option and human capital is private information:

$$E_t[\exp\left(-\frac{\varepsilon^*_{jkt}}{b_t}\right)] = p_{jkt}\left(h,h'\right)^{\frac{1}{b_t}}\Gamma\left(\frac{b_t+1}{b_t}\right). \tag{75}$$

Substituting (75) along with the conditional choice-probability ratios (73) and the retirement probability (74) into (19) yields

$$\begin{aligned}
B_t\left(h,h'\right) &= \Gamma\left(\frac{b_t+1}{b_t}\right)\left\{p_{0t}\left(h,h'\right)^{1+\frac{1}{b_t}}+\sum_{j=1}^{J}\sum_{k=1}^{K}\left[p_{jkt}\left(h,h'\right)^{1+\frac{1}{b_t}}V'_{jkt}(h,h')\right]\right\} \\
&= \Gamma\left(\frac{b_t+1}{b_t}\right)\frac{\left\{1+\sum_{j=1}^{J}\sum_{k=1}^{K}\left[V'_{jkt}(h,h')\right]^{-1-\frac{1}{b_t}}V'_{jkt}(h,h')\right\}}{\left\{1+\sum_{j=1}^{J}\sum_{k=1}^{K}\left[V'_{jkt}(h,h')\right]^{-1}\right\}^{1+\frac{1}{b_t}}},
\end{aligned}$$

which simplifies to (46). ■

Proof of Lemma 10. There are two steps to the proof. First, for any finite positive $\widehat{\rho}$, and any probability density function $\widehat{\varphi}(\varepsilon)$ with the same support as $\widetilde{\varphi}(\varepsilon)$, we define another parameterization, $\widehat{\theta} \in \Theta$. To complete the proof, we show that the model defined by $\widehat{\theta}$ generates the same data as $\widetilde{\theta}$, and is therefore observationally equivalent. Given the compensation process generated by $\widetilde{\theta}$, and our construction in the first step, the conditional choice probabilities of $\widehat{\theta}$ replicate those of $\widetilde{\theta}$. Thus, the only remaining task is to show that the compensation schedule generated by $\widehat{\theta}$ reproduces the schedule generated by $\widetilde{\theta}$. This only leaves us to prove that the contracts are the same, a second step that follows directly from the analysis of the pure moral hazard model in Gayle and Miller (2009). Here, we prove the first step.

For any finite positive $\widehat{\rho}$, let $\widehat{v}_{jk,t+1} \equiv \exp\left[-\widehat{\rho} w_{jkt}(\pi)/b\right]$ and define

$$\widehat{g}_{jkt}(\pi, h) = \frac{\exp\left(\widehat{\rho}\overline{w}_{jkt}/b\right) - \widehat{v}_{jk,t+1}^{-1}}{\exp\left(\widehat{\rho}\overline{w}_{jkt}/b\right) - E_t\left[\widehat{v}_{jk,t+1}(\pi)^{-1}\right]}. \tag{76}$$

For any probability density function $\widehat{\varphi}(\varepsilon)$ with the same support as $\widetilde{\varphi}(\varepsilon)$, let

$$\widehat{E}_t\left[\exp\left(\varepsilon_{jkt}/b\right)\right] \equiv p_{jkt}(h_t)^{-1} \int d_{jkt} \exp\left(\varepsilon_{jkt}/b\right) \widehat{\varphi}(\varepsilon)\, \mathrm{d}\varepsilon$$

denote the conditional expectation of ε_{jkt}/b given the choices observed in the population but integrated with respect to the $\widehat{\varphi}(\varepsilon)$ density rather than $\widetilde{\varphi}(\varepsilon)$. Appealing to Proposition 1 of Hotz and Miller (1993), there exists a mapping $\widehat{q}(p)$ implied by $\widehat{\varphi}(\varepsilon)$ for any conditional valuation function. Starting with $\widehat{A}_t(h) = 1$ for all $t \geq R$, and given the density $\widehat{\varphi}(\varepsilon)$, we now recursively define $\widehat{\alpha}_{jkt}(h)$ and $\widehat{A}_t(h)$ to rationalize the choice probabilities generated by θ^* by repeatedly appealing to Equation (see text) and setting

$$\widehat{\alpha}_{jkt}(h_t) = \exp\left[\widehat{q}_{jk}(p_t(h_t))\right] \widehat{A}_{t+1}\left[\overline{H}_{jk}(h)\right]^{1-b} E_t\left[\widehat{v}_{jk,t+1}(\pi)\right]^{1-b}. \tag{77}$$

Finally, $\widehat{\beta}^*_{jkt}(h)$ is defined by setting

$$\widehat{\beta}^*_{jkt}(h) = \exp\left[\widehat{q}_{jk}(p_t(h))\right] \widehat{A}_{t+1}\left[\overline{H}_{jk}(h)\right]^{1-b} E_t\left[\widehat{v}_{jk,t+1}(\pi_t)\, \widehat{g}_{jkt}(\pi, h)\right]^{1-b}. \tag{78}$$

In this manner we construct another element in the parameter space, $\widehat{\theta} \in \Theta$ defined by:

$$\widehat{\theta} \equiv \left(\widehat{\alpha}_{jkt}(h), \widehat{\beta}^*_{jkt}(h), \widehat{\rho}, \widetilde{f}(\pi), \widehat{g}_{jkt}(\pi, h), \widehat{\varphi}(\varepsilon)\right)$$

The second step now follows from applying Proposition 8 of Gayle and Miller (2009). ■

References

[1] **Acharya, Viral V., and Paolo F. Volpin.** "Corporate Governance Externalities." *Review of Finance* 14 (2010): 1–33.

[2] **Albanesi, Stefania, and Claudia Olivetti.** "Gender and Dynamic Agency: Theory and Evidence on the Compensation of Top Executives" (Working Paper, Columbia University, 2008).

[3] **Alchian, Armen A., and Harold Demsetz.** "Production, Information Costs, and Economic Organization." *American Economic Review* 62 (1972): 777–95.

[4] **Altonji, Joseph C., and Robert A. Shakotko.** "Do Wages Rise with Seniority?" *Review of Economic Studies* 54(1987): 437–59.

[5] **Altug, Sumru, and Robert A. Miller.** "The Effect of Work Experience on Female Wages and Labour Supply." *Review of Economic Studies* 65 (1998): 45–85.

[6] **Antle, Rick, and Abbie Smith.** "Measuring Executive Compensation: Methods and an Application." *Journal of Accounting Research* 23 (1985): 296–325.

[7] **Arcidiacono, Peter, Jane C. Cooley, and Andrew Hussey.** "The Economic Returns to an MBA." *International Economic Review* 49 (2008): 873–99.

[8] **Aron, Debra J.** "Ability, Moral Hazard, Firm Size, and Diversification." *The RAND Journal of Economics* 19, no. 1 (1988): 72–87.

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

[10] **Becker, Gary S.** *Human Capital,* 1st ed. (New York, NY: Columbia University Press, 1964).

[11] **Ben-Porath, Yoram.** "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (1967): 352–65.

[12] **Bell, Linda A.** "Women-Led Firms and the Gender Gap in Top Executive Jobs" (IZA Discussion Paper 1689, July 2005).

[13] **Bertrand, Marianne, and Kevin F. Hallock.** "The Gender Gap in Top Corporate Jobs." *Industrial and Labor Relations Review* 55 (2001): 3–21.

[14] **Bertrand, Marianne, and Sendhill Mullainathan.** "Are CEOs Rewarded for Luck? The Ones without Principals Are." *Quarterly Journal of Economics* 116 (2001): 901–32.

[15] **Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin.** "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica* 74 (2006): 323–64.

[16] **Chevalier, Judith, and Glenn Ellison.** "Career Concerns of Mutual Fund Managers." *The Quarterly Journal of Economics* 114 (1999): 389–432.

[17] **Dewatripont, Mathais, Ian Jewitt, and Jean Tirole.** "The Economics of Career Concerns, Part I: Comparing Information Structures." *Review of Economic Studies* 66 (1999): 183–98.

[18] **Dicks, David L.** "Executive Compensation, Incentives, and the Role for Corporate Governance Regulation" AFA 2010 Atlanta Meetings Paper.

[19] **Dustmann, Christian, and Costas Meghir.** "Wages, Experience and Seniority." *Review of Economic Studies* 72 (2005): 77–108.

[20] **Eckstein, Zvi, and Kenneth I. Wolpin.** "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience." *Review of Economic Studies* 56 (1989): 375–90.

[21] **Ekeland, Ivar, James J. Heckman, and Lars P. Nesheim.** "Identification and Estimation of Hedonic Models." *Journal of Political Economy* 112, no. S1 (2004): S60–S109.

[22] **Fudenberg, Drew, Bengt Holmstrom, and Paul Milgrom.** "Short-Term Contracts and Long-Term Agency Relationships." *Journal of Economic Theory* 50 (1990): 1–31.

[23] **Gabaix, Xavier, and Augustin Landier.** "Why has CEO Pay Increased so Much?" *Quarterly Journal of Economics* 123 (2008): 49–100.

[24] **Garen, John.** "Executive Compensation and Principal–Agent Theory." *Journal of Political Economy* 102 (1994): 1175–99.

[25] **Gayle, George-Levi, and Limor Golan.** "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–97." *Review of Economic Studies* (forthcoming).

[26] **Gayle, George-Levi, Limor Golan, and Robert A. Miller.** "Gender Differences in Executive Compensation and Job Mobility." (Working Paper, Tepper School of Business Working Paper, Carnegie Mellon University, 2011).

[27] **Gayle, George-Levi, and Robert A. Miller.** "Has Moral Hazard Become a More Important Factor in Managerial Compensation?" *American Economic Review* 99 (2009): 1740–69.

[28] **Gayle, George-Levi, and Robert A. Miller.** "Identifying and Testing Models of Managerial Compensation" (Working Paper, Tepper School of Business, Carnegie Mellon University, 2011).

[29] **Gibbons, Robert, and Kevin J. Murphy.** "Optimal Incentive Contract in the Presence of Career Concerns: Theory and Evidence." *Journal of Political Economy* 100 (1992): 468–505.

[30] **Hall, Brian J., and Jeffrey B. Liebman.** "Are CEOS Really Paid Like Bureaucrats?" *The Quarterly Journal of Economics* 113 (1998): 653–80.

[31] **Hallock, Kevin F.** "Reciprocally Interlocking Boards of Directors and Executive Compensation." *Journal of Financial and Quantitative Analysis* 32 (1997): 331–44.

[32] **Haubrich, Joseph G.** "Risk Aversion, Performance Pay, and the Principal–Agent Problem" *Journal of Political Economy* 102 (1994): 258–76.

[33] **Heckman, James J., and Bo E. Honoré.** "The Empirical Content of the Roy Model." *Econometrica* 58 (1990): 1121–49.

[34] **Heckman, James J., Lance J. Lochner, and Christopher Taber.** "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1 (1998): 1–58.

[35] **Heckman, James J., and Guilherme L. Sedlacek.** "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market." *Journal of Political Economy* 93 (1985): 1077–125.

[36] **Holmstrom, Bengt.** 2003 "Managerial Incentive Problems: A Dynamic Perspective." *Review of Economic Studies* 66 (1999): 169–82.

[37] **Hotz, V. Joseph, and Robert A. Miller.** "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies* 60 (1993): 497–529.

[38] **Jensen, Michael, and Kevin J. Murphy.** "Performance Pay and Top-Management Incentives," *Journal of Political Economy* 98 (1990): 225–64.

[39] **Kreps, David M., and Robert Wilson.** "Sequential Equilibria." *Econometrica* 50 (1982): 863–94.

[40] **Kuhnen, Camelia M., and Jeffrey H. Zwiebel.** "Executive Pay, Hidden Compensation and Managerial Entrenchment" (Working Paper, Northwestern University, 2009).

[41] **Lee, Donghoon, and Kenneth I. Wolpin.** "Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica* 74 (2006): 1–46.

[42] **Lucas Jr., Robert E.** "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9 (1978): 508–23.

[43] **Malcomson, James M., and Frans Spinnewyn.** "The Multiperiod Principal–Agent Problem." *Review of Economic Studies* 55 (1988): 391–408.

[44] **Magnac, Thierry, and David Thesmar.** "Identifying Dynamic Discrete Decision Processes." *Econometrica* 70 (2002): 801–16.

[45] **Margiotta, Mary M., and Robert A. Miller.** "Managerial Compensation and the Cost of Moral Hazard." *International Economic Review* 41 (2000): 669–719.

[46] **Miller, Robert A.** "Job Matching and Occupational Choice." *Journal of Political Economy* 92 (1984): 1086–20.

[47] **Mirrlees, James A.** "The Optimal Structure of Incentives and Authority within an Organization." *The Bell Journal of Economics* 7, no. 1 (1976): 105–31.

[48] **Neal, Derek A.** "Industry-Specific Human Capital: Evidence from Displaced Workers." *Journal of Labor Economics* 13 (1995): 653–77.

[49] **Newey, Whitney K., and Daniel McFadden.** "Large Sample Estimation and Hypothesis Testing." Chapter 36 in *Handbook of Econometrics Volume 4* (Amsterdam: Elsevier Science, 1994).

[50] **Postel-Vinay, Fabien, and Jean-Marc Robin.** "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica* 70 (2002): 2295–350.

[51] **Rey, Patrick, and Bernard Salanie.** "Long-Term, Short-Term and Renegotiation." *Econometrica* 58 (1990): 597–620.

[52] **Rose, Nancy L., and Andrea Shepard.** "Diversification and Executive Compensation: Ability Premia or Managerial Entrenchment?" *RAND Journal of Economics* 28 (1997): 489–514.

[53] **Rosen, Sherwin.** "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82 (1974): 34–55.

[54] **Rosen, Sherwin** "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics* 13 (1982): 311–23.

[55] **Rosen, Sherwin.** "Contracts and the Market for Executives." (Working Paper NBER no. 3442, 1990).

[56] **Roy, Andrew D.** "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (1951): 135–46.

[57] **Selody, Karen.** "Is the Risk Worth the Reward for Top Female Executives?" (Working Paper, University of California, Berkeley, 2010).

[58] **Sicherman, Nachum and Oded Galor.** "A Theory of Career Mobility." *Journal of Political Economy* 98 (February, 1990): 169–192.

[59] **Taber, Christopher, and Rune Vejlin.** “Estimation of a Roy/Search/Compensating Differential Model of the Labor Market.” (Working Paper, University of Wisconsin, 2010).

[60] **Terviö, Marko.** “The Difference that CEOs Make: An Assignment Model Approach.” *American Economic Review* 98 (2008): 642–68.

[61] **Topel, Robert H.** “Specific Capital, Mobility and Wages: Wage Rise with Job Seniority.” *Journal of Political Economy* 99 (1991): 145–76.

[62] **Topel, Robert H., and Michael P. Ward.** “Job Mobility and the Careers of Young Men.” *Quarterly Journal of Economics* 107 (1992): 439–79.

[63] **Williamson, Oliver.** “Hierarchical Control and Optimum Firm Size.” *Journal of Political Economy* 75 (1967): 123–38.

TABLE 1: TITLES AND RANKS

Rank	Title(s)
1	Chairman of the Board & Vice Chairman of the Board; Chairman of the Board of a Subsidiary/Region & CEO of Subsidiary/Region; Chairman of a Subsidiary & Vice Chairman of a Subsidiary/Region; Chairman of the Board & Executive of a Subsidiary/Region
2	Chairman & President & CEO of the Company, CEO of the Company
3	President & Chief Operating Officer of the Company; Chairman of the Board & Chief Financial Officer of the Company; Chairman of the Board & Executive Vice President of the Company; Chairman of the Board and Chief Operating Officer of the Company
4	Executive Vice President of the Company; Executive Vice President and Chief Operating Officer of the Company; Executive Vice President and Chief Financial Officer of the Company; Chief Operating Officer of the Company; President of a Subsidiary/Region; Executive Vice President & Other Executive[‡] of the Company; President of a Subsidiary/Region & Executive Vice President of the Company; Executive Vice President of the Company and Chief Operating Officer of a Subsidiary/Region; President and CEO of a Subsidiary/Region; President and Chief Operating Officer of a Subsidiary/Region; President & Executive Vice President of the Company; CEO of a Subsidiary/Region & Executive Vice President of the Company; Senior Vice President of Company
5	Vice President of the Company; Senior Vice President and Other Executive[‡] of the Company; Vice President & Other Executive[‡] of the Company; Chief Financial Officer & Other Executive[‡] of the Company; Senior Vice President and Chief Financial Officer of the Company; Senior Vice President and Chief Operating Officer of the Company; Senior Vice President of the Company & President of a Subsidiary/Region; President & Other Executive[‡] of the Company; Senior Vice President of the Company & CEO of a Subsidiary/Region; CEO of a Subsidiary/Region & Other Executive[‡] of the Company; Chief Operating Officer of a Subsidiary/Region; Vice President and Chief Operating Office of the Company; Vice President of the Company & President of a Subsidiary/Region; Vice President of the Company & CEO of a Subsidiary/Region; Vice President of the Company & Chief Operating Officer of a Subsidiary/Region; Chief Financial Officer of the Company

[‡]Other Executive includes titles that did not occur often enough to warrant their own category and hence were grouped together. These include, but are not limited to, General Counsel, Chief Technology Officer, Chief Information Officer, Chief Marketing Officer, and Consultant.

TABLE 2: SAMPLE CHARACTERISTICS BY FIRM TYPE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Industrial Sectors			Firm Size		Insiders on Board	
Variable	Full Sample	Service	Primary	Consumer	Large	Medium	Small	Large	Small
Rank 1	0.06	0.05	0.05	0.07	0.08	0.05	0.03	0.08	0.03
Rank 2	0.28	0.28	0.30	0.29	0.25	0.30	0.30	0.27	0.29
Rank 3	0.07	0.08	0.07	0.09	0.09	0.07	0.06	0.09	0.06
Rank 4	0.37	0.38	0.35	0.38	0.40	0.39	0.32	0.37	0.37
Rank 5	0.21	0.21	0.23	0.17	0.17	0.19	0.29	0.18	0.25
Exit	0.13	0.14	0.12	0.13	0.12	0.12	0.15	0.12	0.13
Turnover	0.02	0.03	0.02	0.03	0.03	0.02	0.02	0.02	0.03
No College	0.19	0.18	0.16	0.24	0.16	0.21	0.21	0.20	0.17
Bachelors	0.81	0.82	0.84	0.76	0.84	0.79	0.79	0.80	0.83
MBA	0.23	0.24	0.23	0.22	0.27	0.21	0.19	0.22	0.24
MS/MA	0.20	0.23	0.20	0.16	0.18	0.18	0.23	0.19	0.21
PhD	0.19	0.19	0.21	0.15	0.19	0.19	0.18	0.18	0.19
Female	0.04	0.04	0.03	0.06	0.04	0.05	0.04	0.04	0.05
Execdir	0.41	0.45	0.47	0.51	0.40	0.43	0.40	0.50	0.31
Interlocked	0.03	0.03	0.03	0.03	0.02	0.03	0.03	0.04	0.02
Age	52.31	51.97	53.97	52.95	52.82	52.26	51.66	52.70	51.90
	(8.45)	(8.36)	(7.92)	(8.32)	(7.82)	(8.67)	(9.02)	(8.48)	(8.39)
Tenure	13.62	13.49	14.53	14.02	13.13	13.89	14.01	14.20	13.01
	(10.11)	(9.67)	(10.61)	(10.16)	(10.09)	(9.91)	(10.20)	(10.35)	(9.81)
Exec. Exp	16.60	16.70	17.26	17.37	15.81	17.12	17.21	16.93	16.25
	(10.09)	(10.09)	(9.87)	(10.00)	(9.64)	(10.30)	(10.38)	(10.19)	(9.97)
NBE	0.90	0.93	0.89	0.82	0.77	0.91	1.08	0.85	0.95
	(1.30)	(1.31)	(1.30)	(1.23)	(1.18)	(1.30)	(1.43)	(1.28)	(1.33)
NAE	0.84	0.84	0.85	0.87	0.82	0.83	0.86	0.79	0.88
	(1.34)	(1.32)	(1.34)	(1.40)	(1.31)	(1.35)	(1.38)	(1.32)	(1.37)
Ab. Return	−0.02	0.02	−0.04	−0.05	−0.03	−0.02	−0.02	−0.02	−0.03
	(0.53)	(0.66)	(0.44)	(0.49)	(0.42)	(0.53)	(0.66)	(0.53)	(0.53)
Assets	17,827	23,826	9,472	7,945	37,427	4,531	700	23,034	12,255
	(76,423)	(102,268)	(33,105)	(29,092)	(112,077)	(7,640)	(593)	(94,455)	(49,841)
Employees	22.93	15.21	19.40	38.21	43.82	11.12	2.62	24.87	20.84
	(52.80)	(33.50)	(44.40)	(77.40)	(69.00)	(36.20)	(2.30)	(58.80)	(45.30)
Equity	3,018	3,440	2,667	2,298	6,022	1,071	321.6	3,611	2,384
	(8,020)	(8,902)	(7,082)	(6,250)	(11,354)	(1,347)	(288)	(9,688)	(5,654)
Salary	477	487	501	576	614	429	332	506	448
	(329)	(311)	(313)	(388)	(381)	(260)	(203)	(360)	(289)
Compensation	2,551	4,487	2,367	2,537	3,612	2,077	1,499	2,536	2,566
	(18,323)	(23,380)	(12,857)	(19,124)	(22,197)	(15,401)	(14,212)	(19,718)	(16,737)
N	60,300	20,302	21,089	18,190	26,581	15,209	18,510	31,268	29,100

Sources: The data are for top managers from Standard & Poor's ExecuComp database for 1991 through 2006 matched with background data from the Marquis *Who's Who* database. *Note:* Standard deviation in parentheses; Asset and Equity are measured in millions of 2006 US$; Compensation and Salary are measured in thousands of 2006 US$; Employees is measured in thousands; Tenure and Executive Experience (Exec. Exp.) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is an indicator of whether the executive is a member of the board of directors.

Table 3: Executive Characteristics by Rank

Variable	Rank 1		2		3		4		5	
Exit	0.245		0.090		0.116		0.149		0.154	
Turnover	0.027		0.032		0.028		0.021		0.016	
Age	57.798	(8.220)	55.000	(7.433)	51.768	(7.363)	51.184	(8.140)	50.817	(8.804)
Female	0.019		0.015		0.025		0.053		0.06	
Tenure	15.784	(11.708)	14.412	(10.672)	13.388	(10.062)	13.383	(9.754)	13.221	(9.541)
Exec. Exp.	20.331	(11.113)	18.643	(9.638)	15.738	(9.555)	15.664	(9.901)	15.871	(10.124)
NBE	0.706	(1.186)	0.689	(1.118)	0.702	(1.174)	0.959	(1.341)	1.159	(1.427)
NAE	0.887	(1.357)	0.909	(1.374)	0.764	(1.310)	0.799	(1.323)	0.841	(1.350)
Execdir	0.720		0.929		0.675		0.177		0.069	
Interlocked	0.071		0.068		0.026		0.009		0.003	
No College	0.232		0.212		0.236		0.178		0.144	
Bachelors	0.768		0.788		0.764		0.822		0.856	
MBA	0.246		0.255		0.232		0.229		0.196	
MS/MA	0.155		0.172		0.168		0.212		0.214	
PhD	0.151		0.150		0.135		0.183		0.257	
Salary	615	(366)	719	(412)	559	(318)	397	(197)	304	(176)
Compensation	2,945	(26,035)	4,794	(26,701)	3,717	(19,009)	1,844	(11,644)	1,269	(9,438)
Observations	4,358		20,983		5,620		28,271		15,972	

Sources: The data are for top managers from Standard & Poor's ExecuComp database for 1991 through 2006 matched with background data from the Marquis *Who's Who* database. *Note:* Standard deviation in parentheses; Compensation and Salary are measured in thousands of 2006 US$; Tenure and Executive Experience (Exec. Exp.) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is an indicator of whether the executive is a member of the board of directors.

Table 4: Logit Coefficient Estimates of the Likelihood of Exit

Variable	Parameter		Elasticity (%)	
Rank 1 lagged	0.778	(0.054)	55.26	(3.23)
Rank 2 lagged	−0.106	(0.051)	−8.90	(4.28)
Exec. Exp.	0.018	(4.1E−3)	24.44	(5.69)
Exec. Exp. Sq.	−3.6E−4	(8.4E−5)	−11.38	(2.65)
Tenure	0.022	(3.5E−3)	24.66	(3.97)
Tenure Sq.	−3.0E−4	(8.1E−5)	−7.26	(1.97)
Female	0.218	(0.061)	17.42	(4.16)
No College	−0.410	(0.209)	−35.68	(18.90)
MBA	−0.935	(0.210)	−84.46	(20.10)
NBE	0.067	(0.012)	5.54	(0.62)
NAE	0.079	(0.009)	4.49	(0.78)
Age	−0.121	(0.013)	−527.74	(58.58)
Age Sq.	0.001	(1.2E−4)	312.89	(28.81)
Interlocked	−0.615	(0.086)	−55.14	(8.29)
Execdir	−0.736	(0.038)	−64.72	(3.50)
Bond Price	−0.232	(0.020)	−335.05	(28.89)
Constant	4.830	(0.503)		
Observations		51,808		

Sources: The data are for top managers from Standard & Poor's ExecuComp database for 1991 through 2006 matched with background data from the Marquis *Who's Who* database. *Note:* The elasticities are calculated at the mean of variables. For dummy variables, the change is from 0 to 1. Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is an indicator of whether the executive is a member of the board of directors. Other ranks, education types, and interactions are included but are not significant and hence are not reported here.

Table 5: Logit Estimates of the Likelihood of Promotion and Rank Choice

	Rank				Rank				
	1	2	3	4	1	2	3	4	5
Variable	Parameter				Elasticities (%)				
Rank 1 lagged	10.505	6.488	5.935	3.542	111	−290	−345	−585	−939
	(0.322)	(0.283)	(0.282)	(0.274)	(10)	(34)	(34)	(34)	(39)
Rank 2 lagged	7.824	9.687	5.080	3.501	−118	68	−393	−551	−901
	(0.276)	(0.207)	(0.219)	(0.194)	(25)	(4)	(20)	(18)	(22)
Rank 3 lagged	6.659	6.815	8.678	3.498	−88	−72	114	−404	−754
	(0.285)	(0.210)	(0.194)	(0.192)	(24)	(13)	(4)	(15)	(20)
Rank 4 lagged	5.158	4.598	4.682	5.994	−20	−76	−68	63	−536
	(0.225)	(0.134)	(0.112)	(0.056)	(21)	(10)	(10)	(4)	(7)
Exec. Exp.	−0.027	−0.041	−0.021	−0.013	−10	−33	1	14	35
	(0.012)	(0.011)	(0.011)	(0.009)	(15)	(9)	(12)	(7)	(11)
Exec. Exp. Sq.	4.0E−4	5.0E−4	2.0E−4	2.0E−4	6	8	−3	−4	−10
	(3.0E−4)	(2.0E−4)	(2.0E−4)	(2.0E−4)	(6)	(4)	(6)	(4)	(6)
Tenure	−0.026	−0.036	−0.027	−0.011	−10	−23	−11	11	26
	(0.010)	(0.009)	(0.009)	(0.007)	(10)	(6)	(8)	(5)	(8)
Tenure Sq.	0.001	0.001	0.001	3.0E−4	5	7	4	−3	−12
	(2.0E−4)	(2.0E−4)	(2.0E−4)	(2.0E−4)	(5)	(3)	(4)	(2)	(4)
Female	−0.845	−0.737	−0.729	−0.268	−43	−32	−31	15	42
	(0.247)	(0.200)	(0.186)	(0.114)	(20)	(12)	(14)	(7)	(10)
NBE	−0.197	−0.219	−0.172	−0.0577	−9	−1	−7	3	8
	(0.033)	(0.029)	(0.028)	(0.0189)	(2)	(2)	(2)	(1)	(1)
NAE	−0.012	0.019	−0.027	−0.0011	−1	1	−2	−1	−1
	(0.041)	(0.036)	(0.036)	(0.0259)	(2)	(2)	(2)	(1)	(2)
Age	0.160	0.358	0.195	0.0743	−9	1,024	174	−459	−847
	(0.049)	(0.043)	(0.042)	(0.0271)	(188)	(124)	(158)	(86)	(128)
Age Sq.	−0.001	−0.003	−0.002	−7.0E−4	136	−5	−111	236	434
	(0.001)	(0.001)	(4.0E−4)	(3.0E−4)	(89)	(60)	(80)	(44)	(66)
Execdir	1.438	2.279	1.208	0.348	−23	123	17	−70	−105
	(0.105)	(0.092)	(0.091)	(0.076)	(13)	(4)	(5)	(3)	(7)
Bond Price	−0.139	−0.294	−0.144	−0.087	−2	−265	−10	87	235
	(0.047)	(0.042)	(0.041)	(0.030)	(55)	(36)	(46)	(26)	(43)
Constant	−8.682	−8.630	−6.304	−2.437					
	(1.599)	(1.369)	(1.321)	(0.878)					
Observations	58,328	58,328	58,328	58,328					

Sources: The data are for top managers from Standard & Poor's ExecuComp database for 1991 through 2006 matched with background data from the Marquis *Who's Who* database. *Note:* The elasticities are calculated at the mean of variables. For dummy variables, the change is from 0 to 1. Rank 5 is the Base Outcome. Standard error in parentheses; Tenure and Executive Experience (Exec. Exp.) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is an indicator of whether the executive is a member of the board of directors. Other ranks, education types, and interactions are included but are not significant and hence are not reported here.

Table 6: Logit Estimates of the Likelihood of New Employer

Variable	Parameter		Elasticities (%)	
Primary Sector	−0.192	(0.073)	−18.7	(7.1)
Large board	−0.262	(0.058)	−25.7	(5.7)
Rank 1	0.912	(0.257)	86.1	(23.2)
Rank 2	2.420	(0.182)	213.0	(12.6)
Rank 3	1.002	(0.197)	94.7	(17.6)
Rank 2×Female	−1.174	(0.548)	−0.5	(0.2)
Rank 2 Lagged	−1.321	(0.187)	−132.0	(18.6)
Rank 3 Lagged	−0.432	(0.194)	−42.8	(19.1)
Exec. Exp.	0.052	(0.008)	82.8	(13.4)
Exec. Exp. Sq.	−0.001	(1.9E−4)	−20.6	(6.6)
Tenure	−0.227	(0.007	−302.0	(9.0)
Tenure Sq.	0.003	(1.6E−4)	88.1	(4.3)
NBE	−0.130	(0.025)	−11.1	(2.1)
NAE	−0.168	(0.024)	−13.7	(1.9)
Age	0.385	(0.047)	1,948	(239)
Age Sq.	−0.004	(0.001)	−992	(122)
Interlocked	−0.939	(0.286)	−93	(28.6)
Execdir	−1.036	(0.093)	−102	(9.2)
Bond Price	−0.241	(0.036)	−397	(59.4)
Constant	−8.227	(1.382)		
Observations			54,705	

Sources: The data are for top managers from Standard & Poor's ExecuComp database for 1991 through 2006 matched with background data from the Marquis *Who's Who* database. *Note:* The elasticities are calculated at the mean of variables. For dummy variables, the change is from 0 to 1. Standard error in parentheses; Tenure and Executive Experience (Exec. Exp.) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is an indicator of whether the executive is a member of the board of directors.

Table 7: OLS Coefficient Estimates of the Compensation Regression

Rank	π	π^2	Level	Variable	π	π^2	Level
1	9,839	−454	1,055	Interlocked	6,403	−1,496	−299
	(1,690)	(987)	(797)		(995)	(471)	(464)
2	6,007	−789	3,456	Execdir	7,695	−848	845
	(1,394)	(699)	(683)		(570)	(304)	(251)
3	2,627	−164	1,267	Bond Price	−1,521	531	−97
	(1,407)	(605)	(662)		(217)	(110)	(92)
4	1,529	−242	103	Rank 1 Lagged	12,085	−3,054	544
	(926)	(444)	(463)		(1,769)	(987)	(822)
2×Female	–	–	2,668	Rank 2 Lagged	14,640	−2,875	660
			(1,295)		(1,342)	(625)	(658)
Firm				Rank 3 Lagged	4,849	−1,100	597
New Employer	−12,396	2,155	−1,026		(1,389)	(586)	(653)
	(996)	(478)	(1,255)	Exec. Exp.	191	−42	2
Service Sector	3,149	88	777		(26)	(14)	(25)
	(419)	(222)	(198)	Tenure	−23	22	−40
Primary Sector	−3,609	1,537	−633		(25)	(14)	(20)
	(473)	(267)	(198)	NAE	−484	−58	215
Medium Firm	4,079	−253	937		(174)	(93)	(80)
	(437)	(201)	(214)	PhD	−871	83	11
Large Firm	12,703	−2,224	3,697		(464)	(223)	(212)
	(405)	(212)	(190)	Age	17	15	281
Large board	2,683	−1,203	280		(23)	(10)	(85)
	(358)	(176)	(163)	Age sq.	–	–	−3
Firm×Rank							(1)
2×New Employer	–	–	3,840	Constant	21,601	−9,114	−4,359
			(1,459)		(3,859)	(1,914)	(2,716)
3×New Employer	–	–	5,289				
			(1,975)	Observations	50,405	50,405	50,405

Sources: The data are for top managers from Standard & Poor's ExecuComp database for 1991 through 2006 matched with background data from the Marquis *Who's Who* database. *Note:* Compensation is measured in thousands of 2006 US$; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp.) are measured in years; NAE is the number of times the executive changed firms after becoming one of the ranks in our sample. Execdir is an indicator of whether the executive is a member of the board of directors. Other ranks, education types, and interactions are included but are not significant and hence are not reported here.

TABLE 8: COMPENSATING DIFFERENTIAL FOR NONPECUNIARY COST OF DILIGENCE VERSUS EXIT

Variable	Constant	Age-50	Tenure	Exec. Exp.	NBE	NAE	Female	Interlock	Execdir	No College	MBA	MS	PhD
Constant	1.628	0.007	0.016	−0.004	−0.006	0.025	−0.043	−0.074	0.333	−0.085	0.025	−0.047	0.005
	(0.071)	(0.001)	(0.002)	(0.000)	(0.004)	(0.001)	(0.024)	(0.011)	(0.049)	(0.007)	(0.006)	(0.007)	(0.005)
Rank 1	0.205						0.219	−0.125	−0.564				0.060
	(0.063)						(0.020)	(0.010)	(0.042)				(0.003)
Rank 2	0.263						0.347	−0.070	−0.545				0.060
	(0.063)						(0.020)	(0.008)	(0.034)				(0.003)
Rank 3	0.111						−0.072	−0.070	−0.545				0.060
	(0.063)						(0.020)	(0.008)	(0.034)				(0.003)
Rank 4	−0.181												0.060
	(0.063)												(0.003)
Industrial Sector													
Primary	−0.241	−0.006	−0.008	0.003	0.000	−0.009	0.106	0.034	0.005	0.051	−0.008	−0.004	−0.034
	(0.048)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.018)	(0.009)	(0.037)	(0.005)	(0.005)	(0.006)	(0.003)
Service	0.400	0.009	0.008	0.002	−0.012	0.003	0.091	−0.038	0.017	−0.028	0.010	−0.095	−0.021
	(0.050)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.019)	(0.009)	(0.038)	(0.006)	(0.005)	(0.006)	(0.003)
Firm Size													
Medium	−0.373	−0.009	−0.010	0.001	0.021	−0.002	−0.080	0.042	−0.045	0.060	−0.024	0.082	−0.007
	(0.050)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.019)	(0.009)	(0.038)	(0.006)	(0.005)	(0.006)	(0.003)
Large	−0.553	−0.016	−0.012	0.004	0.033	−0.006	−0.063	0.068	−0.067	0.105	−0.052	0.094	−0.010
	(0.049)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.019)	(0.009)	(0.038)	(0.006)	(0.005)	(0.006)	(0.003)
Number of Insider Executives on the Board of Directors													
Large	−0.238	0.000	0.005	−0.003	0.004	0.008	−0.023	−0.036	−0.095	−0.011	0.012	0.004	0.004
	(0.040)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.015)	(0.007)	(0.031)	(0.005)	(0.004)	(0.005)	(0.003)
Turnover													
New Employer	−0.380	0.001	0.008	−0.002	−0.004	0.004	−0.020	0.026	0.111	0.003	0.000	−0.001	0.002
	(0.040)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.015)	(0.007)	(0.031)	(0.005)	(0.004)	(0.005)	(0.003)

Note: Compensation is measured in millions of 2006 US$; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is equal to one if the executive is on the board and zero otherwise.

TABLE 9 : COMPENSATION FOR MARKET DEMAND VERSUS EXIT

Variable	Constant	Age-50	Tenure	Exec. Exp.	NBE	NAE	Female	Interlock	Execdir	No College	MBA	MS	PhD
Constant	−0.569	−0.003	−0.007	0.002	−0.003	−0.010	0.069	−0.034	−0.320	0.036	0.010	−0.009	0.014
	(0.013)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.014)	(0.008)	(0.029)	(0.003)	(0.003)	(0.002)	(0.003)
Rank 1	−0.151						−0.219	0.094	0.458				−0.058
	(0.013)						(0.013)	(0.007)	(0.027)				(0.002)
Rank 2	0.022						−0.181	0.050	0.486				−0.058
	(0.013)						(0.013)	(0.006)	(0.022)				(0.002)
Rank 3	0.019						−0.050	0.050	0.486				−0.058
	(0.013)						(0.013)	(0.006)	(0.022)				(0.002)
Rank 4	0.182												−0.058
	(0.013)												(0.002)
Industrial Sector													
Primary	0.048	0.002	0.004	−0.003	0.001	0.006	−0.124	−0.009	0.017	−0.029	−0.006	0.019	0.032
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)	(0.003)	(0.003)	(0.002)	(0.002)
Service	−0.006	−0.002	0.001	−0.001	0.004	−0.001	−0.045	0.011	0.006	−0.011	0.003	0.042	0.021
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)	(0.003)	(0.003)	(0.002)	(0.002)
Firm Size													
Medium	0.032	0.001	0.002	−0.002	−0.010	0.000	0.029	0.009	0.010	−0.010	0.003	−0.022	0.011
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)	(0.003)	(0.003)	(0.002)	(0.002)
Large	0.170	0.005	0.000	−0.004	−0.020	0.003	−0.003	−0.044	−0.010	−0.046	0.022	−0.033	0.011
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)	(0.003)	(0.003)	(0.002)	(0.002)
Number of Insider Executives on the Board of Directors													
Large	−0.117	0.000	−0.008	0.002	−0.007	−0.009	0.000	0.002	0.052	0.001	−0.005	−0.008	−0.008
	(0.008)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.010)	(0.005)	(0.020)	(0.002)	(0.002)	(0.002)	(0.002)
Turnover													
New Employer	−0.085	0.000	−0.013	0.003	−0.007	−0.010	−0.006	−0.060	−0.111	−0.012	0.007	−0.006	−0.019
	(0.008)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.010)	(0.005)	(0.020)	(0.002)	(0.002)	(0.002)	(0.002)

Note: Compensation is measured in millions of 2006 US$; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is equal to one if the executive is on the board and zero otherwise

Table 10: Value of Human Capital Investment

Variable	Constant	Age-50	Exec. Exp.	NBE	NAE	Female	Interlock	Execdir	No College	MBA	MS	PhD
Constant	−0.2278	0.0013	0.0014	0.0058	0.0050	0.0182	−0.0015	0.0004	−0.0090	−0.0100	0.0055	−0.0043
	(0.0002)	(0.0005)	(0.0006)	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Rank 1	0.0237			0.0001	0.0003	0.0033	0.0001	0.0000	−0.0015	0.0003	−0.0006	0.0002
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0006)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)
Rank 2	−0.0632			0.0006	0.0007	0.0017	−0.0003	0.0000	−0.0003	−0.0005	0.0007	−0.0012
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Rank 3	−0.0372			0.0012	0.0012	0.0070	−0.0002	−0.0001	−0.0027	−0.0010	0.0002	−0.0002
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Rank 4	−0.0062			0.0005	0.0005	0.0026	0.0000	−0.0001	−0.0016	−0.0006	−0.0001	0.0001
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Turnover												
New Employer	−0.0132			0.0006	0.0006	0.0036	−0.0001	−0.0001	−0.0014	−0.0005	0.0002	−0.0003
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)

Note: Compensation is measured in millions of 2006 US$; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is equal to one if the executive is on the board and zero otherwise.

Table 11: Risk Premium From Agency

Variable	Constant	Age-50	Tenure	Exec. Exp.	NBE	NAE	Female	Interlock	Execdir	No College	MBA	MS	PhD
Constant	0.499	−0.046	−0.019	−0.012	0.032	0.190	−0.268	−0.333	−0.507	−0.178	0.035	−0.059	0.128
	(0.736)	(0.005)	(0.004)	(0.002)	(0.011)	(0.005)	(0.195)	(0.049)	(0.224)	(0.026)	(0.017)	(0.029)	(0.017)
Rank 1	0.569					0.000	−0.660	0.140	0.177	0.001	0.000	0.001	0.032
	(0.125)					(0.000)	(0.069)	(0.045)	(0.112)	(0.000)	(0.000)	(0.000)	(0.013)
Rank 2	2.836					−0.001	2.338	0.058	0.081	0.000	0.001	−0.001	0.033
	(0.125)					(0.000)	(0.069)	(0.037)	(0.092)	(0.000)	(0.000)	(0.000)	(0.013)
Rank 3	1.032					−0.002	−1.120	0.058	0.081	0.003	0.001	0.000	0.032
	(0.125)					(0.000)	(0.069)	(0.037)	(0.092)	(0.000)	(0.000)	(0.000)	(0.013)
Rank 4	−0.016					0.000	−0.003	0.000	0.000	0.002	0.001	0.000	0.032
	(0.125)					(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.013)
Industrial Sector													
Primary	−0.037	−0.001	−0.001	0.000	0.012	0.011	0.142	−0.013	0.083	0.025	−0.014	0.058	−0.017
	(0.096)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.061)	(0.039)	(0.100)	(0.023)	(0.015)	(0.026)	(0.012)
Service	0.379	−0.049	−0.003	0.010	0.035	−0.061	−0.595	0.749	0.639	0.325	−0.166	0.355	0.096
	(0.098)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.062)	(0.040)	(0.101)	(0.024)	(0.015)	(0.026)	(0.012)
Firm Size													
Medium	1.032	0.016	0.003	0.004	−0.033	0.007	0.513	−0.240	0.375	−0.042	0.094	−0.118	−0.014
	(0.098)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.062)	(0.040)	(0.101)	(0.024)	(0.015)	(0.026)	(0.012)
Large	3.350	0.030	0.004	0.001	−0.064	0.002	0.495	−0.312	0.531	−0.111	0.126	−0.291	0.010
	(0.097)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.061)	(0.040)	(0.101)	(0.024)	(0.015)	(0.026)	(0.012)
Number of Inside Executives on the Board of Directors													
Large	0.270	0.006	0.006	0.003	−0.022	−0.004	0.049	−0.085	0.088	−0.052	0.047	−0.015	−0.015
	(0.079)	(0.003)	(0.003)	(0.001)	(0.008)	(0.004)	(0.050)	(0.032)	(0.082)	(0.019)	(0.012)	(0.021)	(0.010)
Turnover													
New Employer	0.362	0.008	−0.003	−0.003	0.012	0.025	0.258	−0.053	−0.062	0.053	−0.014	0.053	−0.046
	(0.080)	(0.003)	(0.003)	(0.001)	(0.008)	(0.004)	(0.051)	(0.033)	(0.083)	(0.020)	(0.013)	(0.022)	(0.010)

Note: Compensation is measured in millions of 2006 US$; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is equal to one if the executive is on the board and zero otherwise.

Table 12: Gross Loss to Shareholders from not Providing Executive Incentives

	$E(x(1-g(x))$	New Employer	Female	Individual	Characteristics
Constant	33.5963	6.8678	1.7380	Interlocked	−3.0951
	(0.0367)	(0.0036)	(0.0263)		(0.0100)
Rank 1	−8.0575	1.0166	−1.5638	Execdir	−7.0620
	(0.0056)	(0.0395)	(0.0358)		(0.0051)
Rank 2	−4.2791	2.8547	−1.7018	Exec.Exp.	−0.1339
	(0.0057)	(0.0412)	(0.0359)		(0.0006)
Rank 3	−1.9994	3.3221	−1.5730	Exec.Exp. Sq	0.0001
	(0.0057)	(0.0440)	(0.0361)		(0.0001)
Rank 4	−0.9403	2.8096	−1.3255	Tenure	0.0012
	(0.0058)	(0.0455)	(0.0362)		(0.0005)
Rank 1 Lagged	−6.6667			Tenure Sq.	−0.0001
	(0.0096)				(0.0001)
Rank 2 Lagged	−8.1900			No College	−0.2616
	(0.0067)				(0.0050)
Rank 3 Lagged	−3.5289			MBA	0.0026
	(0.0080)				(0.0045)
Rank 4 Lagged	−0.4527			MS	−0.4054
	(0.0049)				(0.0047)
Industrial Sector				PhD	0.7338
Primary	−3.7273				(0.0049)
	(0.0042)			NAE	0.4477
Service	9.3501				(0.0018)
	(0.0043)			NBE	0.5651
Firm Size					(0.0015)
Medium	−12.9481		0.0093	Age-50	−0.0411
	(0.0044)		(0.0244)		(0.0005)
Large	−25.4104		0.0139	Age-50 Sq	0.0005
	(0.0044)		(0.0221)		(0.0001)
Number of Insider Board Members					
Large	−3.0350				
	(0.0035)				
Bond price	0.9026				
	(0.0021)				

Note: Gross loss to shareholders measured as a percentage of equity value; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is equal to one if the executive is on the board and zero otherwise

TABLE 13: THE NET COMPENSATING DIFFERENTIALS TO EXECUTIVES FROM WORKING VERSUS SHIRKING

Variable	Constant	Age-50	Age-50 Sq	Tenure	Exec. Exp.	NBE	NAE	Female	Interlock	Execdir	No College	MBA	MS	PhD
Constant	9.952	0.053	−0.001	0.110	0.015	−0.067	0.141	1.437	−0.930	−0.151	−0.518	0.250	−0.469	0.069
	(0.888)	(0.019)	(0.001)	(0.027)	(0.000)	(0.066)	(0.031)	(0.530)	(0.190)	(0.002)	(0.097)	(0.089)	(0.101)	(0.079)
Rank 1	1.029					−0.004	−0.004	−0.378	−0.070	0.018	−0.014	0.004	0.002	0.061
	(0.798)					(0.002)	(0.002)	(0.480)	(0.173)	(0.003)	(0.005)	(0.004)	(0.004)	(0.060)
Rank 2	0.759					0.000	0.000	−1.082	−0.058	0.015	−0.001	−0.004	0.016	0.046
	(0.798)					(0.002)	(0.002)	(0.481)	(0.144)	(0.003)	(0.005)	(0.004)	(0.005)	(0.060)
Rank 3	0.307					0.006	0.005	−1.716	−0.063	0.018	−0.027	−0.009	0.010	0.056
	(0.798)					(0.002)	(0.002)	(0.481)	(0.144)	(0.003)	(0.005)	(0.004)	(0.005)	(0.060)
Rank 4	0.039					−0.001	−0.003	−0.120	0.010	0.017	−0.014	−0.004	0.008	0.058
	(0.798)					(0.002)	(0.002)	(0.008)	(0.010)	(0.003)	(0.005)	(0.004)	(0.005)	(0.060)
Industrial Sector														
Primary	−2.599	−0.032	0.001	−0.040		−0.005	−0.080	−0.612	0.427		0.264	−0.164	0.188	−0.050
	(0.605)	(0.016)	(0.001)	(0.023)		(0.055)	(0.026)	(0.419)	(0.145)		(0.079)	(0.074)	(0.082)	(0.054)
Service	3.799	0.060	−0.001	0.080		−0.050	0.074	0.788	−0.616		−0.434	0.122	−0.562	0.030
	(0.628)	(0.017)	(0.001)	(0.024)		(0.057)	(0.027)	(0.427)	(0.149)		(0.082)	(0.076)	(0.085)	(0.055)
Firm Size														
Medium	−3.105	−0.073	0.002	−0.079		0.125	−0.061	−1.041	0.769		0.530	−0.211	0.619	0.054
	(0.628)	(0.017)	(0.001)	(0.024)		(0.057)	(0.027)	(0.427)	(0.149)		(0.082)	(0.076)	(0.085)	(0.055)
Large	−4.500	−0.096	0.002	−0.111		0.153	−0.105	−1.207	0.766		0.653	−0.306	0.645	0.038
	(0.621)	(0.016)	(0.001)	(0.024)		(0.056)	(0.027)	(0.425)	(0.148)		(0.081)	(0.075)	(0.084)	(0.055)
Number of Inside Executives on the board of directors														
Large	−2.182	0.015	−0.001	−0.027		−0.056	−0.077	−0.415	0.149		−0.032	0.076	−0.014	−0.078
	(0.508)	(0.013)	(0.001)	(0.019)		(0.046)	(0.022)	(0.347)	(0.121)		(0.066)	(0.062)	(0.069)	(0.045)
Turnover														
New Employer	−4.755	0.051	−0.001	−0.052		−0.187	−0.189	−2.485	0.318	0.034	0.040	0.076	0.007	−0.259
	(0.514)	(0.013)	(0.001)	(0.019)		(0.048)	(0.023)	(0.355)	(0.130)	(0.003)	(0.071)	(0.066)	(0.073)	(0.049)

Note: Compensation is measured in millions of 2006 US$; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is equal to one if the executive is on the board and zero otherwise.

Table 14: Career Concern Amelioration of Agency Problem

Variable	Constant	Age-50	Age-50 Sq	Tenure	Exec. Exp.	NBE	NAE	Female	Interlock	Execdir	No College	MBA	MS	PhD
Constant	−1.547	0.006	0.001	0.009	0.015	0.059	0.050	0.154	0.170	−0.151	−0.080	−0.106	0.079	−0.044
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Rank 1	0.013					−0.004	−0.004	0.061	0.010	0.018	−0.014	0.004	0.002	0.002
	(0.006)					(0.002)	(0.002)	(0.008)	(0.009)	(0.003)	(0.005)	(0.004)	(0.004)	(0.005)
Rank 2	−0.490					0.000	0.000	−0.198	0.011	0.015	−0.001	−0.004	0.016	−0.012
	(0.006)					(0.002)	(0.002)	(0.009)	(0.010)	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)
Rank 3	−0.671					0.006	0.005	0.182	0.007	0.018	−0.027	−0.009	0.010	−0.002
	(0.006)					(0.002)	(0.002)	(0.009)	(0.010)	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)
Rank 4	−0.242					−0.001	−0.003	−0.120	0.010	0.017	−0.014	−0.004	0.008	0.000
	(0.006)					(0.002)	(0.002)	(0.008)	(0.010)	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)
Turnover														
New Empoyer	−0.101					−0.017	−0.019	−0.150	0.023	0.034	0.019	0.008	0.018	−0.002
	(0.006)					(0.002)	(0.002)	(0.008)	(0.009)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)

Note: Compensation is measured in millions of 2006 US$; Standard error in parentheses; Tenure and Executive Experience (Exec. Exp) are measured in years; NBE (NAE) is the number of times the executive changed firms before (after) becoming one of the ranks in our sample. Execdir is equal to one if the executive is on the board and zero otherwise.

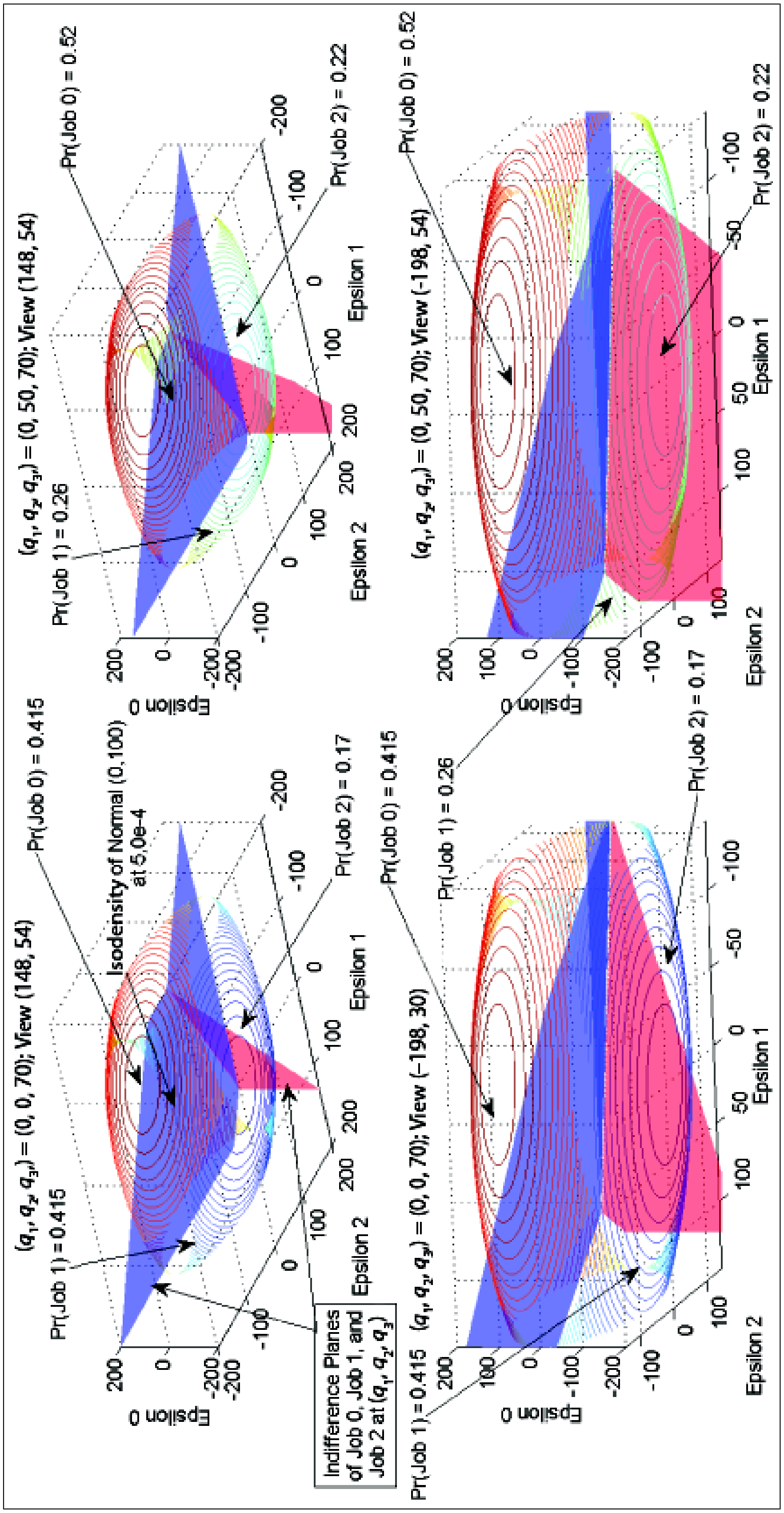

Figure 1: Equilibrium Sorting Using the Inversion Theorem