Interlocked Executives and Insider Board Members: An Empirical Analysis. *

George-Levi Gayle
Department of Economics, Washington University in St. Louis and St. Louis Federal Reserve Bank

Limor Golan
Department of Economics, Washington University in St. Louis and St. Louis Federal Reserve Bank

Robert A. Miller
Tepper School of Business, Carnegie Mellon University

November 27, 2015

Abstract

This paper asked the question of whether the behavior and compensation of interlocked executives and non-independent board of directors are consistent with the hypothesis of governance problem or whether this problem is mitigated by implicit and market incentives. It then analyzes the role of independent board of directors. Empirically, we cannot reject the hypothesis that executives in companies with a large number of non-independent directors on the board receive the same expected compensation as other executives. In our model, every executive has an incentive to work. Placing more of non-independent directors on the board mitigates gross losses to the firm should any one of them shirk because they monitor each other. It also reduces the net benefits from shirking 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). Therefore having a greater non-independent director representation on the board create a more challenging signaling problem to solve thereby raising the risk premium. However, giving more votes on the board to non-independent executives fosters better executive working conditions, which in turn offsets the higher risk premium in pay by a lower certainty-equivalent wage in equilibrium. 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. We then conduct counterfactual policy experiment imposing 50% upper bound on the fraction of insiders on the board and another counterfactual imposing 40% quotas for women on the boards.

*The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

1 Introduction

Measures of executives' network and the independence of their board of directors – e.g. interlocked[1] and the number of insiders on the board of directors – are positively correlated with executive compensation. These positive correlations are often interpreted as evidence of corporate governance problem with the management of publicly traded companies (Hallock, 1997; Bebchuk and Fried, 2003, 2005). Formally Holmstrom (1999) defined the existence of a governance problem as arising when contracts are incomplete. However, as pointed out in Fama (1980) the existence of incompleteness does necessarily give rise to a governance problem as implicit and market incentives – this is career and reputation concerns – will cause the agent to act more in his principal interest than the standard agency approach might suggests. Therefore well networked executives and large insider boards could be a form of implicit incentives (Lee and Persson, 2011) and hence not a sign of governance problem. This paper asked the question of whether the behavior and compensation of interlocked executives and non-independent board of directors are consistent with the hypothesis of governance problem or implicit and market incentives. It then assesses the role of independent boards by conducting counterfactual policy analysis.

There is a large theoretical and empirical literature on what directors do and how Boards of directors should be structured (Hermalin and Weisbach, 1998; Yermack (1996); Dominquez-Martinez et. al (2008); see Adams, Hermalin, and Weisbach, 2010 for survey). The theoretical models in these papers rely on the assumption - as in Holmstrom (1999)- that contracting is necessarily incomplete while the empirical studies uses measures of executive's bargaining power, tenure, and the executive's share holdings and correlated them with measures of board independence. None of these papers however has taken as theoretical model that embodies explicit, career, reputation, and market incentives and empirically quantify the relative importance of these different elements across board and executive network structure. Also the empirical evidence on whether these measures of board independence and executives are consistent with the "capture" or implicit incentives hypothesis are mixed. For example, Hallock (1997) findings that interlock is positively correlated with compensation however after controlling for firm size this correlation is no longer significant. He interpreted in support of the "capture" hypothesis via an empire-building hypothesis. On the other hand Gilson (1990) and Kaplan and Reishus (1990) find that poorly performing CEOs are less likely to gain board seats on other companies than well-performing CEOs. Hence lending support to the career and reputation concerns hypothesis of well-networked executives. Also Rose and Shephard (1997) empirically do not find any support for the entrenchment hypothesis. Empirical study of boards suffers from a number of difficulties. First, one must deal with broader than ideal classifications of directors. Second, nearly all variables of interest are, as discussed, jointly endogenous. Ultimately, much of what one learns about boards is about equilibrium associations. Causality, in the usual sense, is often impossible to

[1] A executive is interlocked if he serves (i) on the board committee that makes his compensation decisions; (ii) 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; or (iii) 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.

determine. To this overcome these concerns this paper develops a general equilibrium model of agency with both career and reputation concerns. It shows that, although board structure arises endogenously as economic actors choose them in response to the governance issues they face; structural differences across boards and type of executives' network have difference implications for executives' behaviors and compensation structure. These differences in behaviors and compensation structure can be used to separate the influence of explicit incentives, career and reputation concerns, and market forces across difference board and executive's network structures.

The model builds on the Gayle, Golan and Miller (2015). Firms are modeled as multilateral contracts between independent agents – executives at in different positions– and the principals with different span of control over the firm's outcomes. This view of the internal organizations of a firm was put forward by Alchian and Demsetz (1972) and Mirrless (1976) and was shown in Gayle, Golan, and Miller (2015) to empirically match the organization of publicly traded companies in the USA. Therefore the directors sign contract with the shareholders as just the other executives. The directors' job could include monitoring the CEO and the CEO could monitor the lower level executives in the C-suite but it is not a chain of command model like in Williamson (1967) and Calvo and Wellisz (1980) in which the directors have a contract with the shareholders and the CEO have a contract with the board of directors. There exist the standard moral hazard problem with call for a second best incentive contract. However, this is supplemented with career and reputation concerns as suggested by Fama (1980). This is modeled as human capital accumulation problem which is priced in a market equilibrium, however, if the executive or directors shirks – not act in the shareholders interest- no human capital is accumulation. However, as the effort the hidden action of the executives and directors this gives rise to also private information. Therefore in equilibrium the market beliefs of about the executives and directors human capital becomes important. We close the model using a sequential equilibrium concept, which requires that these beliefs be consistent in equilibrium. As entrenchment is normally measured by tenure in the firm and experience in similar firm these are endogenous to model through promotion and turnover.

The paper uses detailed data on the compensation and the structure of compensation. It include details on the history of all executives and directors therefore we observe promotion, turnover and exit along with the position and compensation structure in the new firm. This allows us to empirically document that the interlocked executives are more entrenchment, i.e. the have longer tenure in the firms, turnover less, are less likely to retire and more likely to eventually becomes a member the board of directors. We also documents that unconditionally the interlocked executives are paid more however, the structure of their incentive pay is also substantially different from non-interlocked executives. Moreover, once we control for the rank and firm characteristics they are paid less on average. We document that being an interlocked executive or on the board of directors also reduces the probability of exit by 55 and 65 percent respectively. Also being a board member increases the probability of being/becoming CEO. 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. Furthermore executives on the board are paid a premium of about $845,000, but are also more affected by firm abnormal returns. Finally, compensation is more closely tied to firm's performance in firms with more insider board members, and for interlocked executives.

We then structurally estimated the model. The estimation results show that once we adjust for the risk, the certainty equivalent pay in these firm is substantially lower than firms with independent boards ($380k versus $740k). This result is consistent with better work conditions and good governance of firms with boards that are less independent. The risk premium, however, is larger by $280k on average in firms with large number of insiders on the board. Analyzing the risk premium, which is the cost of agency in our model, shows that the net benefit of shirking is lower when there are large number of insiders on the board. This can indicate that monitoring is more efficient in firms with less independent boards and therefore the goals of the executives are more aligned with the goals of the shareholders. The reason that we find that the risk premium is higher in these firms is because the quality of signals of productivity when an executive shirks while the rest of the executives are working diligently is lower. Put it another way, shirking of one executive causes less destruction in these firms which may imply better governance but also implies that the risk premium is higher for executives to work diligently.

At the higher ranks, executives 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 measures of networking opportunities reduce the nonpecuniary costs of a job match, and hence its equilibrium compensation.

Comparing the certainty equivalent pay confirms that the gap between their compensation and the compensation of other executives would be a lot smaller if it was not for the risk component. 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. 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 non-interlocked executive, $710,000. 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. 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 insiders dominate the board, it falls by $2.2 million. Interlocked executives places lower value on career concerns, executive directors higher. However, overall the role of implicit incentives are small even for executive directors. We therefore conclude that the main difference in the pay structure for these executives is the higher costs of agency due to the relative difficulty in assessing their performance relative to other executives, therefore, explicit incentive provided by compensation contract is the most important tool for shareholders to align incentives of board members and interlocked executives with their own goals.

Whether the promotional structure and compensation incentives make for good governance has recently drawn the attention of several empirical researchers, namely Rose and Shephard (1997), Hallock (1997), Bebchuk and Fried (2005), Kuhnen and Zwiebel (2009), Acharya and Volpin (2010), and Dicks (2012). Our study makes a theoretical and an empirical contribution on this topic. Theoretically we have a model in which entrenchment arises endogenously. 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. 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, reduces the net benefits from shirking, 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 signaling problem to solve, thereby raising the risk premium; giving more votes to executives fosters better executive working conditions, which in turn is offset by a lower certainty-equivalent wage in equilibrium. 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.

We find that female executives are less likely to be board members and interlocked executives (with the exception of females in rank 3 which are more likely to be interlocked). We find 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 findings suggest some non-trivial implications of imposing female quotas for boards, despite the fact that we find no evidence for differences in productivity of male and female executives.

Lastly we perform two counterfactual policy exercises. The first one imposes quotas of 40% women in boards (as the policy rules in Norway). The second counterfactual imposes a rule of no more than 50% insiders on boards.

2 Data

The data for our empirical study are 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 titles. Each executive worked for one of the 2,818 firms comprising the (composite) S&P 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 the eight top-paid employees in the firm whose compensation was reported to the Securities and Exchange Commission (SEC). Data on the 2,818 firms for the ExecuComp database were supplemented by the COMPUSTAT North America database and monthly stock price data from the Center for Research in Security Prices 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 Marquis Who's Who, which contains biographies of about 350,000 executives. The matched data provide 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, and education; and a comprehensive description of their career path sequence described by their annual transitions through the possible positions and firms.

We construct a hierarchy consisting of five ranks using a rational ordering over a set of job titles based on transition independent of compensation.[2] Rank 1 includes 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 is the CEO of the company. Rank 3 includes the COO, the CFO, and the chairman of the board of the company if that person holds an executive position in the company other than CEO. Other high-level corporate executives and heads of subsidiaries or regional chiefs comprise Rank 4, while Rank 5 is reserved for lower-level executives. Thus, CEOs are not in Rank 1. Since this hierarchy is based on transitions, the ranking reflects lifecycle considerations, not power or control. The ranking corroborates the institutional use of the term, which emphasizes the supervisory roles of managers over their subordinates. For example, the chairman of the board of directors monitors the CEO of the firm. Interlock and Large Insider Board

Following the literature on corporate governance we construct two measures of governance and executives 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 is 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.

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

[2] The method for constructing the hierarchy, and a detailed description of the titles in each rank, is in Gayle, Golan, and Miller (2012).

We classify firms into three industrial sectors: primary, consumer, and service. Firms are also partitioned by size – large, medium, and small – based on the value of their assets and number of employees over the sample period. A firm is defined as large if both its value of assets and its number of employees are above the median for its sector over the sample period, and as small if both its value of assets and number employees are below the median for its sector over the sample. All other firms are medium sized. We further classify firms by the number of "insiders" on their board relative to the industrial norm. A company is defined as having a large insider board if the number of insiders on its board is above the median for its sector and firm size. Finally, reflecting our focus on executive compensation, firms are classified from the perspective of their executives: New if this is the first year the executive is working in the firm and old if the executive has worked in the firm for more than one year. This variable allows us to capture the effects of executive turnover. Summarizing, there are 36 firm types, differentiated by size, industrial sector, importance of insiders on the board, and whether the executive in question has just joined the firm. 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 executives holding firm options and changes in wealth from holding firm stock relative to a well-diversified market portfolio.[3] 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.

Individual characteristics consist of several dimensions of labor market experience, some demographic background variables, and whether the executive is interlocked.[4] Variables we construct on labor market experience include years of tenure with the firm, years worked as top executive, number of firms in which an executive worked before becoming an executive, and the number of firms in which an executive worked after becoming an executive. We also observe educational qualifications (including MBA, MSc, PhD), gender, and age. Finally, since the price of console bonds plays a role in consumption smoothing in our model, we construct a bond price series from the Federal Reserve Economic Database (FRED). Online Appendix B contains a full description of the construction and a data summary.

[3] Changes in wealth from holding firm stock and options reflect the cost a manager incurs from not being able to fully diversify her 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. See Antle and Smith (1985, 1986), Hall and Liebman (1998), Margiotta and Miller (2000), and Gayle and Miller (2009a,b) for other papers using this measure of total executive compensation.

[4] An executive is classified as interlocked if at least one of the following is true: (i) The executive serves on the board committee that makes her compensation decisions. (ii) The executive serves on the board of another company that has an executive officer serving on the compensation committee of the indicated executive's company. (iii) The executive serves on the compensation committee of another company that has an executive officer serving on the board of the indicated executive's company.

2.1 Preliminary Analysis of the Data

This section documents regularities relating governance, ranks, board membership and being interlocked executives to characteristics such as education experience and entrenchment (exit, turnover and tenure) to compensation. Table 1A and 1B document basic facts on the pay by ranks of executive directors, interlocked executives and firm with large fraction of insiders on the board compared to other executives and firms with small fraction of insiders on the board. The salary component is larger for executive directors than salaries of other executives in all ranks. The salaries of interlocked executives are smaller with the exception of interlocked executives in ranks 4 and 5; however, the majority of interlocked executives are in rank 2. Salaries in all ranks are higher in firms with large insider board. However, unlike the previous literature on interlocked executives we find that unconditionally interlocked executives are paid \$ 3.7 million while non-interlocked executives are paid \$ 2.5 million. This highlights the need to control for rank and the structure of firm. The table presents the different components of compensation. Total compensation of executive directors is larger in general (with the exception of rank 1) than the compensation of other executives. With the exception of ranks 4 and 5 the total compensation of interlocked executives is smaller than that of other executives. With the exception of rank 3 total compensation in firms with large insider board is lower than that of the average pay of executives in firms with smaller insider fraction on the board. To further analyze the pay differentials, tables 2A and 2B document the personal attributes of executives by rank. Interlocked executives and executives directors are older than other executives and there are no age differences by board structure. Interlocked executives and executives directors have more tenure, experience as executives relative to other executives. Retirement probabilities of board members and interlocked executives are lower and they are more likely to transition to other firms (Table 2B). This facts are consistent with these executives being more entrenched. With the exception of level 5, executives in firms with larger fraction of insiders on the board are also more experienced and have more tenure in the firms, they are less likely to quit the executive occupation or move to another firm. These facts can be related to governance quality and our model rationalizes it in the context of an equilibrium model.

Table 3 presents the governance structure by sector, firm size and ranks. Board members are drawn mainly from the top 3 ranks. In all firms the likelihood of becoming executive director is highest for executives at ranks 1 and 2. Executives are more likely to become board members in smaller firms and in the primary sector. The likelihood of becoming interlocked executive is small in all ranks and in all types of firms. It is also larger for executives in the consumer sector (10.5% for rank 1 executives in this sector) and in small firms (15% for rank 1 executives).

To further explore whether the compensation, turnover and rank transitions patterns found in the previous tables are explained by observable characteristics we run regression of compensation, and logit regressions of within firm rank transitions and turnover on the individual and firm characteristics observed in the data. Table 4 presents the results. The compensation regression reveals that controlling for executives characteristics and positions. The results reveal that executives directors have larger fixed pay (by \$845k) on average, interlocked executives have smaller fix pay (by \$299k), and executives in firms with large insider boards have higher fixed pay (by $280k). The coefficients on the excess return and its square represent exposure to risk. Because of the non-linearity of the pay schedule in excess return we include graphs of the pay schedules of executive directors, interlocked executives and executives in firms with large insiders board. Our estimation results from the model, will enable us to recover the risk aversion parameters and compute the certainty equivalent pay in these positions. The logit regressions confirm that executive directors, interlocked executives and executives in firms with large insiders board are less likely to retire, or switch firms, and therefore are more entrenched, even once we control for age, experience and tenure (as well as other characteristics).

3 The Model

The model builds on the Gayle, Golan and Miller (2015). Firms are modeled as multilateral contracts between independent agents – executives at in different positions– and the principals with different span of control over the firm's outcomes. This view of the internal organizations of a firm was put forward by Alchian and Demsetz (1972) and Mirrless (1976) and was shown in Gayle, Golan, and Miller (2015) to empirically match the organization of publicly traded companies in the USA. Therefore the directors sign contract with the shareholders as just the other executives. The directors' job could include monitoring the CEO and the CEO could monitor the lower level executives in the C-suite but it is not a chain of command model like in Williamson (1967) and Calvo and Wellisz (1980) in which the directors have a contract with the shareholders and the CEO have a contract with the board of directors.

There exist the standard moral hazard problem with call for a second best incentive contract. However, this is supplemented with career and reputation concerns as suggested by Fama (1980). This is modeled as human capital accumulation problem which is priced in a market equilibrium, however, if the executive or directors shirks – no act in the shareholders interest- no human capital is accumulation. However, as the effort the hidden action of the executives and directors this gives rise to also private information. Therefore in equilibrium the market beliefs of about the executives and directors human capital becomes important. We close the model using a sequential equilibrium concept, which requires that these beliefs be consistent in equilibrium. As entrenchment is normally measured by tenure in the firm and experience in similar firm these are endogenous to model through promotion and turnover.

According to Fama (1980), insider executives on the board are the top managers and they are the most informed and responsive critics of the firm. The role of outside directors is similar to the role of professional referees who are impartial and can prevent collusion of the insiders on the board. They are disciplined by the market prices for their services. The model in this section utilizes the framework developed in Gayle, Golan and Miller (2015) which includes the building blocks of the model are moral hazard, sorting, nonpecuniary benefits from jobs, human capital and career concerns to capture the role of the board and explain compensation of board members and interlocked executives. Specifically, the model captures different factors affecting assignment and compensation of executives to firms with heterogenous insider board size. The building blocks of the model are parsimoniously combined to allow estimation of the underlying technology- and utility-function parameters rationalizing the observed compensation schedule for executive directors and interlocked executives as well as assess the hypotheses about the role of the composition of insiders and outsiders on the board in different firms in monitoring the firm. Sorting and human capital and modeling careers of executives captures the sorting of executives into board positions. Turnover and career concerns captures the role of market prices in the observed compensation of board members. Interlocked executives and executive directors are accounted for as different positions. Accounting for differences in board structure across firms, together with executives' career concerns and life-cycle and compensation allows us to explain compensation structure, and test for entrenchment. We begin by formulating the model, expected-value-maximizing shareholders are subject to moral hazard from choices made by risk-averse managers, who have private information about their own effort levels. Since human capital accumulated depends on effort levels, it is also private information of the executive. For analytical tractability we assume a complete set of markets exists for publicly disclosed events. Managers accumulate both firm-specific and general human capital through experience on the job. Managers sequentially choose employment, bargain with firms about their compensation, and choose their effort levels, which determine the probability distribution of the returns to the firms. Through this process, managers extract all the rent from their job matches.

3.1 Executives and Firms

A finite number of firms in the executive market are indexed by $j \in \{1, \ldots, J\}$, with $j = 0$ representing retirement. There are K positions within each firm j, indexed by $k \in \{1, \ldots, K\}$ and ranked in hierarchical order. Board member and interlocked executives are considered two different ranks. The different combinations of firms and ranks captures heterogeneity of jobs in the economy. Firms belong to different industries and have different sizes of capital and employment. Thus, the position of a CEO who is also interlocked in a large firm with large insiders on the board, for example, may be different from a CEO position who is not interlocked in a firm with small insiders board, in terms of the tasks performed, skill requirements and nonpecuniary benefits and costs. Let $t \in \{0, 1, \ldots\}$ denote each executive's age, with retirement upon reaching or before age $T < \infty$. To simplify the notation, we assume that executives are infinitely lived. Each manager's background is defined by age t and a vector of human capital h_t, which includes fixed demographic characteristics and indexes work experience.

3.2 Choices

At the beginning of period t, managers choose consumption, c_t, and, for any $t \leq T$, they make employment choices. They negotiate their compensation and sign an employment contract determining how they will be paid. They then choose their effort, which is unobserved by the shareholders. Let $d_{jkt} \in \{0, 1\}$ indicate the manager's choice of 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, working diligently and shirking, denoted by $l_t \in \{0,1\}$, where $l_t = 0$ means the manager shirks at age t and $l_t = 1$ means the manager works. Effort affects the distribution of the firm's returns and the manager's current-period nonpecuniary utility. As in standard moral-hazard models, the goals of the managers and shareholders are not aligned. Therefore, the term shirk refers to managers' activities that benefit the managers but not the shareholders, and working diligently describes effort and activities that are taken to achieve the shareholders' goals.

3.3 Preferences

A manager's preferences depend on her consumption and nonpecuniary utility associated with labor-supply choices. Preferences are characterized by the discounted sum of a time-additively separable, constant absolute risk-aversion (CARA) utility function. The utility function is decomposed into utility from consumption and a nonpecuniary cost-of-working. The nonpecuniary costs of working and shirking is allowed to be different in each rank and firm and is further decomposed into systematic and nonsystematic components. The nonsystematic component captures the manager's firm- and rank-specific idiosyncratic-taste shock, which does not depend on effort. The taste-shock vector is denoted by $\varepsilon_t \equiv (\varepsilon_{0t}, \varepsilon_{11t}, \ldots, \varepsilon_{JKt})$, where the retired manager's taste shock is given by $-\varepsilon_{0t}$, and the taste shock from working in firm j at rank k is ε_{jkt}. The systematic component of the nonpecuniary utility from working depends on the manager's effort, characteristics and experience h, as well as the firm and rank. When $l_t = 1$, the nonpecuniary cost of working is $\alpha_{jkt}(h)$; when $l_t = 0$, work cost is $\beta_{jkt}(h)$. The manager's lifetime utility can thus be summarized as

$$-\sum_{t=1}^{\infty} \delta^t \exp(-\rho c_t) \left[d_{0t} \exp(-\varepsilon_{0t}) + \sum_{j=1}^{J} \sum_{k=1}^{K} d_{jkt} [\alpha_{jkt}(h_t) l_t + \beta_{jkt}(h_t)(1-l_t)] \exp(-\varepsilon_{jkt}) \right], \tag{2}$$

where δ denotes the subjective discount factor and ρ is the constant absolute risk-aversion parameter. The nonpecuniary benefits from retirement are normalized to be equal to 1. We assume there is more disutility from working than from shirking, so $\alpha_{jkt}(h) > \beta_{jkt}(h)$. The difference between $\beta_{jkt}(h_t)$ and $\alpha_{jkt}(h_t)$ captures the divergence between the shareholders' and managers' goals. This formulation of the utility function captures differences across rank–firm nonpecuniary costs, which allows the model to account for different levels of moral hazard between large and small firms and among ranks and industries. The formulation also allows managers with different characteristics to have different disutilities from firm–rank and effort choices. The CARA utility function is commonly used in the literature because the lack of wealth effects makes the dynamic problem more tractable (see Grossman and Hart, 1983; Malcomson and Spinnewyn, 1988; Fudenberg, Holmström, and Milgrom, 1990; Rey and Salanie, 1990). In addition, our data does not include managers' wealth, and the taste shocks are determined independently of effort level and are multiplicatively separable, implying that the log of the indirect utility is linear and additively separable.

3.4 Human-Capital Accumulation and Effort

Human capital is multidimensional and includes skills that depend on education and work experience. We define a vector of time invariant characteristics and skills, h_1, that captures gender and education dummies. We further define a vector to capture the individual's history of rank–firm choices, including retirement, as $h_{2t} = (h_{211t}, \ldots, h_{2JKt})$. Thus, the vector that captures all human capital is $h_t = (h_1, h_{2t})$. We assume that h_{2t} is a three-dimensional vector, $h_{2t} \equiv \left(h_{2t}^{(1)}, h_{2t}^{(2)}, h_{2t}^{(3)}\right)$ and that firm is the cross between two sets of indices: The first, $j_1 \in \{0, 1\}$, denotes whether this is a new firm, $j_1 = 0$, or the executive worked for this firm last period, $j_1 = 1$. The second, $j_2 \in \{1, 2, \ldots, J_2\}$, denotes firm size and industrial sector; hence, $j = j_1 \otimes j_2 \in \{0, 1\} \otimes \{1, 2, \ldots, J_2\}$. Therefore let $h_{2t}^{(1)}$ measure tenure of the executive in the current firm and, hence, capture firm-specific capital. Let $h_{2t}^{(2)}$ measure the number of years of executive experience, and let $h_{2t}^{(3)}$ measure the number of different firms the executive has worked in since becoming an executive. The last two, $h_{2t}^{(2)}$ and $h_{2t}^{(3)}$, are meant to capture the years of general human capital. The second, $h_{2t}^{(2)}$, is standard in the learning-by-doing human-capital–accumulation literature; however, the third, $h_{2t}^{(3)}$, is meant to capture the idea that management may require many different skills and networking, and the greater the number of firms an executive worked in the better or more connected she may be when she becomes, say, CEO or a board member. When the executive works diligently, $l_t = 1$, in rank k of the jth firm works each component in the human capital vector described above is augmented by $\Delta_{jkt} \equiv \left(\Delta_{jkt}^{(1)}, \Delta_{jkt}^{(2)}, \Delta_{jkt}^{(3)}\right)$. When she shirks, $l_t = 1$, it is augmented by $\underline{\Delta}_{jkt} \equiv (\underline{\Delta}_{jkt}^{(1)}, \underline{\Delta}_{jkt}^{(2)}, \underline{\Delta}_{jkt}^{(3)})$. Therefore, human capital evolves according to the transition function

$$h_{2t+1} = \sum\nolimits_{j=1}^{J} \sum\nolimits_{k=1}^{K} d_{jkt} \left[l_t (h_{2t} + \Delta_{jkt}) + (1 - l_t)(h_{2t} + \underline{\Delta}_{jkt}) \right]. \tag{3}$$

For executives who works diligently, $d_{0t} = 0$, and $l_t = 1$, if $j_1 = 0$, that is, the executive works in a new firm, $\Delta_{jkt}^{(1)} = -h_{2t}^{(1)}$, $\Delta_{jkt}^{(2)} = 1$ and $\Delta_{jkt}^{(3)} = 1$. This mean that the executive would lose all her firm-specific capital while still gaining an additional year of executive experience, but would increase the number of firms she worked in. On the other hand, if $j_1 = 1$, then $\Delta_{jkt}^{(1)} = 1$, $\Delta_{jkt}^{(2)} = 1$ and $\Delta_{jkt}^{(3)} = 0$. With this formulation of human capital, the executive is gaining firm-specific capital, but does not increase the number of firms she worked in. However, if she shirks, $l_t = 0$, and chooses a new firm, so $j_1 = 0$, then $\underline{\Delta}_{jkt}^{(1)} = -h_{2t}^{(1)}$, which is the same as if the executive had worked diligently. However, $\underline{\Delta}_{jkt}^{(2)} = 0$ and $\underline{\Delta}_{jkt}^{(3)} = 0$, meaning that the executive would lose all her firm-specific capital and does not gain an additional year of executive experience or an increase the number of firms she worked in if she shirks. On the other hand, if she chooses a j with $j_1 = 1$, then $\underline{\Delta}_{jkt}^{(1)} = 0$, $\Delta_{jkt}^{(2)} = 0$ and $\Delta_{jkt}^{(3)} = 0$. Human capital is private information to the executive because her effort choice is not observed by the firm or the market.

3.5 Firm Technology

In this subsection alone, it is necessary to identify the executive pool explicitly, because firms may employ more than one executive in the same position. To distinguish between lifecycle effects and aggregate technological shocks we also track of workers' ages over calendar time. We now suppose there are $N_{j\tau}$ executives who sort themselves into positions at the j^{th} firm in period τ. Denote by $t(\tau, n)$ the age of the n^{th} executive at calendar time τ, and her human capital at τ by $h_{t(\tau,n)}$. Let $F_{jk}(h_{t(\tau,n)})$ denote the executive's contribution to the j^{th} firm's output in τ if she chooses the k^{th} job with that firm by setting $d_{jkt(\tau,n)} = 1$. Let $\pi_{\tau+1}$ denote a return from an exogenous aggregate productivity shock that affects every firm, and let $\pi_{j,\tau+1}$ denote the (net) excess return to the j^{th} firm. Let $\mathcal{E}_{j\tau}$ denote the value of firm j at the beginning of calendar time τ. Finally, denote by $w_{jk\tau+1}^{(n)}$ the firm's compensation to executive n if she worked at rank k in period τ. We assume the equity of the firm evolves according to the law of motion:[5]

$$\mathcal{E}_{j,\tau+1} \equiv \textstyle\sum_{n=1}^{N} \sum_{k=1}^{K} d_{jkt(\tau,n)} \left[F_{jk}(h_{t(\tau,n)}) - w_{jk\tau+1}^{(n)}\right] + \mathcal{E}_{j\tau}(\pi_{\tau+1} + \pi_{j,\tau+1}). \tag{4}$$

$\mathcal{E}_{j\tau}\pi_{\tau+1}$ is the return on equity attributable to aggregate productivity shocks, and $\mathcal{E}_{j\tau}\pi_{j,\tau+1}$ is the excess return to the firm, $\pi_{j,\tau+1}$, whose probability distribution depends on the effort of all the executives.[6] The first component of the output, the summed expression involving $F_{jk}(h_{t(\tau,n)})$, is additively separable in the productivity of each executive n. It is determined by h though past effort, which is unobserved. To simplify the notation and the equilibrium characterization, we make a further assumption that if $l_1 = 0$, then $F_{jk}(h_t) \equiv \underline{F}$ for all h_t.[7] This initial condition places an upper bound on output, ensuring that firms do not benefit from employing executives who shirked in their first period. This individual factor is deterministic, has a level effect on the executive's marginal product, and is independent of the individual's effort and other executives' characteristics and efforts. The second component, $\pi_{\tau+1}$, captures the effect of aggregate factors on the firm's equity. In standard moral-hazard models, the optimal contract does not depend on the market portfolio or aggregate factors the executive cannot affect, because they are risk averse and a contract depending on such factors imposes additional risk on them without providing any additional incentive. Assuming all dividends are paid when the firm is liquidated, We show that in equilibrium the expected compensation to an executive fully offsets his expected contribution to output. Setting the summed expression to zero and rearranging Equation (4) to make $\pi_{j,\tau+1}$ the subject then yields the standard definition of excess returns in the asset-pricing literature, $\mathcal{E}_{j,\tau+1} / \mathcal{E}_{j\tau} - \pi_{\tau+1}$.

Executive effort only affects the firm through the probability distribution determining $\pi_{j,\tau+1}$. We analyze an equilibrium where every executive works, in which case the value of $\pi_{j,\tau+1}$ is drawn from a probability density function denoted by $f_j(\pi)$. Consistent with standard asset pricing theory, we

---

[5] This formula can be easily modified to allow for dividends to be distributed throughout the life of the firm, but the firm's dividend policy does not affect the compensation paid to managers in our model.

[6] Here, we are abstracting from other costs faced by the firm, such as the wage bill for the nonexecutive work force, by implicitly accounting for them in $\mathcal{E}_{j\tau}\pi_{\tau+1}$.

[7] The human capital of an executive who did not shirk in the first period, but shirks later, evolves according to Equation (3).

normalize the expected value of abnormal returns in equilibrium from everyone working to zero.

If everyone except the k^{th} ranked executive works, conditional on any level of human capital h, the value of $\pi_{j,\tau+1}$ is on $f_j(\pi)g_{jk}(\pi\,|h_t)$. Thus the impact on production from an executive shirking is captured by $g_{jk}(\pi\,|h_t)$, the likelihood ratio for the density when the executive with h in position k shirks while all other executives work, and the density when all executives work diligently. Since equilibrium compensation depends on $\pi_{j,\tau+1}$, the k^{th} ranked executive realizes that if he was the only one to shirk, his expected compensation would depend on $f_j(\pi)g_{jk}(\pi\,|h_t)$, and this consideration ultimately explains why $f_j(\pi)g_{jk}(\pi\,|h_t)$ helps shape equilibrium compensation.

Let $f_{0j}(\pi)$ denote the probability density function for π when the combination of who works and who shirks is chosen to maximize its expected value subject to the constraint that at least two executives shirk. The precise functional form of $f_{0j}(\pi)$ is immaterial in an equilibrium where everyone works, because $f_{0j}(\pi)$ only generates π if two or more executives deviate from their equilibrium action.[8]

In our model a necessary condition for an equilibrium to exist where everyone works is that expected abnormal returns are maximized by everyone working. Formally we assume:

$$0=\int\pi f_j(\pi)\,\mathrm{d}\pi>\max\left\{\int\pi f_j(\pi)g_{jk}(\pi\,|h_t)\,\mathrm{d}\pi,\int\pi f_{0j}(\pi)\,\mathrm{d}\pi\right\}. \tag{5}$$

The potential for conflict between executive and shareholder goals arises in this model from the preferences of executives to shirk rather than work, that is $\alpha_{jkt}(h_t)>\beta_{jkt}(h_t)$, whereas the inequalities in (**??**) show production is greater when all executives work. The likelihood ratio $g_{jk}(\pi|h_t)$ measures the degree to which executive effort can affect a firm's returns, so we interpret it as a measure of their span of control. Since $g_{jk}(\pi|h_t)$ depends on rank within the firm and the type of firm, there is scope to testing how it varies with board structure, managers' human capital, tenure, sector and positions.

Effort is unobserved in our model but $\pi_{j,\tau+1}$ is a signal of effort. In this respect $g_{jk}(\pi|h_t)$ measure the quality of the signal. For example if $g_{jk}(\pi',h)=1$ for some π' then the signal is uninformative about effort. If there existed some π'' in the support of $f_j(\pi)$ such that $g_{jk}(\pi'',h)$ was arbitrarily large, then the signal would so informative that a first best allocation could be achieved, by heavily penalizing all executives if π'' occurs, and paying a constant wage otherwise. Since executives are not paid constant wages, we assume $g_{jk}(\pi|h_t)$ is bounded. We also impose the regularity condition:

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

Intuitively this condition states that if firm performance at the end of the period is truly outstanding, then shareholders are almost certain that all the executives have worked during the period. Our assumptions ensure the existence of an optimal contract with bounded compensation (Mirrlees, 1975),

[8] Margiotta and Miller (2000) assume the distribution of π is the same when two or more executives shirk. In their specialization $f_{0j}(\pi)$ is a primitive, namely the common probability density for all possible work/shirk combinations of the firm's executives when at least two shirk. In our framework we could develop notation for the density functions of all those possible work/shirk combinations, and state $f_{0j}(\pi)$ in terms of those primitives. However this would be a sterile exercise because not even $f_{0j}(\pi)$ is identified in an equilibrium where everyone works, let alone the functions from which it is derived.

and are clearly weaker than the common monotonicity assumption requiring $g_{jk}(\pi|h_t)$ to decline in π.

3.6 Capital Markets, Timing and Information

We assume there exists a complete contingent-claims market for consumption, including all publicly disclosed events.[9] Let ξ_t denote the manager's endowment at age t. We also measure $w_{jk,t(\tau)+1}$, the manager's compensation for employment in position k at firm type j at the beginning of age $t+1$, in units of current consumption. The manager's wealth is endogenously determined by her compensation and cannot be fully insured when compensation depends on the firm's returns $\pi_{j,\tau+1}$.

At the beginning of each period, the outcome of the previous period's production, $\mathcal{E}_{j,\tau}\pi_{j,\tau+1} + \sum_{k=1}^{K} F_{jkt(\tau)}(h_t)$, is observed, as well as the market return, $\pi_{\tau+1}$. The manager's compensation, $w_{jk,t(\tau)+1}$, is paid according to the contract signed. The manager observes her preference taste shocks, ε_{jkt}, at the beginning of the period. After observing the shocks, the manager chooses consumption, c_t; asset portfolio, ξ_t; whether to retire or not; and, if the manager decides not to retire, the firm, rank and effort level. After the manager chooses rank and firm, there is a negotiation stage between each individual manager and the shareholders. We assume that the manager makes a take-it-or-leave-it offer and the shareholders then accept or reject the offer. If no agreement is reached, the manager does not work during that period, and there is no additional hiring by the firm.

The taste shocks and effort level are assumed to be the manager's private information. All other information is symmetric. The shareholders of all firms observe the managers' consumption and asset choices, as well as their rank and firm choices d_{jkt}, and human capital h_t. Together with public disclosure laws, these assumptions are necessary to rule out anonymous trading in the assets of the employing firm. Otherwise, under an incentive scheme in which mangers' compensation is tied to $\pi_{j,\tau+1}$, the idiosyncratic risk imposed on the manger can be undone by trading contingency claims. Since $F_{jkt(\tau)}(h_t)$ cannot be separately observed and given that human capital is the executives' private information, $F_{jkt(\tau)}(h_t)$ is private information. We assume that all accepted and rejected contracts and employment histories are observed by all firms. This simplifies the off-equilibrium-path analysis by reducing the number of observed histories. To summarize, at the beginning of each period $\tau+1$, the market and managers observe the previous period's outcomes, $\mathcal{E}_{j,\tau}(\pi_{\tau+1} + \pi_{j,\tau+1} - 1) + \sum_{k=1}^{K} F_{jkt(\tau)}(h_t)$ and (ξ_t, h_t) while the manager observes $(\xi_t, h_t, \varepsilon_t, \{l_s\}_{t=0}^{t}, F_{jkt(\tau)}(h_t))$.

3.7 Intertemporal Consumption and Employment Choices

The separability of preferences, the executives' absolute risk aversion and the completeness of the capital market allows us to focus on the executives' indirect utility function, which maps their expected utility as a function of the relevant security prices, the portion of their wealth that can be fully diversified, the distribution of any unanticipated changes in their wealth induced by the undiversifiable component of their contingent compensation and the option value of their stock of human capital.

Let $h'_t = (h'_1, h'_{2t})$ denote shareholders' belief about a manger's human capital—that is, the

[9] The framework is developed in Margiotta and Miller (2000). Details are presented in Gayle, Golan and Miller (2014).

manager's reputation—while $h_t = (h_1, h_{2t})$ denotes the manager's actual human capital. If the executive shirks, firm returns are related to the manager's actual human capital and drawn from $g_{jkt(\tau)}(\pi \mid h) f_j(\pi)$, not $g_{jkt(\tau)}(\pi \mid h') f_j(\pi)$. Then conditional-choice probabilities depend on both the manager's actual human capital, h_t, and the manager's reputation, h'_t.

We denote the utility of the present value of compensation by

$$\upsilon'_{jk,t(\tau)+1} \equiv \exp\left(-\rho w_{jk,t(\tau)+1}(h'_t, \pi_t)/b_{\tau+1}\right). \tag{7}$$

Where b_τ denote the price of a bond that, contingent on the history through date τ, pays a unit of consumption from period τ in perpetuity in period-τ prices.

Next, we define the on and off-equilibrium path and choices. 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. Similarly, firms cannot definitively recognize past shirking because individual productivity, $F_{jkt(\tau)}(h_t)$, is not observed separately from the aggregate output of the executive team. The initially shirking executive's choices of job-match profiles do not reveal past shirking either. In the equilibrium, every job history has a strictly positive mass even if no shirking occurs along the equilibrium path. Underlying this result is the assumption that ε_{jkt} has full support and is private information. Therefore, when contracts are only offered for work, shareholders believe that h'_t follows the law of motion $h'_{t+1} = h'_t + \Delta_{jkt}$ in any given history. In truth, if a manager deviates and shirks at age t, her next-period human capital is $h_{t+1} = h_t + \underline{\Delta}_{jkt}$. The contract is based on the manager's reputation, h'_t, not the manager's actual human capital, h_t. Let $p_{jkt}(h, h')$ denote the probability of choosing (j, k) at age t conditional on h. Similarly, we denote the retirement probability by $p_{0t}(h, h')$.

Theorem 3.1 *If the employers' beliefs follow $h'_{t+1} = h' + \Delta_{jkt}$, then job matches d_t and effort levels l_t are picked to sequentially maximize*

$$\varepsilon_{0t} d_{0t} + \sum_{j=1}^{J} \sum_{k=1}^{K} d_{jkt} \left[\varepsilon_{jkt} - \ln V_{jkt}(h, h', b_\tau) - \ln \Gamma\left(\frac{b_{t+1}+1}{b_{t+1}}\right)\right] \tag{8}$$

where:

$$V_{jkt}(h, h', b_\tau) \equiv \min \left\{ \begin{array}{c} \alpha_{jkt}(h)^{\frac{1}{b_t}} p_{0,t+1}\left(h + \Delta_{jkt}, h' + \Delta_{jkt}\right)^{\frac{b_t - 1}{b_t b_{t+1}}} E_t\left[\upsilon'_{jk,t+1}\right]^{1 - \frac{1}{b_t}} \\ \beta_{jkt}(h)^{\frac{1}{b_t}} \left\{ p_{0,t+1}\left(h + \underline{\Delta}_{jkt}, h' + \Delta_{jkt}\right)^{\frac{b_t - 1}{b_t b_{t+1}}} E_t\left[\upsilon'_{jk,t+1} g_{jkt}(\pi \mid h)\right]\right\}^{1 - \frac{1}{b_t}} \end{array} \right\}, \tag{9}$$

The first element of the minimization operator in Equation (9) is proportional to the manager's conditional valuation function, net of lifetime utility conferred by endowment wealth, at age t in position (j, k) with human capital h and reputation h' from choosing to work. The second element is proportional to a conditional-valuation function for a similarly placed manager from choosing to shirk:

She reaps the immediate benefit from shirking since $\beta_{jkt}(h) < \alpha_{jkt}(h)$, but firm returns are drawn from $g_{jkt(\tau)}(\pi \mid h)f(\pi)$ rather than $f_j(\pi)$, affecting the probability distribution of her compensation; her reputation subsequently diverges further from her true human capital. Formally the result follows from four features of the model: absolute risk aversion assumption plus markets for nonlabor income implies separation between consumption smoothing from labor income process and financial wealth (Margiotta and Miller, 2000); satisfies inversion theorem so can write differences in conditional valuation function in terms of choice probabilities (Hotz and Miller, 1993 , Proposition 1); satisfies Type 1 Extreme Value assumption which gives simple functional form of representation; retirement choice satisfies terminal state condition (Hotz and Miller, 1993) so other options can be valued easily with respect to this exit option.

The Type 1 extreme Next, we characterize the firm and rank choice probabilities, and how they change over the lifecycle in an equilibrium in which all executives work diligently. Empirically, these choice probabilities map into the model's parameters, and therefore play an important role in estimation. The vector of choice probability functions, $p_t(h, h) \equiv (p_{11t}(h, h), \ldots, p_{JKt}(h, h))$, that the executive uses to compute $B_t(h, h, b_\tau)$ in Equation (**??**) are precisely the probability functions that characterize her choices when solving the optimization function described by (8).

Suppose all managers work diligently in all periods, $h'_{t+1} = h' + \Delta_{jkt}$ and $h = h'$ and ε_{jkt} is independently and identically distributed as a Type I extreme value then it immediately follows from the top line of (9) that:

$$\ln\left(\frac{p_{jkt}(h,h)}{p_{0t}(h,h)}\right) = -\ln\alpha_{jkt}(h) - (b_\tau - 1)\left(\frac{1}{b_{\tau+1}}\ln p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt}) + \ln\{\Gamma\left(\frac{b_{t+1}+1}{b_{t+1}}\right) E_t[\upsilon_{jk,t(\tau)+1}]\}\right). \tag{10}$$

The generalized dynamic Roy model is augmented by an additional component, the future expected utility attached to a job. In our specific case it simplified to a function of next period probability of working in the executive occupation and it depends on human capital accumulated in the job and firm. If the probability of retirement next period increases for some exogenous reason, this would decrease the probability of choosing any job today. The probability of remaining in the executive occupation depends on the human capital accumulated in a job at time t. Thus, if human capital accumulated while being on the board is valuable, it is captured by the change in the retirement probability. Again, human capital is defined broadly and captures the value of networking for example.

The above Theorem captures the relationship between positions, being a board member and the retirement probabilities; It is tempting to relate human capital of board members and the retirement probabilities using Equation 10; however $p_{0,t+1}(\cdot)$, is an equilibrium object and we cannot find the sign of its derivative with respect to h unambiguously outside of the general equilibrium. While the assumption that $p_{0,t+1}(\cdot)$ is decreasing in human capital may seen reasonable, this predicate relies on the assumption that human capital increases with expected compensation. We are not able to assert that at this stage as compensation is determined in equilibrium, and we have not analyzed the demand side of the executive market as of yet.

3.8 Labor Demand, Optimal Contracting and Equilibrium

The shareholders' objective will be to minimize the executive team's expected aggregate compensation bill because they are concerned about only a single manager shirking conditional on the others working diligently. This is equivalent to minimizing $E_t[w_{jk,t(\tau)+1}(h', \pi) \mid h]$ or, equivalently, $E_t[\ln v_{jk,t(\tau)+1} \mid h']$. The shirking contract minimizes $E_t[\ln v_{jk,t(\tau)+1} \mid h']$ subject to a market participation constraint characterized by the executive employment decision rule summarized in Equation (10).

To elicit diligence from the managers, the shareholder must offer a contract that gives a higher utility than the outside option, Equation (10), and a higher expected utility than shirking provides. These equations therefore represent the participation constraint and incentive compatibility constraints respectively.

$$\begin{aligned} \alpha_{jkt}(h)^{1/(b_\tau-1)} E_t[v_{jk,t+1}] B_{t+1}[h+\Delta_{jkt}, h+\Delta_{jkt}] \quad \leq & \\ \beta_{jkt}(h)^{1/(b_\tau-1)} E_t[v_{jk,t+1} g_{jkt}(\pi \quad \mid \quad h)] B_{t+1}\left[h+\underline{\Delta}_{jkt}, h+\Delta_{jkt}\right]. \end{aligned} \tag{11}$$

Thus, whenever $B_{t+1}[h+\Delta_{jkt}, h+\Delta_{jkt}] < B_{t+1}\left[h+\underline{\Delta}_{jkt}, h+\Delta_{jkt}\right]$, career concerns ameliorate the agency problem.

Fama (1980) argues that markets discipline managers because of reputation, thus there is no need for explicit incentives contract. Holmstom (1982) develops a formal model of reputation and career concerns when it is not possible to write contingent contracts based on output and shows that this argument holds under very restrictive conditions. In our model managers who shirked have less human capital than managers who worked diligently. The equilibrium offers which depend on productivity and skills and are designed for managers who always worked diligently may not be as profitable for managers who shirks. This aspect on our model captures the discipline that the market imposes on managers. In equilibrium, contracts are designed such that they indeed discipline managers to always work diligently, thus signaling and updating only occurs off-the-equilibrium path. We therefore incorporate Fama's logic, however, as pointed out by Holmstrom, explicit incentives are required. In our model, explicit incentives given by contracts are the tool shareholders and the market discipline managers to work diligently in every period. Board members, like any other executive are also disciplined by the market and the demand for the services in a similar way (this point is made by Fama (1980) as well).

The compensation schedule minimizes expected wage payments from employment subject to the participation and incentive-compatibility constraints decomposes into fixed and variable components. Define the variable component part by

$$r_{jk,t+1}(h,\pi) \equiv \frac{b_{\tau+1}}{\rho} \ln\left[1 - \eta(h)\{g_{jkt}(\pi \mid h) - \left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)}\right]^{1/(b_{\tau(t)}-1)} \left(\frac{p_{0t+1}[h+\Delta_{jkt}, h+\Delta_{jkt}]}{p_{0t+1}[h+\underline{\Delta}_{jkt}, h+\Delta_{jkt}]}\right)^{\frac{1}{b_{\tau+1}}}\right], \tag{12}$$

where $\eta(h)$ is the unique positive root to

$$\int \left[\eta^{-1} + \left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)} \right]^{1/(b_\tau - 1)} \left(\frac{p_{0t+1}[h+\Delta_{jkt}, h+\Delta_{jkt}]}{p_{0t+1}[h+\underline{\Delta}_{jkt}, h+\Delta_{jkt}]} \right)^{\frac{1}{b_{\tau+1}}} - g_{jkt}(\pi \mid h) \right]^{-1} f_j(\pi)\, d\pi = 1. \tag{13}$$

For a manager who worked diligently up to period $t-1$, the difference between the risk premiums in the basic and the extended models is the value of human capital attained by diligent work relative to the value of human capital attained if the manager shirks.

The optimal contract is the sum of the compensating-equivalent wage and the variable component defined in the optimal contract. The certainty equivalent wage, when $h = h'$, is given by:

$$w^*_{jk,t+1}(h) = \frac{b_{\tau+1}}{\rho} \left\{ \frac{\ln \alpha_{jkt}(h)}{b_\tau - 1} + \frac{\ln\left(p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt})\Gamma\left[1+\frac{1}{b_{\tau+1}}\right]\right)}{b_{\tau+1}} + \frac{1}{b_\tau - 1} \ln \left(\frac{p_{jkt}(h,h)}{p_{0t}(h,h)} \right) \right\}. \tag{14}$$

There are three sources of pay differentials required to attract a manager of characteristics h with probability $p_t(h,h)$. Differentials in the certainty-equivalent wage arise because jobs differ in the value and nonpecuniary costs of working, $\alpha_{jkt}(h)$, and in the value of human capital provided by the job, $p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt})$. In addition, jobs are different in the agency-risk premium, which is determined by the likelihood ratio, and by the relative disutility of working versus shirking in a particular job. The agency risk premium is denoted by $r_{jk,t+1}(h,\pi)$.

It can be shown that 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

$$w_{jk,t+1}(h,\pi) = w^*_{jk,t+1}(h) + r_{jk,t+1}(h,\pi). \tag{15}$$

In the model shirking executives affect the firm's future returns, both directly through F_{jkt}, and also, since $h \neq h'$ for shirking 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 potentially cheaper to implement because more than one signal is used to achieve incentive compatibility in any given period.

The model is a signaling game because executives have private information over their human capital, and their choices and accepted offers can potentially provide information on their human capital. However, given the support of the realization of output and the support of the taste shock, all outcomes and job–rank choices are consistent with the beliefs that no manager has shirked. Thus, job–rank choices and output realizations do not serve as a signal. However, the contracts executives offer may serve to signal their level of human capital.

Theorem 5.3 in Gayle, Golan and Miller (2015), establishes existence of a sequential-equilibrium with one-period contracts, in which the expected compensation of the executives' equals their marginal productivity. That is,

$$E_t \left[w_{jkt(\tau)+1}(h', \pi) \,|h \right] = F_{jk}(h'). \tag{16}$$

4 Empirical Strategy

Having completed the model section we now turn to the empirical strategy that allow us to separate the effect of the various driving forces behind the empirical regularities observed in our data through the lens of our theoretical model. In this section we first outline the formal identification of the our model and then turn to the estimation of the structural parameters.

4.1 Identification

Our data consist of matched panel data on firms and their executives in different time periods, consisting of job-match choices d_{jkt} over the firms j and ranks k, compensation w_{jkt} indexed by age t, executive demographic information and employment histories h_{it}, excess firm returns $\pi_{j\tau}$ indexed by calendar time τ, and bond prices b_τ, again indexed by calendar time. The model is characterized by its preference and technology parameters. The preference parameters include the coefficient of risk aversion ρ, the disutility from working $\alpha_{jkt}(h_t)$, the disutility from shirking $\beta_{jkt}(h_t)$, and an idiosyncratic taste shock associated with each job match $G(\varepsilon_t)$. The technology parameters are the marginal product of work $F_{jk}(h)$, the probability density function of excess returns when every executive works, $f_j(\pi)$, and the likelihood ratio $g_{jk}(\pi\,|h_t)$ that essentially defines the density $f_j(\pi)g_{jk}(\pi\,|h_t)$ when everybody except from one executive in rank k at firm j works.

There are potentially two situations to investigate, depending on whether or not it is optimal to pay executives a constant wage. The latter arises when career concerns are so pronounced that the incentive-compatibility constraint is not binding, meaning (11) is satisfied, or when the cost-minimizing risk premium is so high relative to the net losses from shirking that executives are optimally paid to shirk. All the executives in our data receive compensation awards that depend on excess firm returns, leading us to focus on the former situation, when it is optimal for executives to work because the incentive-compatibility constraint is met with equality in equilibrium.

We assume the data are generated by an equilibrium in which every executive works. Thus, $F_{jk}(h)$ is identified from the conditional expectation of $w_{jk,t+1}$ on d_{jkt}, h_t, and t using the rent extraction condition equation (16); $f_j(\pi)$ is identified from observations on $\pi_{j\tau}$. Since Magnac and Thesmar (2002) have shown that the distribution of unobserved idiosyncractic shocks is not identified nonparametrically in dynamic discrete choice models, we have assumed throughout the paper that distribution of ε_t is type 1 extreme value. This leaves only ρ, $\alpha_{jkt}(h_t)$, $\beta_{jkt}(h_t)$, and $g_{jk}(\pi|h_t)$.

The identification of this general class of models are established in Gayle and Miller (2015) and Gayle et al. (2015) and the reader is referred to these sources for the general results. In this paper we focus on why structural differences across boards and type of executives' network have difference implications for executives' behaviors and compensation structure and how these differences in behaviors and compensation structure can be used to separate the influence of explicit incentives, career and reputation concerns, and market forces across different board and executive's network structures.

Our approach to identification mimics the one we used to explain the model. First, we analyze identification of ρ and $\alpha_{jkt}(h_t)$, the preference parameters that generate the job-match choices observed in the data, when (i) the equilibrium choice is to work each period and (ii) the compensation schedule inducing the effort choice is given. These are summarized by the two main equations exploited in identification and estimation. The first equation is the equilibrium sorting equation which we use to establish identification in dynamic Roy models where there is human capital accumulation. The restrictions from this equilibrium allows us to separate the effect of market, reputation, and agency on the level of pay across different board and executive's network structures. Intuitively different board and executives's network structures essentially provide different lotteries over which executives are making choices. Therefore when an executives switch from a firm with one type of board and executives's network structures we are able to identify there risk preference (ρ) and nonpecuniary benefits ($\alpha_{jkt}(h_t)$) once the compensation profile — in terms of level and variability — are different across these different board and executives's network structures. Hence, the importance of having match data with the same executive moving across firms. The second equation used in identification is optimal contract in equation (15). There we will show – using the results from Gayle and Miller (2015) – how $\beta_{jkt}(h_t)$ and $g_{jk}(\pi|h_t)$ are identified and how to identify career concerns.

4.1.1 Identification of market, reputation, and nonpecuniary taste.

The conditional-choice probability (CCP) vector, $p_t(h)$, is identified by the conditional expectation of $d_{ijk\tau}$, on $(h_{i\tau}, t_{i\tau}, b_\tau)$. Exponentiating equation (10) and then raising it to the power of $1/b_\tau$ yields[10]

$$\alpha_{jkt}(h)^{\frac{1}{b_\tau}}\left\{E_t[\upsilon_{jk,t+1}]p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt})^{\frac{1}{b_{t=1}}}\Gamma\left[(b_{t+1}+1)/b_{t+1}\right]\right\}^{1-\frac{1}{b_\tau}} = \left(\frac{p_{jkt}(h,h)}{p_{0t}(h,h)}\right)^{\frac{1}{b_\tau}}. \tag{17}$$

Rearranging (17) we obtain

$$\alpha_{jkt}(h_t) = \left(\frac{p_{jkt}(h,h)}{p_{0t}(h,h)}\right)\frac{1}{p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt})^{b_\tau-1}\Gamma\left[\frac{b_{t+1}+1}{b_{t+1}}\right]^{b_\tau-1}}E\left[e^{-\rho w_{jk,t+1}(h,\pi)/b_{\tau+1}}|h_t, j\right]^{1-b_\tau}. \tag{18}$$

Equation (18) is an equilibrium sorting condition characterized by $E_t[\upsilon_{jk,t+1}]$ that accounts for certainty equivalent pay, the value of human capital $p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt})\Gamma\left[\frac{b_{t+1}+1}{b_{t+1}}\right]$, a shrinkage factor that raises the value of job matches, and a market-clearing condition captured by $\left(\frac{p_{jkt}(h,h)}{p_{0t}(h,h)}\right)$ that equilibrates the idiosyncratic individual taste disturbances.

The compensation schedules offered by different ranks, firms, board and executives's network structures can be interpreted as choices over lotteries with different nonpecuniary characteristics. Thus, (18) can be used to identify both $\alpha_{jkt}(h_t)$ and ρ when exclusion restrictions exist that limit the dependence of the taste parameters on variables the help determine the contract. Define $z_{jkt}(h, b_\tau, b_{\tau+1})$ as

$$z_{jkt}(h, b_\tau, b_{\tau+1}) \equiv \Gamma\left(\frac{b_{\tau+1}+1}{b_{\tau+1}}\right)^{-1} p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt})^{\frac{-1}{b_{\tau+1}}}\left[\frac{p_{0t}(h,h)}{p_{jkt}(h,h)}\right]^{\frac{1}{(b_\tau-1)}} \tag{19}$$

since $p_{jkt}(h,h)$, $p_{0t}(h,h)$, and ${}^{-1}p_{0,t+1}(h+\Delta_{jkt}, h+\Delta_{jkt})$ are identified from the conditional expec-

[10] Henceforth, the dependence of $A_t(h)$ and $B_t(h,h')$ on b_τ is made explicit. In identification and estimation, b_τ plays a critical role; for example, in Gayle and Miller (2009b) the exclusion restriction on b_τ is one of the main sources of identification.

tation of $d_{ijk\tau}$, on $(h_{i\tau}, t_{i\tau}, b_\tau)$, so is $z_{jkt}(h, b_\tau, b_{\tau+1})$. Identification of ρ and $\alpha_{jkt}(h)$ then follow from assumptions that some components of (j, k, t, h, b_τ) affect $z_{jkt}(h, b_\tau, b_{\tau+1})$ but neither ρ nor $\alpha_{jkt}(h)$. Note that all the elements in (j, k, t, h, b_τ) belong to the information set of the executive at the beginning of each age period t that affects her choices. This can be ascertained by checking for variation in the CCP vector. Hence, they qualify as valid instruments if they do not affect preferences as well. In this paper, we assume that (i) ρ is independent of an executive's human capital and (ii) that the nonpecuniary cost of switching firms or ranks does not depend on some dimension of human-capital accumulation. In estimation, we use previous ranks as an instrument. Similarly, b_τ is a valid instrument if, as we later assume, ρ and $\alpha_{jkt}(h)$ are independent of the aggregate state of the economy.

Let x denote a vector of instruments constructed from (h, j, k, b_τ) for each observation, and define the unconditional density of π as $f(\pi)$. Substituting $z_{jkt}(h, b_\tau, b_{\tau+1})$ into (18), rearranging to make $z_{jkt}(h, b_\tau, b_{\tau+1})$ the subject of the equation, and taking expectations conditional on x yields

$$E[z_{jkt}(h, b_\tau, b_{\tau+1})|x] = E\left[\alpha_{jkt}(h)^{\frac{1}{b_\tau - 1}} \exp\left(\tfrac{-\rho w_{jk,t+1}(\pi,h)}{b_{\tau+1}}\right) \tfrac{f_j(\pi)}{f(\pi)} |x\right]. \tag{20}$$

Thus, ρ and $\alpha_{jkt}(h)$ are identified from the conditional expectations function (20).

4.2 Identification of explicit and implicit incentives

From the data the equilibrium compensation schedule, $w_{jk,t+1}(h_t, \pi)$, is identified by the conditional expectation of individual observations of compensation on $(d_{jkt}, \pi_{j\tau}, h_t, t, b_\tau)$.[11] The finite-upper-bound property of $r_{jk,t+1}(h, \pi)$ and the optimal compensation schedule in equation (15) imply that compensation is bounded and the executive's maximum compensation is

$$\lim_{\pi \to \infty} w_{jk,t+1}(h, \pi) = w^*_{jk,t+1}(h) + \overline{r}_{jk,t+1}(h) \equiv \overline{w}_{jk,t+1}(h). \tag{21}$$

Thus, $\overline{w}_{jk,t+1}(h_t)$ is identified by the maximum of $w_{jk,t+1}$ conditional on $(d_{jkt}, h_t, t, b_\tau)$.

Gayle et al. (2016) demonstrates that, in equilibrium, $g_{jk}(\pi\,|h_t)$ is a mapping of the identified functions $p_t\,(h)$, $w_{jk,t+1}(h_t, \pi)$, $\overline{w}_{jk,t+1}(h_t)$, and ρ. Intuitively, (22) shows $g_{jk}(\pi\,|h_t)$ is identified from the curvature of $w_{jkt+1}(h_t, \pi)$. Therefore in equilibrium

$$g_{jk}(\pi|h_t) = \frac{e^{\rho\overline{w}_{jk,t+1}(h_t)/b_{\tau+1}} - e^{\rho w_{jk,t+1}(h_t,\pi)/b_{\tau+1}}}{e^{\rho\overline{w}_{jk,t+1}(h_t)/b_{\tau+1}} - E[e^{\rho w_{jk,t+1}(h,\pi)/b_{\tau+1}}|h_t, j]}. \tag{22}$$

So identification purposes is instructive to a virtual shirking parameter as

$$\beta^*_{jkt}(h) \equiv \beta_{jkt}(h) \left\{ \frac{p_{0,t+1}\left[h+\Delta_{jkt}, h+\Delta_{jkt}, b_\tau\right]}{p_{0,t+1}\left[h+\underline{\Delta}_{jkt}, h+\Delta_{jkt}, b_\tau\right]} \right\}^{(b_\tau - 1)}. \tag{23}$$

Having identified the working preference parameter $\alpha_{jkt}(h_t)$ from (18) and the likelihood ratio $g_{jk}(\pi|h_t)$ from (22), the shirking preference parameter $\beta^*_{jkt}(h_t)$ is now identified from the incentive-compatibility

[11] In this way, we allow for observations on compensation to be measured with independent error.

constraint (**??**), which holds with equality when compensation varies with π:

$$\beta_{jkt}^{*}(h)=\left(\frac{p_{jkt}(h,h)}{p_{0t}(h,h)}\right)\frac{1}{p_{0,t+1}(h+\Delta_{jkt},h+\Delta_{jkt})^{b_{\tau}-1}\Gamma\left[\frac{b_{t+1}+1}{b_{t+1}}\right]^{b_{\tau}-1}}E\left[e^{\rho w_{jk,t\ +1}(h,\pi)/b_{\tau+1}}g_{jk}(\pi|h_{t})|h,j\right]^{1-b_{\tau}}. \quad (24)$$

By exactly the same logic, that ρ, $g_{jk}(\pi|h)$, $\alpha_{jkt}(h)$, and $\beta_{jkt}^{*}(h)$ are identified. Note that the virtual shrinking parameter is a combination of the explicit incentives $(\beta_{jkt}(h))$ and the implicitly incentives $\left\{p_{0,t+1}\left[h+\Delta_{jkt},h+\Delta_{jkt},b_{\tau}\right]/p_{0,t+1}\left[h+\underline{\Delta}_{jkt},h+\Delta_{jkt},b_{\tau}\right]\right\}^{(b_{\tau}-1)}$. While $p_{0,t+1}\left[h+\Delta_{jkt},h+\Delta_{jkt},b_{\tau}\right]$ is identified on the equilibrium paths $p_{0,t+1}\left[h+\underline{\Delta}_{jkt},h+\Delta_{jkt},b_{\tau}\right]$ is counterfactual is not. So in order to identify the explicit from the implicit incentives we need to identify $p_{0,t+1}\left[h+\underline{\Delta}_{jkt},h+\Delta_{jkt},b_{\tau}\right]$.. Imposing exclusion restrictions on preferences or the technology of human capital accumulation does, however, distinguish the explicit incentives from the implicit incentives component. To illustrate, consider the following three restrictions:

(i) Suppose $\beta_{jkt}(h)$ does not depend on the executive's age, meaning $\beta_{jkt}(h)=\beta_{jk}(h)$ for all t, and there is a maximum retirement age T. Recalling that at age T there is no investment value from human capital or career concerns, then

$$p_{0,T}\left[h+\Delta_{jkT-1},h+\Delta_{jkT-1},b_{\tau(T)}\right]=p_{0,T}\left[h+\underline{\Delta}_{jkT-1},h+\Delta_{jkT-1},b_{\tau(T)}\right]=1.$$

In this case, the shirking parameter is identified from (23) as:

$$\beta_{jk}(h)=\beta_{jk,T-1}^{*}(h)=\left(\frac{p_{jkT-1}(h,h)}{p_{0T-1}(h,h)}\right)E\left[e^{\rho w_{jkT}(h,\pi)/b_{\tau(T)}}g_{jk}(\pi|h)\,|h,j,k\right]^{1-b_{\tau}}.$$

Intuitively, models with with or without implicit incentives have exactly the same predictions if the executive is of age $T-1$ and has not shirked before, so the distinction between $\beta_{jk}(h)$ and $\beta_{jk}^{*}(h)$ is moot. Having identified $\beta_{jk}(h)$, the continuation value associated with shirking the first time is also identified from (23) for all $t\leq T-2$ as

$$p_{0,t+1}\left[h+\Delta_{jkt},h+\Delta_{jkt},b_{\tau(t+1)}\right]=\left[\beta_{jkt}^{*}(h)/\beta_{jk}(h)\right]^{\frac{1}{b_{\tau}-1}}p_{0,t+1}\left[h+\underline{\Delta}_{jkt},h+\Delta_{jkt},b_{\tau(t+1)}\right].$$

In this way, the importance of career concerns at younger ages can be compared by showing how the identified continuation value of shirking for the first time varies over the lifecycle. Note that the basic model does have empirical content against the extension that nests it: Under the null hypothesis of no career concerns, $p_{0,t+1}\left[h+\Delta_{jkt},h+\Delta_{jkt},b_{\tau(t+1)}\right]=p_{0,t+1}\left[h+\underline{\Delta}_{jkt},h+\Delta_{jkt},b_{\tau(t+1)}\right]$.

(ii) Similarly, suppose $\beta_{jkt}(h)$ is independent of aggregate shocks in the economy, more specifically, bond prices b_{τ}. In this case, given (j,k,t,h) and two bond prices $b_{\tau'}\neq b_{\tau''}$, equation (23) yields two equations in three unknowns – namely, $\beta_{jkt}^{*}(h)$, $p_{0,t+1}\left[h+\Delta_{jkt},h+\Delta_{jkt},b_{\tau'(t+1)}\right]$, and $p_{0,t+1}\left[h+\Delta_{jkt},h+\Delta_{jkt},b_{\tau''(t+1)}\right]$. Relative to the normalization $p_{0,t+1}\left[h+\Delta_{jkt},h+\Delta_{jkt},b_{\tau'(t+1)}\right]=1$, the other two parameters are identified.

(iii) If $\underline{\Delta}_{jkt}$ is known, then $p_{0,t+1}\left[h+\underline{\Delta}_{jkt},h+\Delta_{jkt},b_{\tau(t+1)}\right]$ can be numerically calculated in recursive fashion starting from $t=T$ using equation (9). The parameter $\beta_{jkt}^{*}(h)$ now follows from

(23).

4.3 Estimation

We use a four step procedure, which directly follows the approach of our identification strategy, to estimate and test our models:

1. Flexibly estimate $w_{jkt}(\pi, h)$, $\overline{w}_{jkt}(h)$, $f_j(\pi)$, $f(\pi)$, and $p_{jkt}(h)$.

2. Estimate ρ and $\alpha_{jkt}(h)$ from sample moments formed from population moments implied by (20), replacing $w_{jkt}(\pi, h)$, $\overline{w}_{jkt}(h)$, $f_j(\pi)$, $f(\pi)$, and $p_{jkt}(h)$ with their estimates obtained from Step 1.

3. Use the formulas from equations (22) and (24) to estimate $g_{jk}(\pi\,|h)$ and $\beta^*_{jkt}(h)$ by replacing ρ with its estimate from Step 2 and $w_{jkt}(\pi, h)$, $\overline{w}_{jkt}(h)$, $f_j(\pi)$, and $p_{jkt}(h)$ with their estimates from Step 1.

4. Numerically calculate $p_{0,t+1}\left[h + \underline{\Delta}_{jkt}, h + \Delta_{jkt}, b_{\tau(t+1)}\right]$ recursively, assuming that $\beta_{jkt}(h)$ is independent of b_τ and that $\underline{\Delta}_{jkt}$ is known, and test the implied overidentifying restrictions.

Step 1. The state space for the dynamic system is the Cartesian product of the executive's age, t, and personal background, $h_t \in \{1, \ldots, H\}$, at the beginning of each period, as well as a vector that includes her employer firm during the last period, $j_{t-1} \in \{1, \ldots, 36\}$, management rank last period, $k_{t-1} \in \{0, 1, \ldots, 5\}$, fixed components (such as cohort, gender, and education), and other variable components (such as measures of executive experience). Job matches in our model follow a stochastic law of motion, $p_{jkt}(h_t)$ and $p_{0t}(h_t)$. We estimate a multinomial logit model of firm type and position transitions with some (but not all) interactions for exit, promotions, and turnover. In estimation, we exploit Bayes' rule: Given background h, the (joint) probability, $p_{jkt}(h_t)$, is the product of the probability of choosing the jth firm conditional on choosing the kth rank, and the (marginal) probability of choosing Rank k. The compensation schedule, $w_{jkt(\tau)}(\pi, h)$, is estimated using a polynomial, and the boundary condition, $\overline{w}_{jkt(\tau)}(h)$, is estimated using the maximum of $w_{jkt(\tau)}(\pi, h)$ over π. Finally, $f_j(\pi)$ and $f(\pi)$ are estimated using kernel density estimators with normal kernel and the Silverman rule of thumb for the bandwidth.

Step 2. To estimate ρ and $\alpha_{jkt}(h)$, we exploit the exclusion restrictions discussed in the identification section by forming population moments from the conditional expectation function (20). We approximate $z_{jkt}(h)$ by substituting the Step 1 estimates of the conditional-choice probabilities, $p_{0t}(h)$, $p_{jkt}(h)$ and $p_{0,t+1}\left(\overline{H}_{jk}(h)\right)$ into (19). Sample analogs for the CCP vector, the compensation schedule, and conditional and unconditional densities of the abnormal return from Step 1 are substituted into Equation (20). Consistent estimates of ρ and $\alpha_{jkt}(h)$ are then obtained from the approximate sample moments along with (consistent estimates of their) standard errors adjusted for the pre-estimation.

We specify $\alpha_{jkt}(h)$ as a log-linear function of 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, and indicators for board membership, interlocked, no college degree, MBA, MS/MA, PhD, and gender. We estimate an unrestricted version of the model that allows $\alpha_{jkt}(h)$ and ρ to be fully interacted with rank and firm type. This allows us to test whether ρ is a function of firm size, a possibility that might arise if our assumption of absolute risk aversion is violated (Baker and Hall, 2004). We interact these 16 variables with rank and firm type to form $\alpha_{jkt}(h)$. We also permit the risk-aversion parameter to vary by the 36 firm types, but not by rank. In total, there are $(16 \times 5 + 1) \times 36 = 2{,}916$ parameters to be estimated. Equation (20) 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 use bond prices and the lag of Ranks 1 through 4 as instruments, adding another $5 \times 20 \times 36 = 3{,}600$ moment conditions. After rejecting the null hypothesis that ρ varies with firm size, we impose these and other nonrejected restrictions on the results and reestimate the model. These restrictions are a common ρ for all firm types and that the effect of rank and firm type in $\alpha_{jkt}(h)$ is additive. This reduces the number of parameters to $(16 \times 36 + 5 \times 16 + 1) = 657$. We obtain similar results from both the restricted and unrestricted versions; hence, only the restricted version is reported.

Step 3. We form $\widehat{w}(h_t, \pi)$, the nonparametric estimates of the compensation schedule, as a polynomial expansion from Step 1, using them in conjunction with our estimate of the risk-aversion parameter obtained from Step 2. We approximate the conditional expectation, $E_t[\exp(-\widehat{\rho}\widehat{w}\,(h_t, \pi)\,/b_{\tau+1}]$, by integration using the nonparametrically estimated density of π for a given j, from Step 1, and compute $\overline{w}_{jk,t+1}(h)$ using the maximum $\widehat{w}\,(h_t, \pi)$ for each value of (j, k, t, h). Finally, our estimate of $g_{jk}(\pi|h)$ is obtained by substituting our estimates of $\overline{w}_{jk,t+1}(h)$, ρ and $E_t[v_{jk,t+1}(\rho, \pi)]$ into equation (22). The sample analog of the CCP vector, $\widehat{w}(h_t, \pi)$,and the estimates of $g_{jk}(\pi|h)$ are now substituted into a sample average of equation (24) to obtain an estimate for $\beta^*_{jkt}(h)$.

Step 4. Estimates of $\beta_{jkt}(h)$ and $p_{0,t+1}\left[h + \underline{\Delta}_{jkt}, h + \Delta_{jkt}, b_{\tau(t+1)}\right]$ are obtained recursively. Noting that $p_{0,T+1}\left[h + \underline{\Delta}_{jkT}, h + \Delta_{jkT}, b_{\tau(t+1)}\right] \equiv 1$ and substituting our estimated risk-aversion parameter and conditional choice probabilities into equation (**??**) yields $\beta_{jkT}(h)$. Substituting $\beta_{jkT}(h)$ into equation (9) yields $V_{jkT}(h, h', b_\tau)$ and hence $p_{0,T}\,(h, h', b_\tau)$, using equation (27). More generally, given $p_{0,t+1}\left[h + \underline{\Delta}_{jkt}, h + \Delta_{jkt}, b_{\tau(t+1)}\right]$, $\beta_{jkt}(h)$ is obtained from equation (**??**); hence, estimates of $V_{jkt}(h, h', b_\tau)$ and $p_{0,t+1}\left[h + \underline{\Delta}_{jkt}, h + \Delta_{jkt}, b_{\tau(t+1)}\right]$ are produced from equations (9) and (27), respectively.

5 Governance, Interlocked Executives, and Board Members

This section presents our estimates of different components comprising the sources of pay differentials between executives and executives who are also board members or interlocked executives. We also present the components of the pay and how they differ by the composition of the board. These estimates shed light on governance practices and the role of market and agency in explaining the pay of interlocked executives and board members. Furthermore, it relates the fraction of insider on the board, the agency problem and the market pay differential between firms with different board structure. The structural parameters of the utility function are presented in the Appendix.

5.1 Governance and Board Structure

Figure 1 presents the differences in compensation components of firms with large and small fraction of insiders on the board. Panel A presents the expected pay and the risk premium for firms with large and small fraction of insiders on the board; Panel B presents the certainty equivalent pay and its components: compensation for non-pecuniary costs of diligent work, compensating differentials pay to meet demand and compensating differentials for human capital and reputation. The expected compensation is the sum of certainty-equivalent pay and a risk premium. The estimation results demonstrate that while the differences in expected pay between firms with large fraction and small fraction of insiders on the board is small (60K more in firms with small fraction of insiders on the board), firms with a large fraction of insiders on the board pay substantially lower certainty equivalent wage for firms with large insider boards: \$380K versus \$740K in firms with small fraction of insiders on the board.

What drives the differences in the certainty equivalent pay? The certainty equivalent pay can be decomposed into three additive components:

$$w^*_{jkt(\tau)+1}(h) = \Delta^{\alpha}_{jkt}(h) + \Delta^{A}_{jkt}(h) - \Delta^{q}_{jkt}(h), \tag{25}$$

where $\Delta^{\alpha}_{jkt}(h)$ is a compensating differential due to the nonpecuniary utility gain or loss incurred by working in (j,k) relative to the outside option, $\Delta^{A}_{jkt}(h)$ is the investment value of (j,k) from accumulating human capital, and $\Delta^{q}_{jkt}(h)$ is a compensating differential that induces selection on the unobserved idiosyncratic preference shocks:

$$\begin{aligned}
\Delta^{\alpha}_{jkt}(h) &\equiv [\rho(b_\tau - 1)]^{-1}\, b_{\tau+1} \ln \alpha_{jkt}(h) \\
\Delta^{A}_{jkt}(h) &\equiv \rho^{-1} b_{\tau+1} \ln A_{t+1}\left[\overline{H}_{jk}(h), b_{\tau+1}\right] \\
\Delta^{q}_{jkt}(h) &\equiv [\rho(b_\tau - 1)]^{-1}\, b_{\tau+1} q_{jk}[p_t(h,h)]
\end{aligned}$$

Note that $q_{jk}[p_t(h,h)]$ is the value of the disturbance $\varepsilon_{jkt} - \varepsilon_{0t}$ that makes the marginal executive in (j,k) indifferent between that position and her outside option at market-clearing pay. Following the literature, we call $q_{jk}[p_t(h,b_\tau)]$ the demand effect.

The main difference between the types of firms driving the difference in certainty equivalent pay is the smaller compensation for diligent work ($\Delta^{\alpha}_{jkt}(h)$) in firms with large fraction of insiders on the board. (1.27 million versus 1.52 Million). Moreover, firms with large fraction of insiders on the board pay less compensating differentials in order to meet market demand (that is, the marginal manager recruited to work has lower disutility from working in the firm and needs \$110K less in compensation for the disutility from the job). This can be indicative of better governance practices improving working conditions in firms with large fraction of insiders on the board. However, the risk premium is larger in those firms (2.51 Million versus 2.23 Million), bringing the expected pay to similar levels.

We then compare the differences in the components of the risk premium: The risk premium is a compensating differential to risk-averse executives for bearing risk in the form of firm-denominated securities. It measures the costs of agency. In our model, it is measured by the difference between expected compensation and certainty equivalent pay defined in equation (14). From (16) expected compensation is the expected value of the executive's marginal product:

$$\Delta^r_{jkt}(h) \equiv E_t\left[r_{jk,t+1}(h,\pi)\right] = F_{jk}(h) - w^*_{jk,t+1}(h). \tag{26}$$

Further look into the compensation for agency risk premium indicates that the net benefit of shirking in firms with larger fraction of insiders on boards is smaller by \$2.2M on average. The gross loss to shareholders from shirking in firms with large fraction of insiders on the boards is on average \$3M smaller. The latter finding also explain the larger risk premium paid by firms with large fractions of insiders on the board. The fact that the gross loss from shirking is smaller for these firms implies that the likelihood ratio is flatter, that is the signals are less informative of shirking. Therefore, conditional on providing incentives for diligent work, when signals are less informative the pay has to be more strongly tied to firm performance and the risk premium is larger. Both evidence are consistent with the role of board in monitoring showing that large fraction of insider boards improve monitoring in firms.

5.2 Compensation, Markets and Incentives

Interlocked Executives Figure 2 documents the decomposition of the pay for interlocked executives using the estimation results. In our model executives are paid their marginal product, since the expected pay of interlocked executives is lower (\$2.43M versus \$2.86M for non-interlocked executives), their marginal productivity is lower in the firm. However, since the pay is composed of variable pay and fixed salary we compute the risk premium and the certainty equivalent. Our estimation results show that the certainty equivalent pay of interlocked executives is on average \$710K relative to \$570K which is the certainty equivalent pay of non-interlocked executives. The main cause for the lower certainty equivalent is the cost of diligent work; it is lower by \$110K for interlocked relative to non-interlocked executives. We did not find significant differences in the value of human capital. We find that the risk-agency premium is lower by \$280K for interlocked executives.

A further look at the component of the agency-risk premium reveals that compensating differential for the value of shirking relative to working diligently is lower by \$910K for interlocked executive. That is, interlocked executives goals are more closely aligned with those of shareholders than goals of other executives reducing the agency-risk premium. This is perhaps an indication of better monitoring for interlocked executives. On the other-hand, the degree to which career concerns ameliorate the moral hazard problem is lower for interlocked executives by \$200K. The loss to shareholders from not providing incentives to work is smaller for interlock executive (loss of 22% versus 25%). The latter implies that providing incentives for interlocked executives is more costly as the signals are less informative, mitigating the large effect of the more aligned preferences of interlocked executives.

Board members Executive directors have a lower expected pay than other executives. Our estimates show that their certainty equivalent pay is also lower on average. Figure 3 shows that on average the certainty equivalent of executive directors is \$635k versus \$710k for non board members executives. The reason for the lower certainty equivalent paid to executive directors is lower non-pecuniary costs, both the systematic part of the non-pecuniary costs of diligent work which requires lower compensating differentials (by 40k) and lower compensation for the idiosyncratic taste for the position which is implied by lower compensating differentials firms pay to meet demand for board members. Board members, however, value human capital slightly less than other executives requiring an extra 4k of compensating differentials. Tables 5-7 show the components of the certainty equivalent pay vary by ranks. Table 5 shows that only at the higher ranks, executives 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. There is greater net demand (Table 6) 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.

The agency risk premium for board members is lower by \$203k. Looking further into the agency risk premium reveals a greater divergence of the shareholders and board members: the compensating differentials for working diligently versus shirking is \$130k higher for executive directors than for the rest of the executives, career concerns ameliorate the problem of moral hazard for executive directors more than it does for other executive (2.01M versus 1.88M). Like interlocked executives, board members also cause less destruction of firm value if they are not provided with incentives than other executives (17.5% gross loss of value versus 24.6). Again this might be because board members are more closely monitored than other executives. The latter two findings explain the lower agency-risk premium of board member.

Entrenchment While the pay for interlocked executives and executive directors is lower than the pay for non-interlocked or non-board members executives, they are more entrenched in two ways; they have a lower probability of turnover and the probability of exit is 55% smaller. Thus, they are also older on average. Our estimates show that the main reason for the observed lower exit probability is the lower non-pecuniary costs (both the cost of the diligent work and the fact that the idiosyncratic disutility from working is lower for interlocked executives). What explains the lower probability of turnover for these executives? First, the non-pecuniary costs are higher; it requires 26k increase in premium for switching for interlocked executives and 111K for executive directors. Second, the net demand for new interlocked executives is lower than the demand for new hires among other executives. Lastly, note that if a new hire is an interlocked executive or executive director the divergence between the goals of the executive and shareholders grows by 318k for interlocked executives and by 34K for executive directors relative to executives in these positions that are not new hires.

Females Our counterfactuals consider quotas for females on boards hence we explore the gender differences revealed in the estimates. Figure 4 panels A-C present the pay decomposition for females. We find that female executives receive higher certainty equivalent pay than male executives. 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. Looking at the different components of the certainty equivalent pay reveals that there is lower net demand (thus, the compensating differentials for females for unobserved factor is lower than that of male executives) for females. The exception is demand for low ranks in the consumer and service sectors.

However, 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 (see Table 5). 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. This pattern may reflect superior outside options, in other labor markets and retirement, for female executives (see discussion in Gayle, Golan and Miller 2012). Lastly, the value of human capital is lower for female executive, requiring larger compensating differentials. 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.

What explains the lower risk premium for females? Looking at Figure 4 panel C, we find significant that there the differences the gross loss from shirking is slightly smaller for females, which if anything should increase the risk premium. Moreover, we find that generally female executives place lower value on career concerns, which is consistent with higher exit rates. An exception is female executives in rank 2 and female executives who join new firms where they place higher value of career concerns than men. Thus, this also increases the risk premium . However, the reason for the lower risk premium is that the net benefit from shirking is lower overall for females, implying that their goals are more aligned with the goals of shareholders.

6 The Role of Insiders on the Board and Female Representation

In this section we consider alternative board structures. We first consider a requirement of having 50% outsiders on the board. The counterfactual will allow us to assess to role and impact of having boards with large fraction of insiders. The second counterfactual requires quotas of females in boards (for example, Norway's 40% female representation on boards). As we documented females behavior and compensation differs, most notably, they are more likely to exit the executive occupation than males. However, their representation on boards is low, and policies mandating females quotas may change exit behavior as well as career paths. Similarly, requiring 50% of outsiders on the board, may change the composition of boards, compensation and career choices. *The work on this section is still oncoming.*

7 Conclusion

This paper estimates a model of executive compensation assessing the role of insiders on the board in governance and analyzing the compensation of board members and interlocked executives. We then perform counterfactual policy analysis first imposing a rule that at least 50% of the board member have to be outsiders, and second imposing quotas for females on boards.

We first document that controlling for ranks and other executive and firm characteristics, interlocked and board members executives are not paid more than other executives. 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. 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, reduces the net benefits from shirking, 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; giving more votes to executives fosters better executive working conditions, which in turn is offset by a lower certainty-equivalent wage in equilibrium. 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.

Despite the fact that their pay is not larger than that of other executives, board members and interlocked executives are more entrenched than other executives. We also document that their compensation structure is different. Our models allows us to uncover the reasons for these differences. Our estimation results reveal that the certainty equivalent pay is substantially smaller than that of other executives. The main reason for the lower pay is the lower non-pecuniary costs of working diligently in these positions. The lower non-pecuniary costs of working also rationalizes why these executives are less likely to exit the executive profession and therefore are more entrenched. While this findings provide further support to the argument that these executives are not extracting higher pay by exploiting existing rules, we cannot rule out the case that there are other unobserved payments not included in compensation packages. The risk premium for executive directors is lower despite the fact that there is a greater divergence between their goals and the shareholders goals. However, the greater career concerns ameliorate the agency problem. Thus our findings support the view that implicit incentives play important role in aligning shareholders and board members goals. Interlocked executives and shareholders, however, have more closely aligned goals, but their career concerns are smaller. Thus, explicit incentive provided by formal compensation contracts are more important than the implicit incentives relative to non-interlocked executives.

To further analyze the role of imposing quotas of females on board, we first find that behavior and compensation of female executives differ from that of male executives. Our empirical results show that, after controlling for other observed characteristics including rank, women are paid the same expected compensation as their male counterparts. We find 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 implying their goals are more closely aligned to the goals of the shareholders. 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 (2012). The higher estimates of certainty equivalent is consistent with females having higher outside options relative to the value of working diligently. 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. Nevertheless, imposing quotas for females on board may change exit behavior and choices of females.

8 Appendix: The Extreme-Value Distribution

In our structural estimation, we assume throughout that ε_t is distributed as a type 1 extreme value. The computational advantages of parameterizing $G(\varepsilon)$ this way are most evident from Lemma .1 below, where we provide formulas for $B_t(h, h')$, the value of human capital on and off the equilibrium path, and also an expression for marginal disturbances, $q_{jk}[p_t(h)]$.

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

$$B_t\left(h, h'\right) = p_{0t}\left(h, h'\right)^{\frac{1}{b_t}} \Gamma\left[\left(b_t + 1\right) / b_t\right]. \tag{27}$$

where $p_{0t}(h, h')$ is the probability that the optimal choice is retirement.

The IIA property of type 1 extreme values implies that the marginal idiosyncratic shock for a manager who is indifferent between the 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 retiring. This ratio does not depend on the other components of 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.

Proof. Denoting the probability density function of $\varepsilon^*_{jkt} \equiv d_{jk}\varepsilon_{jkt}$ by $d\overline{G}(\varepsilon^*_{jkt})$, we first derive an expression for $E[\exp(-\varepsilon^*_{jkt}/b_t)]$ and then use it in our derivation of the formula for $A_t\left(h_t\right)$:

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

$$W_{jkt} \equiv \ln \alpha_{jkt} + (b_t - 1) \ln A_{t+1}\left[\overline{H}_{jk}(h)\right] + (b_t - 1) \log\left\{E_t\left[\upsilon_{jk,t+1}\right]\right\}. \tag{28}$$

Then (j, k) is chosen at t if $\varepsilon_{jkt} + W_{jkt}$ is maximal for all (j', k'). Let $G\left(\varepsilon_{11t}, \ldots, \varepsilon_{JKt}\right)$ denote the probability distribution function for $\left(\varepsilon_{11t}, \ldots, \varepsilon_{JKt}\right)$ and $G_{jk}\left(\varepsilon_{11t}, \ldots, \varepsilon_{JKt}\right)$ its derivative with respect to ε_{jkt}. Since $G\left(\varepsilon_{11t}, \ldots, \varepsilon_{JKt}\right)$ is the product of independently distributed standard Type 1 extreme-value probability distributions in our model,

$$G_{jk}\left(\varepsilon_{11t}, \ldots, \varepsilon_{JKt}\right) = \exp\left(-\varepsilon_{jkt}\right) \prod\nolimits_{(j',k')} \exp\left[-\exp\left(-\varepsilon_{j'k't}\right)\right]. \tag{29}$$

Using the well-known fact that

$$W_{jkt} - W_{j'k't} = \log p_{jkt} - \log p_{j'k't}, \tag{30}$$

it now follows from (29) and (30) that

$$G_{jk}(\varepsilon_{jkt} + W_{jkt} - W_{11t}, \ldots, \varepsilon_{jkt} + W_{jkt} + W_{JKt}) = \exp[-\varepsilon_{jkt} - \exp(-\varepsilon_{jkt} - \log p_{jkt})]. \tag{31}$$

From Equation (28) and Theorem 4.2 in the main text, the conditional-choice probability for (j,k) can be expressed as

$$p_{jkt} = \int_{-\infty}^{\infty} G_{jk}\left(\varepsilon_{jkt} + W_{jkt} - W_{11t}, \ldots, \varepsilon_{jkt} + W_{jkt} + W_{JKt}\right) \mathrm{d}\varepsilon_{jkt}. \tag{32}$$

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

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

To derive $E[\exp(-\varepsilon^*_{jkt}/b_t)]$, we draw from Equations (15) and (17) of Chapter 21 of Johnston and Kotz (1970, 277–278), 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

$$E_t\left[\exp\left(\varepsilon^*_{jkt}/b_t\right)\right] = \exp\left(\log p_{jkt}(h,h')^{1/b_t}\right)\Gamma\left[\left(b_t+1\right)/b_t\right] = p_{jkt}(h,h')^{1/b_t}\Gamma\left[\left(b_t+1\right)/b_t\right]. \tag{33}$$

2. To prove (27), 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 (**??**) and (5.6) in the main text,

$$V'_{jkt}(h,h') = \left[\frac{p_{0t}(h,h')}{p_{jkt}(h,h')}\right]^{1/b_t}. \tag{34}$$

Summing over (j,k) and rearranging, we obtain

$$p_{0t}(h,h') = \left\{1 + \sum_{j=1}^{J}\sum_{k=1}^{K}\left[V'_{jkt}(h,h')\right]^{-1}\right\}^{-1}. \tag{35}$$

Substituting (33) along with the conditional-choice probability ratios (34) and the retirement probability (35) into (5.4) yields

$$B_t(h,h') = p_{0t}(h,h')^{1+\frac{1}{b_t}}\,\Gamma\left(\frac{b_t+1}{b_t}\right) + \sum_{j=1}^{J}\sum_{k=1}^{K}\left[p_{jkt}(h,h')^{1+\frac{1}{b_t}}\,\Gamma\left(\frac{b_t+1}{b_t}\right)\left[\frac{p_{0t}(h,h')}{p_{jkt}(h,h')}\right]^{1/b_t}\right]$$

$$= p_{0t}(h,h')^{\frac{1}{b_t}}\,\Gamma\left[\frac{b_t+1}{b_t}\right],$$

which simplifies to (27).

■

References

[1] Acharya, V. V., & Volpin, P. F. (2010). Corporate governance externalities. *Review of Finance*, 14, 1-33.

[2] Adams, R. B., Hermalin, B. E., & Weisbach, M. S. (2010). The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey. *Journal of Economic Literature*, 48(1), 58-107.

[3] Albanesi, Stefania, and Claudia Olivetti. 2008. Gender and dynamic agency: Theory and evidence on the compensation of top executives. Unpublished manuscript, Columbia University.

[4] Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. *The American Economic Review,* 62(5), 777–795.

[5] Antle, R., & Smith, A. (1985). Measuring executive compensation: Methods and an application. *Journal of Accounting Research,* 23(1), 296–325.

[6] Antle, R., & Smith, A. (1986). An empirical investigation of the relative performance evaluation of corporate executives. *Journal of Accounting Research,* 1–39.

[7] Baker, G. P., & Hall, B. J. (2004). CEO incentives and firm size. *Journal of Labor Economics,* 22(4), 767–798.

[8] Bebchuk, L. A., & Fried, J. M. (2003). Executive compensation as an agency problem (No. w9813). National Bureau of Economic Research.

[9] Bebchuk, L. A., & Fried, J. M. (2005). Pay without performance: Overview of the issues. Journal of Applied Corporate Finance, 17(4), 8-23.

[10] Bell, Linda A. 2005. Women-led firms and the gender gap in top executive jobs. IZA Discussion Paper 1689, July.

[11] Bertrand, M., & Hallock, K. F. (2001). The gender gap in top corporate jobs. *Industrial & Labor Relations Review*, 55(1), 3-21.

[12] Calvo, G. A., & Wellisz, S. (1980). Technology, entrepreneurs, and firm size. *The Quarterly Journal of Economics,* 95(4), 663–677.

[13] Dewatripont, M., Jewitt, I., & Tirole, J. (1999). The economics of career concerns, Part II: Application to missions and accountability of government agencies. *The Review of Economic Studies,* 66(1), 199–217.

[14] Dicks, D. L. (2012). Executive Compensation and the Role for Corporate Governance Regulation. *Review of Financial Studie*s, 25(6), 1971-2004.

[15] Dominguez-Martinez, S., Swank, O. H., & Visser, B. (2008). In defense of boards. *Journal of Economics & Management Strategy*, 17(3), 667-682.

[16] Fama, E. F. (1980). Agency Problems and the Theory of the Firm. *The Journal of Political Economy*, 88(2), 288-307.

[17] Fudenberg, D., Holmström, B., & Milgrom, P. (1990). Short-term contracts and long-term agency relationships. *Journal of Economic Theory,* 50, 1–31.

[18] Gayle, G. L., Golan, L., & Miller, R. A. (2012). Gender differences in executive compensation and job mobility. *Journal of Labor Economics,* 30(4), 829–872.

[19] Gayle, G. L., Golan, L., & Miller, R. A. (2015). Promotion, Turnover and Compensation in the Executive Labor Market. *Econometrica, 83(6), 2293-2369.*

[20] Gayle, G.-L., & Miller, R. A. (2009a). Insider information and performance pay. *CESifo Economic Studies,* 55(3–4), 515–541.

[21] Gayle, G.-L., & Miller, R. A. (2009b). Has moral hazard become a more important factor in managerial compensation? *American Economic Review,* 99(5). 1740–1769.

[22] Gayle, G.-L., & Miller, R. A. (2015). Identifying and testing models of managerial compensation. *The Review of Economic Studies,* 82, 1074-1118.

[23] Gilson, S. C. (1990). Bankruptcy, boards, banks, and blockholders: Evidence on changes in corporate ownership and control when firms default. *Journal of Financial Economics*, 27(2), 355-387.

[24] Grossman, S., & Hart, O. (1983). An analysis of the principal–agent problem. *Econometrica,* 51, 7–46.

[25] Hallock, K. F. (1997). Reciprocally Interlocking Boards of Directors and Executive Compensation *The Journal of Financial and Quantitative Analysis*, 32(3), 331-344.

[26] Hermalin, B. E., & Weisbach, M. S. (1998). Endogenously chosen boards of directors and their monitoring of the CEO. *American Economic Review*, 96-118.

[27] Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies,* 66(1), 169–182.

[28] Hotz, V. J., & Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies,* 60(3), 497–529.

[29] Johnston, N. L., & Kotz, S. (1970). *Distributions in statistics.* New York: Wiley.

[30] Kaplan, S. N., & Reishus, D. (1990). Outside directorships and corporate performance. *Journal of Financial Economics*, 27(2), 389-410.

[31] Kuhnen, C. M., & Zwiebel, J. (2009). Executive pay, hidden compensation, and managerial entrenchment. Northwestern University working paper.

[32] Lee, S., & Persson, P. (2011). Authority versus Loyalty: Social Incentives and Governance. Columbia University.

[33] Magnac, T., & Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica,* 70(2), 801–816.

[34] Malcomson, J. M., & Spinnewyn, F. (1988). The multiperiod principal–agent problem. *Review of Economic Studies,* 55, 391–407.

[35] Margiotta, M., & Miller, R. A. (2000). Managerial compensation and the cost of moral hazard. *International Economic Review,* 41, 669–719.

[36] Mirrlees, J. A. (1975). The theory of moral hazard and unobservable behavior, Part 1. Working Paper. Oxford University.

[37] Mirrlees, J. A. (1976). The optimal structure of incentives and authority within an organization. *The Bell Journal of Economics,* 7(1), 105–131.

[38] Rey, P., & Salanie, B. (1990). Long-term, short-term and renegotiation: On the value of commitment in contracting. *Econometrica,* 58(3), 597–619.

[39] Rose, N. L., & Shepard, A. (1997). Firm Diversification and CEO Compensation: Managerial Ability or Executive Entrenchment?. *RAND Journal of Economics*, 28(3), 489-514.

[40] Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers,* New Series, 3(2), 135–146.

[41] Selody, Karen. 2010. Is the risk worth the reward for female executives? Unpublished manuscript, Department of Economics, University of California, Berkeley.

[42] Williamson, O. E. (1967). Hierarchical control and optimum firm size. *The Journal of Political Economy,* 75, 123–138.

[43] Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics*, 40(2), 185-211.

Table 1 A:. Pay and Compensation Structure Comparison by Rank

VARIABLES	RANK	ALL EXECUTIVES	EXECUTIVE DIRECTORS	INTERLOCKED EXECUIVES	BOARD WITH A LARGE NUMBER OF INSIDERS
Salary	1	594	608	504	616
		(351)	(358)	(362)	(369)
	2	678	695	623	699
		(412)	(403)	(415)	(465)
	3	519	565	511	542
		(314)	(332)	(331)	(358)
	4	368	450	396	388
		(179)	(241)	(256)	(201)
	5	285	365	286	291
		(150)	(221)	(197)	(173)
Bonus	1	705	674	313	777
		(1500)	(1524)	(619)	(1676)
	2	725	748	565	840
		(1782)	(1826)	(1519)	(2286)
	3	608	688	434	698
		(1695)	(1955)	(666)	(1710)
	4	292	409	235	343
		(866)	(1030)	(440)	(836)
	5	178	304	207	192
		(426)	(743)	(408)	(550)
Number of shares owned	1	2021	2592	2877	2187
		(8819)	(10341)	(7903)	(8557)
	2	1812	1923	2484	2095
		(11071)	(11514)	(5880)	(10198)
	3	527	637	1801	651
		(2197)	(2508)	(6024)	(2514)
	4	288	418	548	337
		(1713)	(1721)	(1029)	(1448)
	5	174	338	153	204
		(1012)	(1098)	(506)	(1144)
Value of restricted shares granted	1	359	305	103	369
		(1439)	(1284)	(541)	(1358)
	2	456	480	252	468
		(2155)	(2229)	(1657)	(2174)
	3	413	450	132	465
		(4708)	(5913)	(966)	(5857)
	4	157	211	135	178
		(896)	(1901)	(595)	(1131)
	5	83	106	60	88
		(441)	(587)	(246)	(512)
Number of observations	1	6581	4535	516	4740
	2	28526	26188	1933	14186
	3	8858	5344	222	5522
	4	61131	7961	442	29742
	5	37594	1899	114	15934

Note: Standard deviations are enclosed in parenthesis.

Table 1 B:. Pay and Compensation Structure Comparison by Rank

VARIABLES	RANK	ALL EXECUTIVES	EXECUTIVE DIRECTORS	INTERLOCKED EXECUIVES	BOARD WITH A LARGE NUMBER OF INSIDERS
Values of options held	1	6632	7443	3636	6330
		(19552)	(21922)	(15324)	(17187)
	2	10137	10708	5168	9575
		(28211)	(28780)	(19640)	(26693)
	3	6312	6691	2629	6190
		(21514)	(22501)	(6356)	(20468)
	4	2487	3227	1437	2352
		(9376)	(11364)	(4301)	(6610)
	5	1617	2522	929	1393
		(8596)	(8334)	(2448)	(6569)
Change in wealth from options held	1	71	-177	-199	31
		(14612)	(16383)	(17946)	(14567)
	2	548	654.1	886	521
		(19344)	(19233)	(15766)	(17284)
	3	757	925	457	821
		(17156)	(17308)	(3592)	(15987)
	4	324	387	387	326
		(6927)	(8905)	(2493)	(5071)
	5	262	385	185	212.7
		(6661)	(6246)	(1998)	(6294)
Change in wealth from restricted shares held	1	7327	11214	3721	8765
		(376316)	(454053)	(84732)	(434834)
	2	9888	10584	2220	4523
		(940336)	(973186)	(677919)	(304058)
	3	830	1759	-4611	1604
		(75660)	(94848)	(44960)	(94115)
	4	1469	4156	794	1889
		(123066)	(243655)	(18033)	(139730)
	5	811	404	566	331.6
		(76359)	(17996)	(4200)	(52716)
Total compensation	1	2693	2632	2106	2596
		(25325)	(28312)	(31256)	(26622)
	2	4294	4586	3517	4191
		(25520)	(26159)	(29869)	(27043)
	3	3247	3744	618	3296
		(17708)	(19350)	(23587)	(18785)
	4	1662	2392	1962	1660
		(10979)	(13203)	(12218)	(11511)
	5	1153	2155	1469	1027
		(9091)	(12153)	(4793)	(9157)
Number of observations	1	6581	4535	516	4740
	2	28526	26188	1933	14186
	3	8858	5344	222	5522
	4	61131	7961	442	29742
	5	37594	1899	114	15934

Note: Standard deviations are enclosed in parenthesis.

Table 2A. Education and Personal Attributes Comparison by Rank

VARIABLE	RANK	ALL EXECUTIVES	EXECUTIVE DIRECTORS	INTERLOCKED EXECUTIVES	BOARD WITH A LARGE NUMBER OF INSIDERS
Age	1	59.2	60.82	64.05	59.73
		(9.86)	(9.85)	(10.40)	(10.24)
	2	55.28	55.47	57.27	55.96
		(7.85)	(7.72)	(9.23)	(8.41)
	3	52.11	52.59	54.19	52.48
		(8.05)	(7.38)	(9.77)	(7.88)
	4	51.94	52.25	53.34	51.95
		(9.47)	(7.57)	(8.55)	(9.06)
	5	51.88	52.46	57.2	51.58
		(10.62)	(8.01)	(12.33)	(9.84)
		(0.40)	(0.38)	(0.36)	(0.40)
Years of tenure in the firm	1	16.44	17.74	18.2	17.22
		(12.63)	(12.84)	(13.49)	(12.95)
	2	14.23	14.42	16.49	15.39
		(10.90)	(10.87)	(11.19)	(11.07)
	3	13.23	13.79	12.24	13.79
		(10.47)	(10.48)	(9.23)	(10.47)
	4	13.19	14.99	16.35	13.7
		(10.40)	(10.55)	(12.43)	(10.66)
	5	13.23	13.98	14.72	13.03
		(10.32)	(9.51)	(10.54)	(9.81)
Years of executive experience	1	21.32	22.9	24.51	21.9
		(12.24)	(12.51)	(12.24)	(12.73)
	2	18.86	19.03	21.78	19.65
		(9.91)	(9.86)	(10.20)	(10.07)
	3	15.69	16.45	19	16.2
		(9.91)	(9.44)	(11.53)	(9.81)
	4	15.56	15.79	17.48	15.76
		(10.65)	(9.66)	(11.16)	(10.53)
	5	15.95	16.26	19.59	15.62
		(11.11)	(9.88)	(12.41)	(10.43)
Number of firms worked for before becoming an executive	1	0.683	0.678	0.824	0.694
		(1.17)	(1.16)	(1.69)	(1.20)
	2	0.686	0.679	0.673	0.697
		(1.12)	(1.11)	(1.14)	(1.17)
	3	0.686	0.675	0.684	0.672
		(1.18)	(1.13)	(0.98)	(1.12)
	4	0.89	0.79	0.7	0.859
		(1.32)	(1.17)	(1.13)	(1.29)
	5	1.077	0.951	1.136	0.992
		(1.42)	(1.29)	(1.18)	(1.34)
Number of firms worked for after becoming an executive	1	0.899	0.896	0.971	0.886
		(1.38)	(1.38)	(1.51)	(1.39)
	2	0.912	0.912	0.917	0.865
		(1.38)	(1.38)	(1.43)	(1.32)
	3	0.734	0.745	0.954	0.721
		(1.29)	(1.31)	(1.57)	(1.31)
	4	0.761	0.608	0.568	0.739
		(1.31)	(1.15)	(1.16)	(1.32)
	5	0.797	0.716	1.288	0.766
		(1.34)	(1.25)	(2.13)	(1.30)
Number of observations	1	4,812	3,430	375	3,489
	2	21,283	19,725	1,498	10,561
	3	5,953	3,822	152	3,709
	4	32,550	5,028	273	16,275
	5	18,508	1,105	59	7,844

Note: Standard deviations are enclosed in parenthesis.

Table 2 B. Education and Personal Attributes Comparison by Rank

VARIABLE	RANK	ALL EXECUTIVES	EXECUTIVE DIRECTORS	INTERLOCKED EXECUTIVES	BOARD WITH A LARGE NUMBER OF INSIDERS
Retirement from Executive occupation	1	0.246	0.249	0.171	0.218
	2	0.096	0.094	0.061	0.097
	3	0.137	0.105	0.037	0.125
	4	0.168	0.126	0.102	0.162
	5	0.168	0.135	0.139	0.164
Firm-to-firm transition	1	0.024	0.018	0.006	0.018
	2	0.031	0.030	0.009	0.022
	3	0.027	0.020	0.014	0.020
	4	0.017	0.008	0.007	0.014
	5	0.012	0.004	0.001	0.011
College graduate	1	0.755	0.745	0.752	0.745
	2	0.786	0.787	0.778	0.772
	3	0.752	0.751	0.697	0.75
	4	0.789	0.796	0.832	0.786
	5	0.823	0.753	0.966	0.813
Masters of business administration	1	0.238	0.223	0.224	0.238
	2	0.254	0.256	0.241	0.225
	3	0.230	0.225	0.204	0.223
	4	0.226	0.222	0.198	0.225
	5	0.191	0.226	0.237	0.189
Masters of science	1	0.158	0.154	0.141	0.149
	2	0.172	0.17	0.168	0.163
	3	0.168	0.179	0.211	0.159
	4	0.202	0.199	0.227	0.202
	5	0.205	0.175	0.153	0.197
Phd	1	0.148	0.157	0.221	0.148
	2	0.149	0.149	0.156	0.145
	3	0.132	0.142	0.204	0.139
	4	0.170	0.169	0.139	0.172
	5	0.248	0.156	0.271	0.243
Professional certification	1	0.152	0.143	0.123	0.155
	2	0.141	0.138	0.095	0.142
	3	0.152	0.155	0.178	0.157
	4	0.234	0.226	0.165	0.239
	5	0.333	0.227	0.407	0.324
Female	1	0.018	0.009	0.016	0.018
	2	0.015	0.014	0.011	0.011
	3	0.027	0.022	0.050	0.026
	4	0.058	0.029	0.025	0.046
	5	0.068	0.042	0.123	0.062
Number of observations	1	4812	3430	375	3489
	2	21283	19725	1498	10561
	3	5953	3822	152	3709
	4	32550	5028	273	16275
	5	18508	1105	59	7844

Note: Standard deviations are enclosed in parenthesis.

Table 3. Governance Structure by Firm Type and Rank

VARIABLES	RANK	ALL FIRMS	LARGE FIRMS	MEDIUM SIZE FIRMS	SMALL FIRMS	PRIMARY SECTOR FIRMS	CONSUMER SSECTOR FIRMS
	1	0.689	0.588	0.757	0.885	0.807	0.795
Executive	2	0.918	0.930	0.921	0.906	0.981	0.972
directors	3	0.603	0.597	0.604	0.612	0.693	0.689
	4	0.130	0.122	0.133	0.138	0.136	0.171
	5	0.051	0.051	0.060	0.046	0.050	0.084
	1	0.078	0.047	0.091	0.149	0.084	0.105
Interlocked	2	0.068	0.057	0.071	0.074	0.058	0.073
executives	3	0.025	0.016	0.032	0.033	0.022	0.030
	4	0.007	0.004	0.010	0.009	0.005	0.009
	5	0.003	0.002	0.005	0.003	0.003	0.003
Firms with	1	0.720	0.696	0.739	0.774	0.637	0.773
large	2	0.497	0.495	0.502	0.502	0.432	0.557
numbers	3	0.623	0.604	0.631	0.653	0.582	0.680
of insiders	4	0.487	0.479	0.488	0.501	0.419	0.546
on board	5	0.424	0.395	0.412	0.447	0.389	0.476
	1	6581	3671	1554	1338	1516	1646
Number	2	28526	9685	7391	11333	7996	6113
of	3	8858	3899	2276	2648	2102	2298
observations	4	61131	24660	16885	19323	15820	12905
	5	37594	10202	8759	18574	11338	6347

Table 4: Compensation and transitions

	Compensation			Within firm transition					Firm-to-firm transition	Retirement
Variable	Fixed pay	excess return	excess return sq.	R1	R2	R3	R4	R5		
						Ranks				
constant	-4359	21601	-9114							
	(2716)	(3859)	(1,914)							
R4	103	1529	-242	-20.3	-76	-67.80	63.40	-536		
	(463)	(926)	(444)	(21)	(9.80)	(10)	(4)	(7)		
R3	1267	2627	-164	-88.1	-72	114	-404	-754	94.70	
	(662)	(1407)	(605)	(24)	(12.80)	(4)	(15)	(20)	(18)	
R2	3456	6007	-789	-118	68	-393	-551	-901	213	-8.90
	(683)	(1,394)	(699)	(25)	(3.90)	(20)	(18)	(22)	(13)	(4)
R1	1055	9839	-454	111	123	-345	-585	-939	86.1	55.26
	(797)	(1690)	(987)	(10)	(4)	(34)	(34)	(39)	(23)	(3)
				Governance variables						
Executive	845	7,695	-848	-22.8	123	15.6	-70.4	-105	-102	-64.72
directors	(251)	(570)	(304)	(13.2)	(4.03)	(5.32)	(3.39)	(6.50)	(9)	(4)
Board with a large	280	2,683	-1203						-25.7	
number of insiders	(163)	(570)	(176)						(6)	
Interlocked	-299	6403	-1496						-93	-55
executives	(464)	(995)	(471)						(29)	(8)
				Human Capital and Demographic						
Years of executive	2	191	-42	-9.66	-32.8	1.01	14.04	34.76	82.8	24.44
Experience	(25)	(26)	(14)	(15)	(9)	(12)	(7)	(11)	(13)	(6)
Years of executive				6.09	8.22	-3.08	-3.79	-10.1	-20.6	-11.38
Experience Squared				(6)	(4)	(6)	(3)	(5)	(7)	(3)
Years of tenue	-40	-23	22	-10.43	-23.03	-10.56	10.7	25.8	-302	24.66
with firm	(20)	(25)	(14)	(9)	(6)	(8)	(5)	(8)	(9)	(4)
Years of tenue				4.8	7.28	4.19	-2.79	-11.85	88.1	-7.26
with firm squared				(4)	(3)	(4)	(2)	(4)	(4)	(2)
Number of firms worked				-8.99	-0.91	-6.86	3.11	8.11	-11.1	5.54
for before becoming an executive				(2)	(2)	(2)	(1)	(1)	(2)	(1)
Number of firms worked	215	-484	-58	-1.23	1.35	-2.43	-0.34	-0.25	-13.7	4.49
for after becoming an executive	(80)	(174)	(93)	(2)	(1)	(2)	(1)	(2)	(2)	1
Age	281	17	15	-9.01	1024	174	-459	-847	1948	-528
	(85)	(23)	(10)	(188)	(124)	(158)	(86)	(128)	(239)	(59)
Age Squared	-3.05			136.00	5.20	-111.00	236.00	434.00	-992	313
	(0.80)			(88.70)	(60.30)	(80.30)	(44.40)	(65.80)	(122)	(28.81)
Female Dummy										17.42
Dummy										(4)
Female Dummy	2668								-0.51	
X R 2	(1,295)								(0.24)	

Table 5: Compensating Differential for Nonpecuniary Cost of Diligence versus Exit

Variable	Constant	Age-50	Tenure	Eex. Exp	NBE	NAE	Female	Interlock	Execidir
Constant	1.628	0.007	0.016	-0.004	-0.006	0.025	-0.043	-0.074	0.333
	(0.071)	(0.001)	(0.002)	(0.000)	(0.004)	(0.001)	(0.024)	(0.011)	(0.049)
Rank 1	0.205						0.219	-0.125	-0.564
	(0.063)						(0.020)	(0.010)	(0.042)
Rank 2	0.263						0.347	-0.070	-0.545
	(0.063)						(0.020)	(0.008)	(0.034)
Rank 3	0.111						-0.072	-0.070	-0.545
	(0.063)						(0.020)	(0.008)	(0.034)
Rank 4	-0.181								
	(0.063)								
Industrial Sector									
Primary	-0.241	-0.006	-0.008	0.003	0.000	-0.009	0.106	0.034	0.005
	(0.048)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.018)	(0.009)	(0.037)
Service	0.400	0.009	0.008	0.002	-0.012	0.003	0.091	-0.038	0.017
	(0.050)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.019)	(0.009)	(0.038)
Firm Size									
Medium	-0.373	-0.009	-0.010	0.001	0.021	-0.002	-0.080	0.042	-0.045
	(0.050)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.019)	(0.009)	(0.038)
Large	-0.553	-0.016	-0.012	0.004	0.033	-0.006	-0.063	0.068	-0.067
	(0.049)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.019)	(0.009)	(0.038)
Board with a large number of insiders									
Large	-0.238	0.000	0.005	-0.003	0.004	0.008	-0.023	-0.036	-0.095
	(0.040)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.015)	(0.007)	(0.031)
Firm-to-firm transition									
New Employer	-0.380	0.001	0.008	-0.002	-0.004	0.004	-0.020	0.026	0.111
	(0.040)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.015)	(0.007)	(0.031)

Table 6: Compensation for Market Demand Versus Exit

Variable	Constant	Age-50	Tenure	Eex. Exp	NBE	NAE	Female	Interlock	Execidir
Constant	-0.569	-0.003	-0.007	0.002	-0.003	-0.010	0.069	-0.034	-0.320
	(0.013)	(0.001)	(0.001)	(0.000)	(0.003)	(0.001)	(0.014)	(0.008)	(0.029)
Rank 1	-0.151						-0.219	0.094	0.458
	(0.013)						(0.013)	(0.007)	(0.027)
Rank 2	0.022						-0.181	0.050	0.486
	(0.013)						(0.013)	(0.006)	(0.022)
Rank 3	0.019						-0.050	0.050	0.486
	(0.013)						(0.013)	(0.006)	(0.022)
Rank 4	0.182								
	(0.013)								
Industrial Sector									
Primary	0.048	0.002	0.004	-0.003	0.001	0.006	-0.124	-0.009	0.017
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)
Service	-0.006	-0.002	0.001	-0.001	0.004	-0.001	-0.045	0.011	0.006
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)
Firm Size									
Medium	0.032	0.001	0.002	-0.002	-0.010	0.000	0.029	0.009	0.010
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)
Large	0.170	0.005	0.000	-0.004	-0.020	0.003	-0.003	-0.044	-0.010
	(0.010)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.012)	(0.007)	(0.024)
Board with a large number of insiders									
Large	-0.117	0.000	-0.008	0.002	-0.007	-0.009	0.000	0.002	0.052
	(0.008)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.010)	(0.005)	(0.020)
Firm-to-firm transition									
New Employer	-0.085	0.000	-0.013	0.003	-0.007	-0.010	-0.006	-0.060	-0.111
	(0.008)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.010)	(0.005)	(0.020)

Table 7: Value of Human Capital Investment

Variable	Constant	Age-50	Eex. Exp	NBE	NAE	Female	Interlock	Execidir
Constant	-0.2278	0.0013	0.0014	0.0058	0.0050	0.0182	-0.0015	0.0004
	(0.0002)	(0.0005)	(0.0006)	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0001)
Rank 1	0.0237			0.0001	0.0003	0.0033	0.0001	0.0000
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0006)	(0.0002)
Rank 2	-0.0632			0.0006	0.0007	0.0017	-0.0003	0.0000
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)
Rank 3	-0.0372			0.0012	0.0012	0.0070	-0.0002	-0.0001
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)
Rank 4	-0.0062			0.0005	0.0005	0.0026	0.0000	-0.0001
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)
Firm-to-firm transition								
New Employer	-0.0132			0.0006	0.0006	0.0036	-0.0001	-0.0001
	(0.0005)			(0.0001)	(0.0001)	(0.0006)	(0.0007)	(0.0002)

Table 8: Risk Premium From Agency

Variable	Constant	Age-50	Tenure	Eex. Exp	NBE	NAE	Female	Interlock	Execidir
Constant	0.499	-0.046	-0.019	-0.012	0.032	0.190	-0.268	-0.333	-0.507
	(0.736)	(0.005)	(0.004)	(0.002)	(0.011)	(0.005)	(0.195)	(0.049)	(0.224)
Rank 1	0.569					0.000	-0.660	0.140	0.177
	(0.125)					(0.000)	(0.069)	(0.045)	(0.112)
Rank 2	2.836					-0.001	2.338	0.058	0.081
	(0.125)					(0.000)	(0.069)	(0.037)	(0.092)
Rank 3	1.032					-0.002	-1.120	0.058	0.081
	(0.125)					(0.000)	(0.069)	(0.037)	(0.092)
Rank 4	-0.016					0.000	-0.003	0.000	0.000
	(0.125)					(0.000)	(0.001)	(0.001)	(0.000)
Industrial Sector									
Primary	-0.037	-0.001	-0.001	0.000	0.012	0.011	0.142	-0.013	0.083
	(0.096)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.061)	(0.039)	(0.100)
Service	0.379	-0.049	-0.003	0.010	0.035	-0.061	-0.595	0.749	0.639
	(0.098)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.062)	(0.040)	(0.101)
Firm Size									
Medium	1.032	0.016	0.003	0.004	-0.033	0.007	0.513	-0.240	0.375
	(0.098)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.062)	(0.040)	(0.101)
Large	3.350	0.030	0.004	0.001	-0.064	0.002	0.495	-0.312	0.531
	(0.097)	(0.004)	(0.003)	(0.002)	(0.010)	(0.004)	(0.061)	(0.040)	(0.101)
Board with a large number of insiders									
Large	0.270	0.006	0.006	0.003	-0.022	-0.004	0.049	-0.085	0.088
	(0.079)	(0.003)	(0.003)	(0.001)	(0.008)	(0.004)	(0.050)	(0.032)	(0.082)
Firm-to-firm transition									
New Employer	0.362	0.008	-0.003	-0.003	0.012	0.025	0.258	-0.053	-0.062
	(0.080)	(0.003)	(0.003)	(0.001)	(0.008)	(0.004)	(0.051)	(0.033)	(0.083)

Table 9: Gross Loss to Shareholders from not Providing Executive Incentives

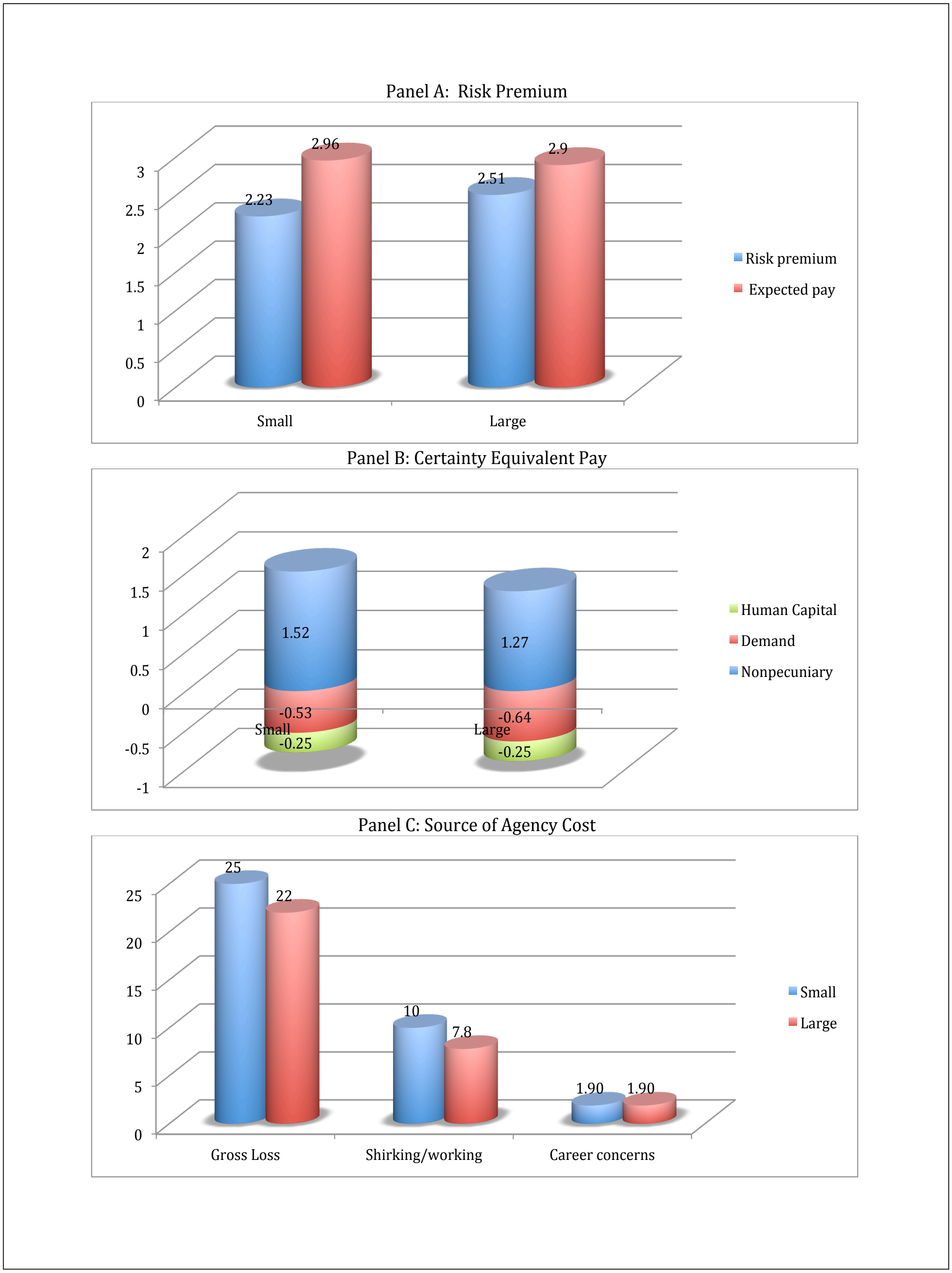
	$E(x(1-g(x)))$	New Employer	Female	Individual	Characteristics
Constant	33.5963	6.8678	1.7380	Interlocked	-3.0951
	(0.0367)	(0.0036)	(0.0263)		(0.0100)
Rank 1	-8.0575	1.0166	-1.5638	Execdir	-7.0620
	(0.0056)	(0.0395)	(0.0358)		(0.0051)
Rank 2	-4.2791	2.8547	-1.7018	Exec.Exp.	-0.1339
	(0.0057)	(0.0412)	(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			NAE	0.4477
	(0.0067)				(0.0018)
Rank 3 Lagged	-3.5289			NBE	0.5651
	(0.0080)				(0.0015)
Rank 4 Lagged	-0.4527			Age-50	-0.0411
	(0.0049)				(0.0005)
Industrial Sector				Age-50 Sq	0.0005
Primary	-3.7273				(0.0001)
	(0.0042)				
Service	9.3501				
	(0.0043)				
Firm Size					
Medium	-12.9481		0.0093		
	(0.0044)		(0.0244)		
Large	-25.4104		0.0139		
	(0.0044)		(0.0221)		
Board with a large number of insiders					
Large	-3.0350				
	(0.0035)				

Table 10: The Net Compensating Differentials to Executives from Working Versus Shirking

Variable	Constant	Age-50	Age-50 Sq	Tenure	Eex. Exp	NBE	NAE	Female	Interlock	Execidir
Constant	9.952	0.053	-0.001	0.110	0.015	-0.067	0.141	1.437	-0.930	-0.151
	(0.888)	(0.019)	(0.001)	(0.027)	(0.000)	(0.066)	(0.031)	(0.530)	(0.190)	(0.002)
Rank 1	1.029					-0.004	-0.004	-0.378	-0.070	0.018
	(0.798)					(0.002)	(0.002)	(0.480)	(0.173)	(0.003)
Rank 2	0.759					0.000	0.000	-1.082	-0.058	0.015
	(0.798)					(0.002)	(0.002)	(0.481)	(0.144)	(0.003)
Rank 3	0.307					0.006	0.005	-1.716	-0.063	0.018
	(0.798)					(0.002)	(0.002)	(0.481)	(0.144)	(0.003)
Rank 4	0.039					-0.001	-0.003	-0.120	0.010	0.017
	(0.798)					(0.002)	(0.002)	(0.008)	(0.010)	(0.003)
Industrial Sector										
Primary	-2.599	-0.032	0.001	-0.040		-0.005	-0.080	-0.612	0.427	
	(0.605)	(0.016)	(0.001)	(0.023)		(0.055)	(0.026)	(0.419)	(0.145)	
Service	3.799	0.060	-0.001	0.080		-0.050	0.074	0.788	-0.616	
	(0.628)	(0.017)	(0.001)	(0.024)		(0.057)	(0.027)	(0.427)	(0.149)	
Firm Size										
Medium	-3.105	-0.073	0.002	-0.079		0.125	-0.061	-1.041	0.769	
	(0.628)	(0.017)	(0.001)	(0.024)		(0.057)	(0.027)	(0.427)	(0.149)	
Large	-4.500	-0.096	0.002	-0.111		0.153	-0.105	-1.207	0.766	
	(0.621)	(0.016)	(0.001)	(0.024)		(0.056)	(0.027)	(0.425)	(0.148)	
Number of Insde Executves on the board ofdirectors										
Large	-2.182	0.015	-0.001	-0.027		-0.056	-0.077	-0.415	0.149	
	(0.508)	(0.013)	(0.001)	(0.019)		(0.046)	(0.022)	(0.347)	(0.121)	
Firm-to-firm transition										
New firm	-4.755	0.051	-0.001	-0.052		-0.187	-0.189	-2.485	0.318	0.034
	(0.514)	(0.013)	(0.001)	(0.019)		(0.048)	(0.023)	(0.355)	(0.130)	(0.003)

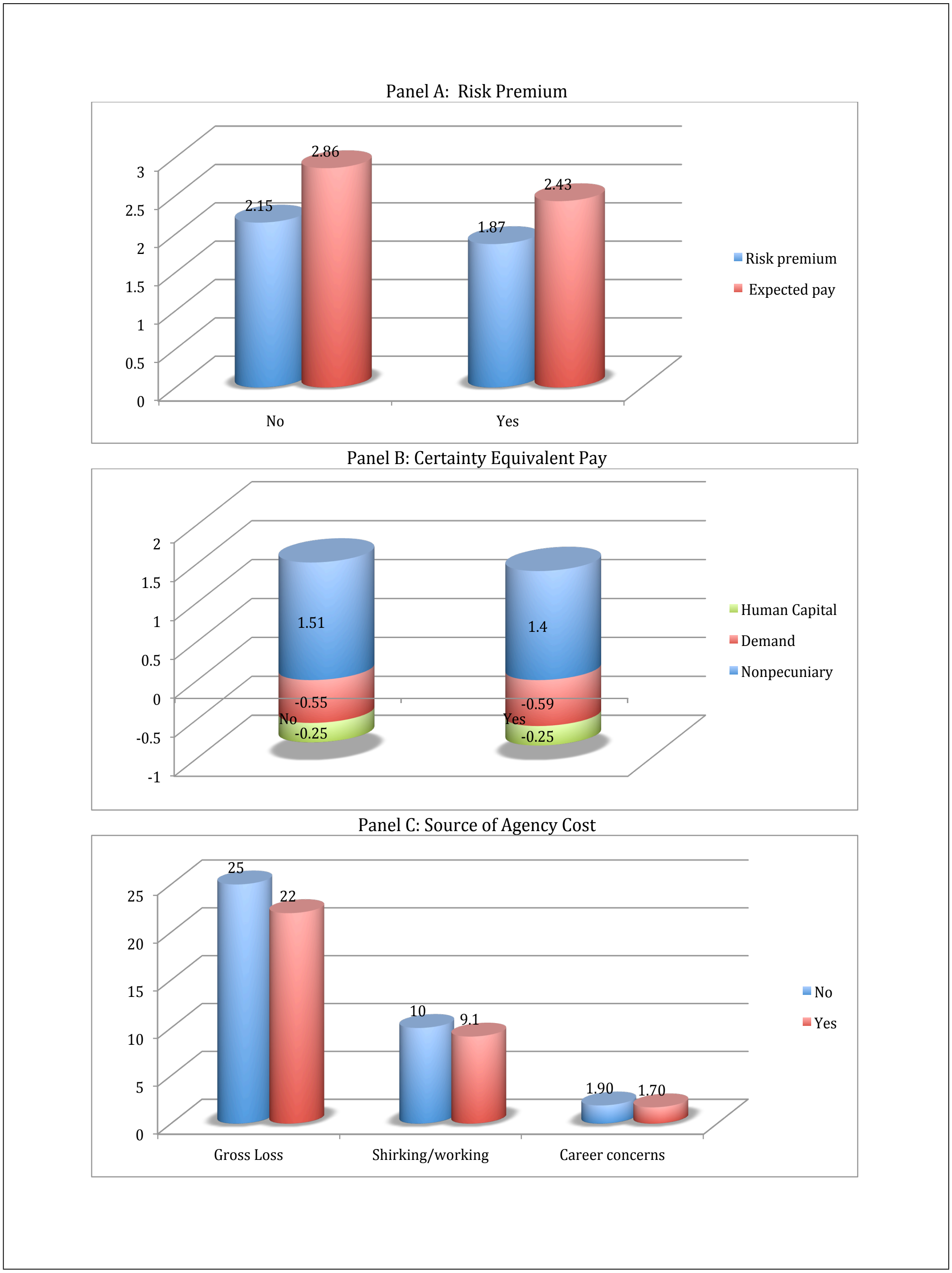
Table 11: Career Concern Amelioration of Agency Problem

Variable	Constant	Age-50	Age-50 Sq	Tenure	Eex. Exp	NBE	NAE	Female	Interlock	Execidir
Constant	-1.547	0.006	0.001	0.009	0.015	0.059	0.050	0.154	0.170	-0.151
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)	(0.002)
Rank 1	0.013					-0.004	-0.004	0.061	0.010	0.018
	(0.006)					(0.002)	(0.002)	(0.008)	(0.009)	(0.003)
Rank 2	-0.490					0.000	0.000	-0.198	0.011	0.015
	(0.006)					(0.002)	(0.002)	(0.009)	(0.010)	(0.003)
Rank 3	-0.671					0.006	0.005	0.182	0.007	0.018
	(0.006)					(0.002)	(0.002)	(0.009)	(0.010)	(0.003)
Rank 4	-0.242					-0.001	-0.003	-0.120	0.010	0.017
	(0.006)					(0.002)	(0.002)	(0.008)	(0.010)	(0.003)
Firm-to-firm transition										
New firm	-0.101					-0.017	-0.019	-0.150	0.023	0.034
	(0.006)					(0.002)	(0.002)	(0.008)	(0.009)	(0.003)



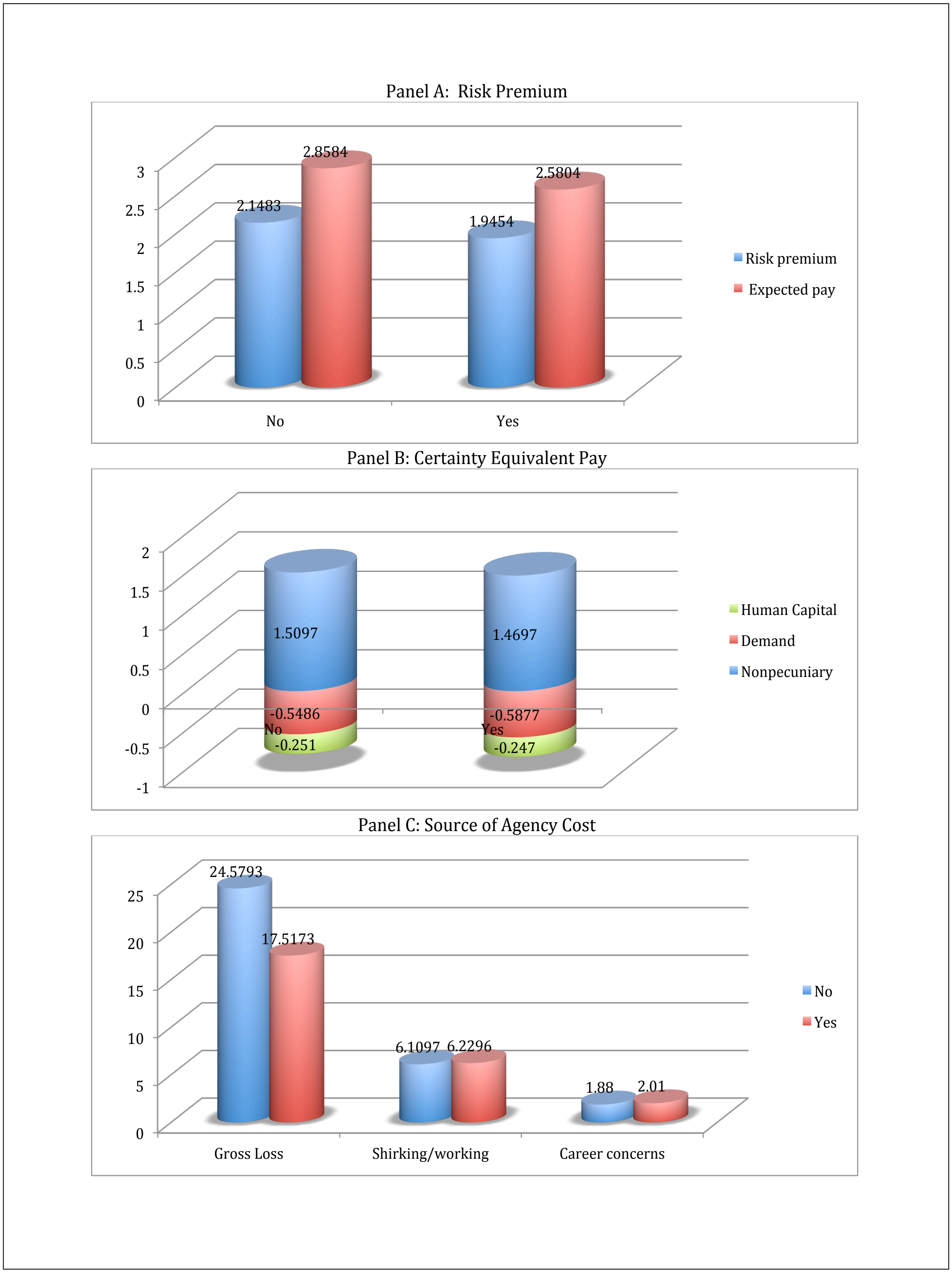
Note: Gross loss is the percentage of the firm value lost if an executive shirks instead of working. Loss of equity is the firm value lost if an executive shirks instead of working. Nonpecuniary benefit is the value to an executive of shirking relative to working. Career concerns measures the extent to which career concerns ameliorate the agency problem.

Figure 1: Governance Pay Decomposition: No. of Insider on the board.



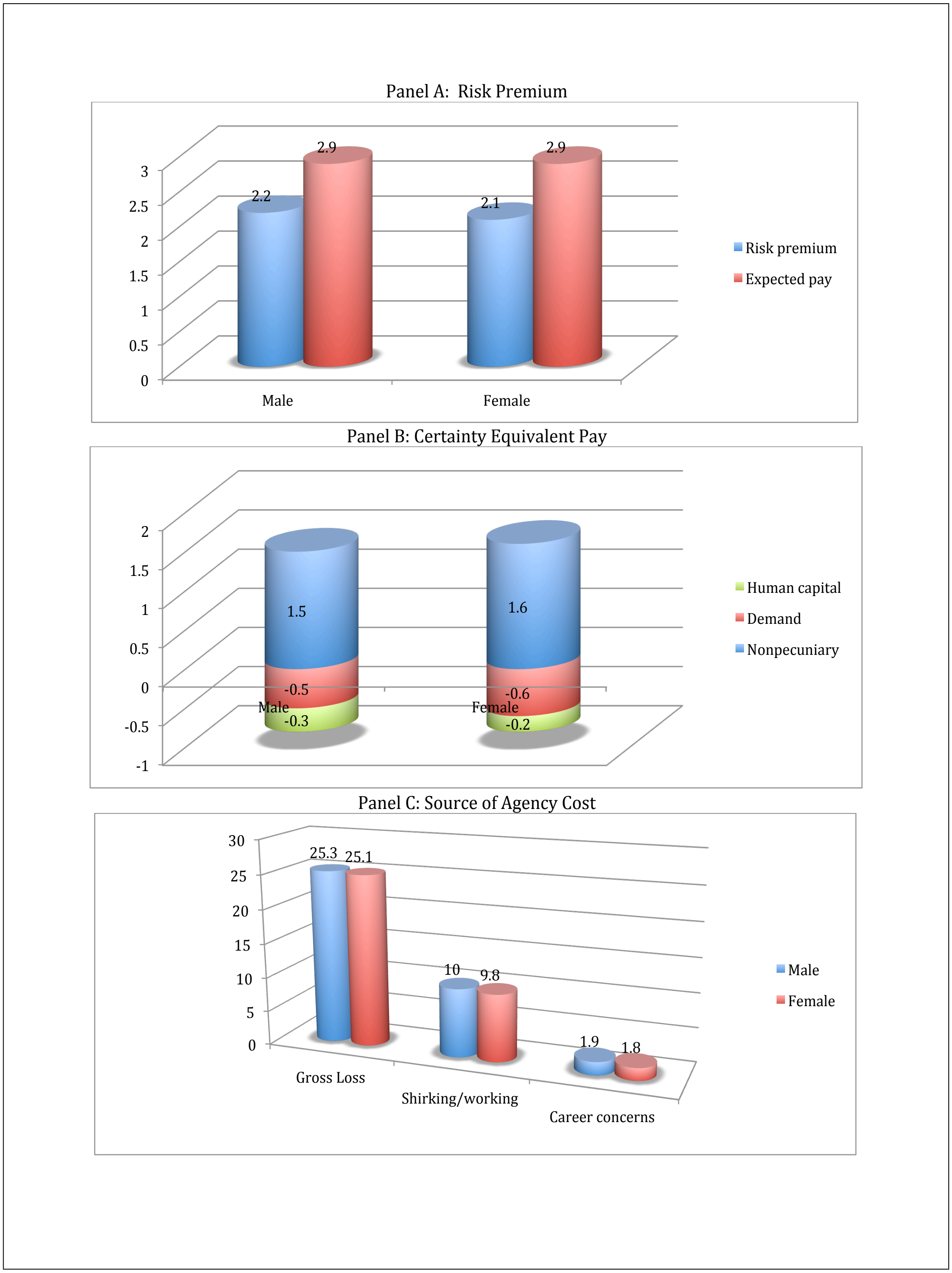
Note: Gross loss is the percentage of the firm value lost if an executive shirks instead of working. Loss of equity is the firm value lost if an executive shirks instead of working. Nonpecuniary benefit is the value to an executive of shirking relative to working. Career concerns measures the extent to which career concerns ameliorate the agency problem.

Figure 2: Governance Pay Decomposition: Interlocked Executives



Note: Gross loss is the percentage of the firm value lost if an executive shirks instead of working. Loss of equity is the firm value lost if an executive shirks instead of working. Nonpecuniary benefit is the value to an executive of shirking relative to working. Career concerns measures the extent to which career concerns ameliorate the agency problem.

Figure 3: Governance Pay Decomposition: Executive Director



Note: Gross loss is the percentage of the firm value lost if an executive shirks instead of working. Loss of equity is the firm value lost if an executive shirks instead of working. Nonpecuniary benefit is the value to an executive of shirking relative to working. Career concerns measures the extent to which career concerns ameliorate the agency problem.

Figure 4: Governance Pay Decomposition: Gender