Compensation, Moral Hazard, and Talent Misallocation in the Market for CEOs

Liang Chen
Department of Economics, Wuhan University
February 9, 2017

Abstract

This paper develops a structural model to quantify the efficiency loss arising from both risk sharing and talent misallocation due to the presence of moral hazard in the market for CEOs. I estimate the model parameters characterizing CEOs’ preferences, firms’ technologies, and their productive type distributions (firms’ size and CEOs’ talent) from the data on firms’ market value, financial returns and CEO compensation. The estimation results show that more leveraged firms and/or firms with more assets tend to end up with less-talented CEOs. By using the estimates of model parameters, I conduct counterfactuals to show that the talent misallocation led to a loss up to USD 1.6 billion dollars for the 1000 largest US firms in 2011. This amount is almost three times the loss arising from the risk-sharing inefficiency.

Key Words: CEO Compensation, Moral Hazard, Talent Misallocation

JEL Codes: J21, J30, J31, L11

---

1I am very grateful to my supervisor Eugene Choo for his invaluable guidance. I also thank the seminar participants at Carnegie Mellon University, Nanjing University, University of Calgary, University of Toronto, and CEA 2013 annual conference for their helpful comments. All remaining errors are my own.
Compensation will be used to attract, retain, and motivate employees and to reward the achievement of business results through the delivery of competitive pay and incentive programs. (2007 proxy statement, the Ford Motor company)

1 Introduction

CEOs receive very large compensation to create value for their firms. Furthermore, a large part of their compensation is linked to the firm’s performance. This type of performance-based pay is commonly attributed to the agency theory, which shows that optimal contracting compensation depending on a firm’s performance can solve the moral hazard problem present in the market. Moral hazard arises when CEOs have interests that conflict with those of shareholders and when their actions are not observable to shareholders. CEOs have incentives to pursue their interests. In this situation, performance-based pay can align a CEO’s interests with those of shareholders. Previous studies have focused on deriving the optimal contracting compensation, testing the agency theory, and estimating the agency cost that firms need pay to CEOs, most of these studies considered the assignment of CEOs to firms as given.

This paper proposes a structural model to illustrate how moral hazard affects the assignment of CEOs to firms in the market equilibrium. I estimate the model parameters to quantify the efficiency loss from both risk sharing and the inefficient assignment of CEOs across firms arising from solving the moral hazard problem through the optimal contracting compensation. A recent theory suggests that moral hazard has important implications on assignment patterns in various markets. However, most existing empirical papers ignore

---

2 According to the estimates from the American Federation of Labor and Congress of Industrial Organizations (AFL-CIO), CEOs from S&P 500 firms received the mean compensation of 12.3 million dollars in 2012 and about 54% of the compensation is from stocks and options awards.

3 The ways in which CEOs pursue their own interests include consuming perks, investing based on their own interests, and etc.

the impact of moral hazard on the assignment patterns.\textsuperscript{5} This paper tries to link the theory with empirics and quantitatively investigates how moral hazard affects the assignment of CEOs to firms. Understanding this question may guide us in implementing other substitute policies for solving the moral hazard problem, such as corporate governance.\textsuperscript{6}

My model describes that in a competitive labor market, risk-averse CEOs differing in talent are hired by risk-neutral firms varying in three characteristics: size, risk and the nonpecuniary benefits of managerial effort. The production output is stochastic and complementary to CEOs’ talent and firms’ size. Moreover, CEOs have private information on their efforts, and all characteristics of CEOs and firms are observable to all market participants but not to researchers.\textsuperscript{7} Under these circumstances, moral hazard arises, and CEOs have incentives to pursue their interests. I assume that firms employ the optimal contracting compensation to align the interests of CEOs to those of shareholders. The optimal contracting compensation drives the assignment of CEOs to firms in the market equilibrium.

I solve the model under three main simplification assumptions: (i) Constant Relative Risk Aversion (CRRA) preferences of CEOs, (ii) two levels of efforts that managers can choose from, and (iii) complementarity of the production function between firms’ size and CEOs’ talent. Under these assumptions, the optimal contracting compensation aligning the interests of CEOs to those of shareholders is the product of a certainty equivalent and a risk premium term, where the former depends on firms’ nonpecuniary benefits of managerial effort and CEOs’ talent and the latter depends on the risk of firms. Riskier firms and/or firms involving higher nonpecuniary benefits of managerial effort then need to pay more compensation to more-talented CEOs. For a given firm, although hiring a more-talented CEO results in greater production, it also incurs higher CEO compensation. Therefore, the equilibrium assignment of CEOs to firms depends not only on firms’ size but also on firms’

\textsuperscript{5}To the best of my knowledge, the only exception is Gayle et al. (2015). They consider the interaction between moral hazard and sorting of managers across and within firms by ignoring the competition for scarce managerial talent.

\textsuperscript{6}Core et al. (1999) showed that firms with weaker governance have greater agency problem.

\textsuperscript{7}The background information on top managers, such as managers’ education and working experience is public to everyone and public traded firms’ information is also quite available to everyone.
nonpecuniary benefits of managerial effort and their risk. I show that the equilibrium assignment would exhibit positive assortative sorting by managerial talent and firms’ “effective size”, which adjusts their actual size by their risk and nonpecuniary benefits of managerial effort. This equilibrium assignment may be inefficient because the efficient assignment exhibits positive assortative sorting by firms’ actual size and managerial talent when the production technology is complementary between the firms’ actual size and managerial talent.

The primary goal of this paper is to quantify the efficiency loss from the inefficient assignment of CEOs to firms in the market equilibrium. In doing so, I estimate the model parameters characterizing CEOs’ preferences, firms’ technologies, and their productive type (firms’ size and CEOs’ talent) distributions. I recover the distributions of firms’ size and CEOs’ talent from their observed incomes, firms’ profit and CEO compensation. From the literature on labor, we know that distributions of productive types for firms and workers can be inferred from their income distributions (see e.g., Sattinger (1979), Sattinger (1993), and Terviö (2008)). The production function is then identified from the productive type distributions of firms and CEOs by using the complementarity assumption and data on financial returns of firms. Finally, I follow Gayle et al. (2015) to identify managers’ preferences by exploiting the optimal compensation contract, managers’ participation and incentive compatibility constraints.

The model is estimated using data on the 1000 largest US firms from S&P Compustat databases for year 2011. The estimation results show that firms’ size is highly skewed to the right, whereas CEOs’ talent does not differ substantially. These estimation results are consistent with findings in literature (see, e.g., Terviö (2008) and Jung and Subramanian (2013)). I also find that more leveraged firms and/or firms having fewer employees are riskier. CEOs working for firms with more employees and assets incur more costs from exerting effort. Thus, more leveraged firms and/or firms with more assets have smaller effective size, and they tend to end up with less-talented CEOs. The relative risk-aversion parameter of CEOs is estimated to be 1.18, which is smaller than the estimates found in previous studies (see,
e.g., Friend and Blume (1975), Mankiw (1985), and Garcia et al. (2003)). This may be because CEOs are assumed to be relative risk averse, and they are usually wealthier than employees and workers in other roles.

By using the estimated parameters, I conduct counterfactuals to quantify the efficiency loss arising from both risk sharing and inefficient assignment of CEOs to firms due to moral hazard in the market equilibrium. In particular, I quantify three measures of the efficiency loss due to moral hazard. First, I estimate the total loss arising from risk-sharing inefficiency over all firms, which is about USD 588 million. Second, I quantify the aggregate loss from the CEOs’ misallocation in the market equilibrium, which ranges between USD 0.6 and 1.6 billion for all firms. This loss can be almost three times as large as that from the risk-sharing inefficiency. The loss is 10.2%-27.7% of the total compensation for all CEOs. Third, the aggregate loss from a random assignment of CEOs to firms is estimated to be between USD 184.9-416.9 billion. This loss is 1.72%-3.94% of the total market value for all firms. These results suggest that the assignment of CEOs to firms is important, and studies focusing solely on risk sharing may severely underestimate the efficiency loss arising from moral hazard.

Related Literature
This paper is closely related to two lines of literature. First, this paper contributes to the growing literature on executive compensation. One stream of this literature focuses on the optimal contracting for motivating executives in the presence of moral hazard. Theoretical papers include Mirrlees (1976), H¨ olmstrom (1979); Holmstrom (1982), Grossman and Hart (1983) and Holmstrom and Milgrom (1987) on the static optimal contracting, and Fudenberg et al. (1990) and Edmans et al. (2012) on the dynamic optimal contracting. The implication of these papers is that executive compensation should be linked with firms’ performance to provide executives with incentives to maximize the firms’ value. This effect has been empirically examined and tested by using the detailed data on executive compensation and firm performance. For example, seminal works by Jensen and Murphy (1990), Hall and Liebman
(1998) and Aggarwal et al. (1999) found significant managerial pay-performance sensitivities. Garen et al. (1994) found that the executive compensation structure is consistent with the intuitions of the contract theory. Gayle and Miller (2009) estimated a principal-agent model of moral hazard to show that more complex firms have greater agency problem and pay their executives more.

Another stream of the literature employs assignment models to explain the observed variation and trend of executive compensation in past decades. This literature suggests that the patterns of executive compensation can be explained by the size of firms because executives generate more value at larger firms. Terviö (2008) found that the variation in CEO compensation can be explained by the difference in firm size. Gabaix and Landier (2008) showed that the six-fold increase in executive compensation in the US since 1980 could be attributed to the simultaneous growth of the firm size. With a few exceptions, most previous papers examined moral hazard and the assignment of executives separately. This paper incorporates moral hazard into an assignment model to study how it affects the assignment of CEOs to firms and to evaluate the importance of this impact.

Two closely related papers are Edmans and Gabaix (2011) and Gayle et al. (2015), which also investigate moral hazard and the assignment of executives together. Edmans and Gabaix (2011) provided a theory to show how moral hazard can distort the efficient assignment of CEOs to firms and to calibrate the loss from this distortion. This paper aimed to draw theoretical conclusions regarding CEO compensation. My paper differs from it in terms of the empirical features. While my model is built for empirical purpose and the model parameters are identified and estimated, my analysis can be used to quantify the loss from inefficient talent assignment due to moral hazard for each firm. Gayle et al. (2015) developed a generalized Roy model with human capital accumulation, moral hazard, and career concerns to decompose the firm-size pay gap. They found that total compensation increasing with firm size is mainly attributed to incentive pay in terms of risk premium. Their model is a much more sophisticated dynamic model that includes many important
features in the executive market. However, although they considered that executives are sorted into different firms and position ranks within firms, they did not consider competition for scarce talented executives. My focus is on how moral hazard affects firms to compete for executives, as a result of which the distributions of firm size and executive talent play important roles in my analysis.

Second, this paper is also related to the literature on identifying and estimating matching and principal-agent models. The literature focuses mainly on the pure matching or pure principal-agent models. For instance, with regard to matching models, Choo and Siow (2006) first proposed the identification of a marriage matching model with transferable utility. Fox (2008) and Galichon and Salanié (2010, 2012) extended Choo and Siow (2006) to more general cases. Hitsch et al. (2010), Menzel (2015), and Agarwal (2015) presented the identification and estimation of the matching models with nontransferable utility. Margiotta and Miller (2000) estimated a pure principal-agent model to quantify the costs of moral hazard. Perrigne and Vuong (2011), Luo et al. (2014), Gayle et al. (2015), Li (2015), and Luo (2015) exploited predictions from static contracting theory to achieve the nonparametric identification in models of moral hazard and adverse selection and nonlinear pricing in various settings. This paper estimates a matching model in the presence of moral hazard, which allows me to empirically investigate the importance of moral hazard in the assignment of CEOs to firms.

The remainder of the paper is organized as follows. Section 2 presents the theoretical model on which my empirical analysis is based. Section 3 introduces data on the US labor market for CEOs in 2011. Section 4 proposes a new empirical method to recover firms’ production and CEOs’ preferences in a matching market with moral hazard. Section 5 presents the estimation and counterfactual results. Finally, section 6 presents the conclusions of this study. All proofs are collected in the Appendix.
2 Proposed Model

This section lays out a static assignment model of CEOs (also referred to as managers hereafter) to firms featuring moral hazard. This model is simplified so as to mainly focus on the importance of moral hazard. Risk-averse managers have private information on their own effort levels, which is the only market friction in the model. In each period, they make the employment decisions, propose cost-minimizing contracts to risk-neutral shareholders of firms, and then select their effort levels. Firms choose their managers to maximize their expected value subject to moral hazard from not observing the effort levels of managers.

2.1 Managers and Firms

There is a continuum of risk-averse managers in a competitive labor market for CEOs. Managers are assumed to differ only in talent, which refers to the combination of all their characteristics that contribute to production, such as background, education, working experience, and so on. I employ a quantile function to capture the distribution of managerial talent. Let $T(m)$ denote the talent of a quantile $m$ manager.\(^8\) There is a continuum of risk-neutral firms in the market.\(^9\) I follow Edmans and Gabaix (2011) and characterize firms by three characteristics: size, risk, and nonpecuniary benefits of managerial effort. Firm size refers to the combination of firms’ all characteristics that contribute to production, such as assets, reputation, capitalization ability, and so on. Its distribution is also captured by a quantile function, with $S(n)$ denoting the size of a quantile $n$ firm. The risk of firms is reflected by the volatility of their production output. The nonpecuniary benefits of managerial effort, are assumed to be firm-specific, meaning that managers incur different costs when serving for firms with different sizes, locations, industries, etc.

---

\(^8\)As noted in Terviö (2008), this quantile representation is more intuitive and tractable than traditional distribution functions in empirical applications.

\(^9\)The continuous distributions of managers and firms rule out the match-specific rents and therefore, any need to model bargaining between firms and CEOs.
2.2 Preferences

Managers’ preferences are captured by Constant Relative Risk Aversion (CRRA) utility functions. I assume managers do not have existing wealth. The utility received by a manager with talent quantile $T(m)$ serving for a firm of size $S(n)$ is then written as,

$$U(w_{nm}) = \frac{[\alpha_{ne} \cdot w_{nm}]^{1-\rho}}{1-\rho}, \text{ for } e \in \{1, 2\},$$

where $w_{nm}$ is the compensation of manager $m$ serving for firm $n$. $\alpha_{ne}$ is a firm-specific nonpecuniary benefits from making effort for managers serving for firm $n$, where $e = 1$ represents shirking and $e = 2$ represents working. As in standard moral hazard models, shirking and working refer to activities that benefit to managers and firms, respectively, meaning that the goals of managers and shareholders of firms are not aligned. I assume that working is more costly, and therefore, managers receives less nonpecuniary benefits when they work, $\alpha_{n1} > \alpha_{n2}$. $\rho$ is the risk-aversion parameter, which is restricted to be greater than one ($\rho > 1$).\(^{10}\) If the manager $T(m)$ is not hired by any firm, he/she receives a reservation utility from his/her outside options, denoted by $[\alpha_{m0}]^{1-\rho}/[1-\rho]$, where $\alpha_{m0}$ is the nonpecuniary benefit of manager $m$ from outside options.

I use CRRA utility functions instead of the Constant Absolute Risk Aversion (CARA) utility functions typically used in literature for two reasons. First, it helps to solve the complex assignment problem where firms differ in multidimensional characteristics, thereby allowing me to aggregate the multidimensional characteristics of firms into a single index. Second, this choice is supported by the evidence provided in previous studies (see, e.g., Friend and Blume (1975); Garcia et al. (2003); Edmans et al. (2009)). It suggests that managers’ utility from exerting effort tends to increase proportionally with their consumptions, which can be derived from CRRA utility functions.

\(^{10}\)In principle, the risk aversion parameter $\rho$ can be any value greater than zero. This restriction gives us tractable optimal compensation contracts for estimation. The estimates of the risk aversion parameter vary widely in the literature with most of them are higher than one (see e.g., Friend and Blume (1975); Weber (1975); Garcia et al. (2003); Mankiw (1985)).
2.3 Firm Technology

The production from the combination of a manager with talent $T(m)$ and a firm with size $S(n)$ is assumed to be multiplicatively separable between their contributions,

$$Y(S(n), T(m), e) = S(n) \cdot T(m) \cdot [1 + x(e)],$$

which is complementary between the manager and the firm.\(^{11}\)

Managerial effort affects production through the stochastic variable $x(e)$. Its probability density function depends on the effort exerted by the manager. For firm $n$, conditional on shirking of its manager, $e = 1$, the probability density function of $x$ is denoted by $f_{n1}(x)$; conditional on working, $e = 2$, the probability density function of $x$ is denoted by $f_{n2}(x)$. Because $x$ is the only stochastic variable in the production function, variances of $f_{n1}(x)$ and $f_{n2}(x)$ reflect the risk of firms. Hereafter, I use the symbol $E[\bullet]$ to represent the expectation taken with respect to $f_{n2}(x)$, $E[\bullet] = \int \bullet f_{n2}(x)dx$. I follow the literature and define the likelihood ratio function, $g_{n}(x) \equiv f_{n1}(x)/f_{n2}(x)$. It is nonnegative for all $x$, $g_{n}(x) \geq 0$, and satisfies $E[g_{n}(x)] = \int f_{n1}(x)dx = 1$.\(^{12}\)

I now make three assumptions on the two probability density functions, $f_{n1}(x)$ and $f_{n2}(x)$. These assumptions follow Gayle et al. (2015), and they are crucial for solving and estimating the model. First, I normalize the expected value of stochastic variable $x$ conditional on working of their managers to be zero for all firms.

**Assumption 1.** $E_{n}[x]$ is normalized to be zero for any firm $n$, that is, $E_{n}[x|e = 2] = 0$.

Second, I assume that each firm prefers its manager working to shirking, that is, the expectation of $x$ is larger when the manager works than when the manager shirks.

**Assumption 2.** $E_{n}[x] \equiv \int xf_{n2}(x)dx > \int xf_{n1}(x)dx \equiv E_{n}[xg_{n}(x)]$ for any firm $n$.

\(^{11}\)I refer readers to Terviö (2008) for the micro-foundations of it.

\(^{12}\) $g_{n}(x)$ can be interpreted as the signal that firm $n$ receives about managerial effort. $g_{n}(x) = 0$ implies working while $g_{n}(x) = \infty$ implies shirking. When $g_{n}(x) = 1$, the signal is useless such that the compensation does not depend on $x$ for any firm $n$. 

10
This assumption, together with the manager’s preference for shirking to working, $\alpha_{n1} > \alpha_{n2}$, provides the conflicts of interests between managers and firms. Finally, I assume that a very large value of $x$ is extremely unlikely to be obtained if the manager shirks.\footnote{As in Gayle et al. (2015), this assumption rules out the possibility of setting a contract that is arbitrarily close to the first-best contract. Mirrlees (1999) first notes that we can always find a contract that is close to the first-best resource allocation by severely punishing the agent when $g(x)$ takes an extremely high value. He also proposes that assuming $g(x)$ is decreasing can rule out this possibility, which is stronger than Assumption 3.}

**Assumption 3.** $\lim_{x \to \infty} g_n(x) = 0$ and $g_n(x)$ is bounded for any firm $n$.

### 2.4 Information and Timing

**Information.** Each manager has private information on his/her own effort level. All other information is symmetric. Everyone observes the talent of all available managers in the market, and therefore, the managerial talent distribution. The background information on top managers, such as managers’ education and working experience, is public to everyone. This information can help shareholders to identify managers’ talent. Everyone also observes the size, nonpecuniary benefits, and risk of all firms in the market. In this study, we focus on publicly traded firms, whose performance information is easily available in the financial market.\footnote{These assumptions rule out any modeling on signaling.} Therefore, the hidden action moral hazard is the only market friction in the model.

**Timeline.** At the beginning of each period, all information on all managers’ talent, and all firms’ size, nonpecuniary benefits, and risk are released to everyone. Each manager then first decides whether to retire or not; and if he/she decides to not retire, he/she proposes take-it-or-leave-it performance-based compensation contracts to all firms. Given the compensation contracts, every firm makes a choice on accepting a manager’s offer to maximize its expected profit, which equals the expected production net of the expected managerial compensation. If a manager chooses to retire, he/she obtains a reservation utility from her outside option. If a manager chooses to propose contracts and is accepted by a firm, he/she then chooses his/her effort level of either shirking or working. Finally, the stochastic variable $x$ in the
production is realized. Each manager gets paid according to the accepted compensation contract. Each firm receives the expected profit accordingly.

2.5 Cost-Minimizing Contracts

To solve the equilibrium assignment of CEOs to firms, I now solve the cost-minimizing contracts of all managers to all firms, corresponding to both effort levels of shirking and working. Consider a talent $T(m)$ manager’s contracting problem of designing cost-minimizing contracts for a firm with size $S(n)$.

The contract corresponding to shirking minimizes the expected compensation, $\int w_{nm}(x)f_{n1}(x)dx$, subject to the participation constraint,

$$\int \left[ \alpha_{n1} w_{nm}(x) \right]^{[1-\rho]} f_{n1}(x) dx \geq \frac{\alpha_{m0}^{[1-\rho]}}{1-\rho},$$

which ensures that the manager at least receives his/her reservation utility from outside options. At the optimum, the participation constraint holds with equality. Thus, the cost-minimizing contract corresponding to shirking must be a flat compensation, $\alpha_{m0}/\alpha_{n1}$.

The contract corresponding to working minimizes the expected compensation, $\int w_{nm}(x)f_{n2}(x)dx$, subject to (1) the participation constraint (the manager at least receives his/her reservation utility from outside options), and (2) the incentive compatibility constraint (the manager at least receives the utility from shirking). The cost-minimizing contract can be obtained by solving

$$\min_{w_{nm}(x)} \int w_{nm}(x)f_{n2}(x)dx$$

subject to

$$\int \left[ \alpha_{n2} w_{nm}(x) \right]^{[1-\rho]} f_{n2}(x) dx \geq \frac{\alpha_{m0}^{[1-\rho]}}{1-\rho},$$

and

$$\int \left[ \alpha_{n1} w_{nm}(x) \right]^{[1-\rho]} f_{n2}(x) dx \geq \int \left[ \alpha_{n1} w_{nm}(x) \right]^{[1-\rho]} f_{n1}(x) dx.$$
working.

**Proposition 1.** The cost-minimizing compensation contract that corresponds to working is given by

\[
 w_{wm}(x) = \frac{\alpha_{m0}}{\alpha_{n2}} \left[ 1 + \lambda_n \left[ \left( \frac{\alpha_{n1}}{\alpha_{n2}} \right)^{\rho-1} - g_n(x) \right] \right]^{1/\rho-1}
\]

where \( \lambda_n \) is the unique positive solution of the equation,

\[
 \int \frac{f_{n2}(x)}{1 + \lambda_n \left[ \left( \frac{\alpha_{n1}}{\alpha_{n2}} \right)^{\rho-1} - g_n(x) \right]} \, dx = 1.
\]

**Proof.** See Appendix 1.

Equation (6) shows that the cost-minimizing contract is the product of a fixed certainty equivalent compensation, \( \alpha_{m0}/\alpha_{n2} \), and a risk premium term, \( \left[ 1 + \lambda_n \left[ \left( \frac{\alpha_{n1}}{\alpha_{n2}} \right)^{\rho-1} - g_n(x) \right] \right]^{1/\rho-1} \). Firms need to pay a higher certainty equivalent compensation to a more talented CEO in equilibrium. Risky firms and/or firms involving high nonpecuniary benefits of managerial effort need to pay a high-risk premium to motivate their hired CEOs. Thus, these firms need to pay a particular high-risk premium to hire a very talented CEO.

In a pure moral hazard model, the certainty equivalent compensation, \( \alpha_{m0}/\alpha_{n2} \), is exogenous. In my model, this certainty equivalent compensation is endogenously determined, depending on the distributions of managerial talent and firm size. Intuitively, firms compete to obtain the scarce talent of managers in the labor market. The assignment of CEOs to firms would drive a more talented manager to obtain higher certain equivalent compensation. The rank of a manager’s certain equivalent compensation should then be consistent with the rank of his/her talent. Therefore, the distributions of managerial talent and firm size play important roles in determining the managerial compensation through the certainty equivalent compensation.

Whether firms will accept the contract corresponding to shirking or working depends on
which one maximizes their expected profits. Hereafter, I focus on the case in which all firms choose the contract that corresponds to working, because in the data, I observe that all firms pay performance-based compensation to their managers. As we know, any contracts corresponding to shirking must be a flat compensation.

### 2.6 Assignment of CEOs to Firms

When the production is complementary between firms’ actual size and managerial talent, efficiency requires positive assortative sorting by them, which maximizes the total expected production output. In the absence of moral hazard, the equilibrium allocation of managers to firms exhibits this sorting.\(^\text{15}\) However, when moral hazard is present, efficient positive assortative sorting may not be achieved in equilibrium. In the following, I solve the equilibrium assignment of managers to firms and show how moral hazard can distort the equilibrium assignment away from efficient sorting.\(^\text{16}\)

In equilibrium, each firm chooses a manager to maximize its expected profit, and the expected production net of the expected compensation paid to the manager under his/her cost-minimizing contracts. Formally, a size \(S(n)\) firm’s profit maximization problem is given by

\[
\max_m E[Y(S(n), T(m), e)] - E_n[w_{nm}(x)],
\]

where the first term is the expected production from hiring a talent \(T(m)\) manager and the second term, the expected compensation paid to the manager. For a firm with size \(S(n)\), hiring a more talented manager \((T(m)\) is higher) results in more expected production \((E[Y(S(n), T(m), e)]\) is higher). However, if the firm is very risky and/or involves high cost for the manager to manager, hiring a more talent manager might be very costly \((E_n[w_{nm}(x)]\)

---

\(^{15}\text{See e.g., Sattinger (1993); Gabaix and Landier (2008); Terviö (2008).}\)

\(^{16}\text{Edmans and Gabaix (2011) provided a similar theory of CEO misallocation using a different model setup.}\)
is high). Therefore, it is not obvious whether this firm should hire a more talented manager. The difficulty in solving this assignment problem is that firms have multidimensional characteristics.

To solve the equilibrium assignment of CEOs to firms, I collapse the multidimensional characteristics of firms into one index. To do so, I first derive the expected utility of managers as a function of the expected compensation that firms pay to their managers. Consider the expected utility of a manager with talent $T(m)$ hired by a size $S(n)$ firm. By using the optimal compensation contract given by (6) in Proposition 1 and the participation constraint (4), a manager’s expected utility can be written as,

$$E_n[U(w_{nm}^o(x))] = \frac{[\alpha_{n2}E_n[w_{nm}^o(x)] e^{-\chi_n}]^{1-\rho}}{1-\rho},$$

where $\chi_n$ is defined by

$$\chi_n \equiv \ln \left[ E_n \left[ \left[ 1 + \lambda_n \left[ \frac{\alpha_{n1}/\alpha_{n2}}{\alpha_{n2}} - g_n(x) \right] \right]^{\frac{1}{\rho-1}} \right] \right],$$

where $\chi_n$ is the risk premium that the firm pays to its manager in the sense that $\chi_n = \ln [\alpha_{n2}E_n[w_{nm}^o(x)]] - \ln [U^{-1} [E_n[U(w_{nm}^o(x))]]]$. $\chi_n$ increases with the variance of $g_n(x)$ and the nonpecuniary benefits $\alpha_{n1}/\alpha_{n2}$, implying that riskier firms (variance of $g_n(x)$ is larger) or firms involving higher relative nonpecuniary benefits of managerial effort ($\alpha_{n1}/\alpha_{n2}$ is larger) need to pay more risk premiums in terms of $\chi_n$.

From the manager’s perspective, $\chi_n$ is the total loss that the manager would incur from working and sharing risk with the firm. If the firm is highly risky or involves high relative nonpecuniary benefits of managerial effort, the manager suffers a high loss. After adjusting for the loss, the firm pays a fixed compensation to its manager, $v_{nm} \equiv E_n[w_{nm}^o(x)] e^{-\chi_n}$, referred to as “effective compensation”. This gives the manager the same expected utility as that under the optimal compensation contract inducing him to work. The proof of Proposition 1 shows that the participation constraint (4) is met with equality, and therefore, I also
have $v_{nm} = \alpha_{m0}/\alpha_{n2}$, which depends on the parameter in the reservation utility from outside options, $\alpha_{m0}$. The outside options of the manager are interpreted as the opportunities to manage firms whose sizes are next to firm $n$ in the competitive labor market. More talented managers should obtain higher effective compensations in equilibrium as they can obtain more reservation utility from outside options.

Formally, the effective compensation of a manager with talent $T(m)$ also has a quantile $m$, which can then be written as a function of the quantile, $v(m)$. If a firm with size $S(n)$ wishes to hire a manager with talent $T(m)$, it has to pay him/her an effective compensation $v(m)$ and thus an expected compensation of $E_n[w_{nm}(x)] = v(m)e^{\chi_n}$. Therefore, a size $S(n)$ firm’s profit maximization problem can be rewritten as

$$
\max_m E[Y(S(n), T(m), e)] - E_n[w_{nm}(x)]
= S(n)T(m) - v(m)e^{\chi_n}
= e^{\chi_n}[S(n)e^{-\chi_n}T(m) - v(m)],
$$

(10)

where the first equality is derived from the normalization $E_n[x(e)|e = 2] = 0$. Because $e^{\chi_n}$ is positive, maximizing (10) is equivalent to maximizing the effective expected output $S(n)e^{-\chi_n}T(m) - v(m)$. It suggests that a firm with actual size $S(n)$ now acts like a firm with size $S(n)e^{-\chi_n}$; this is referred to as the “effective size” of firms. I denote the effective size by $\hat{S}(r) \equiv S(n)e^{-\chi_n}$, where $r$ is the quantile of firm $n$’s effective size. Because the effective expected output given in (10) is complementary between the firm’s effective size and the manager’s talent, the allocation of managers to firms in equilibrium would exhibit positive assortative sorting by firms’ effective size and managers’ talent.

Therefore, the profiles of the effective incomes of managers and firms supporting this
allocation must satisfy two types of conditions in equilibrium,

\begin{align}
(11) \quad \tilde{S}(r)T(r) - v(r) & \geq \tilde{S}(r)T(m) - v(m), \quad \forall \ r, m \in [0, 1], \\
(12) \quad \tilde{S}(r)T(r) - v(r) & \geq \pi^0, \quad \forall \ r \in [0, 1], \\
(13) \quad v(r) & \geq v^0, \quad \forall \ r \in [0, 1],
\end{align}

where \( \tilde{S}(r)T(r) \) is the expected effective output, which is divided between the manager and the firm. \((\pi^0, v^0)\) is the lowest effective compensation and effective profit that managers and firms could obtain from the opportunities within the labor market for managers. The lowest active firm-manager pair \((r = 0)\) is the one that just breaks even with the lowest incomes from the alternative opportunities within the market, \( \tilde{S}(0)T(0) = v^0 + \pi^0 \). The first type conditions (11) guarantee each firm must prefer hiring a manager whose talent has the same quantile as that of the firm’s effective size to hiring any other managers at the equilibrium effective compensations. The second type conditions (12) and (13) guarantee that all firms and managers will be active in the labor market for CEOs.

By solving (11)-(13), I derive Proposition 2, which states the assignment of managers to firms and gives the profiles of their effective incomes supporting the assignment in equilibrium.

**Proposition 2.** When moral hazard is present, the equilibrium assignment of managers to firms exhibits positive assortative sorting by firms’ effective size \( \tilde{S}(r) \) and managers’ talent \( T(r) \). The profiles of managers’ effective compensations and firms’ effective profits supporting equilibrium are given by,

\begin{align}
(14) \quad v(r) = v^0 + \int_0^r \tilde{S}(j)T'(j) dj,
\end{align}
and

\[ \pi(r) = \pi^0 + \int_0^r \tilde{S}'(j)T(j)dj. \] (15)

**Proof.** See Appendix 1.

The expressions of the effective incomes given by (14) and (15) show that all inframarginal manager-firm pairs \((r > 0)\) produce an effective output over the sum of the lowest effective incomes from the opportunities within the labor market. The effective incomes depend on the distributions of firms’ effective size and managers’ talent. The effective incomes of firms and managers, given by (14) and (15), respectively, equal their marginal contributions to effective output in the competitive equilibrium.

The key assumption allows me to solve the complex multidimensional sorting problem is the CRRA utility function. Under this assumption, I obtain the multiplicative form of the optimal compensation inducing managers to work. The multiplicative form helps me to aggregate the three firm characteristics into a single index, called the effective size of firms. I can then unambiguously rank firms and solve the equilibrium allocation of managers to firms, which exhibits positive assortative sorting by firms’ effective size and managers’ talent. This equilibrium assignment may or may not be efficient to maximize the total expected production output for all firms. From the complementary specification of firms’ production function given by (2), we know that the efficient allocation should exhibit positive sorting in firms’ actual size and managers’ talent. Therefore, the magnitude of the inefficiency in the equilibrium allocation exhibiting positive sorting between firms’ effective size and managers’ talent depends on the difference between the distributions of firms’ actual size and that of firms’ effective size. My primary goal is to quantify and evaluate the loss from such an inefficiency.
3 Data

I estimate the model parameters using a sample comprised the 1000 largest publicly traded US firms in terms of market value and their hired CEOs in the S&P Compustat databases for 2011. Data on executive compensation is collected from the S&P Compustat Execucomp database. I only extract the information on CEO compensation for this study. The compensation data are supplemented by data on firms’ characteristics from the S&P Compustat North America database and firms’ financial returns calculated from monthly stock price data from the Center for Research in Security Prices (CRSP) database. Data on the firms’ characteristics are used to introduce heterogeneity in firms’ risk and nonpecuniary benefits of managerial effort. Data on the monthly stock prices are used to construct the abnormal returns of firms.

Industrial-level factors can affect firms’ risk and, in turn their managerial compensation. To account for this while keeping the analysis simple, I follow Gayle et al. (2015) and classify firms in the sample into three industrial sectors according to GICS code. The first is called the primary sector, and it includes firms in energy (GICS: 1010), materials (1510), industrials (2010, 2020, 2030), and utility (5510). The second is called the consumer goods sector, and it includes firms in consumer discretionary (2510, 2520, 2530, 2540, 2550) and consumer staples (3010, 3020, 3030). The third is called the services sector, and it includes firms in healthcare (3510, 3520), financial service (4010, 4020, 4030, 4040), information technology and telecommunication services (410, 4520, 4030, 4040, 5010).

3.1 CEO Compensation

As in Antle and Smith (1985, 1986), I measure CEO compensation as the sum of salary, bonuses, the value of restricted stocks and options granted, the changes in the wealth from holding stocks and options, and value of retirement and long-term compensation schemes. Compensation includes all costs to shareholders of employing a CEO and the total compen-
sation a CEO receives from serving a firm. When a firm is making a hiring decision, it cares about the costs of employing a CEO. And a CEO cares about total compensation associated with the employment. At the same time, a CEO cares about the total compensation associated with employment.

Table 1: CROSS-SECTIONAL STATISTICS ON COMPONENTS OF COMPENSATION

(In thousands of US $ (2011); standard deviations in parentheses)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Primary</th>
<th>Consumer</th>
<th>Services</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salary and bonus</td>
<td>1090.6</td>
<td>979.1</td>
<td>862.3</td>
<td>996.0</td>
</tr>
<tr>
<td></td>
<td>(1585.2)</td>
<td>(620.1)</td>
<td>(646.6)</td>
<td>(1211.8)</td>
</tr>
<tr>
<td>Value of restricted stocks</td>
<td>2377.4</td>
<td>2763.4</td>
<td>2414.7</td>
<td>2454.3</td>
</tr>
<tr>
<td>awarded</td>
<td>(2843.2)</td>
<td>(3541.3)</td>
<td>(2918.8)</td>
<td>(2995.4)</td>
</tr>
<tr>
<td>Value of options awarded</td>
<td>1090.8</td>
<td>1368.4</td>
<td>921.3</td>
<td>1080.7</td>
</tr>
<tr>
<td></td>
<td>(2027.7)</td>
<td>(3655.4)</td>
<td>(1603.4)</td>
<td>(2269.5)</td>
</tr>
<tr>
<td>Change in value of stock</td>
<td>1716.6</td>
<td>1891.9</td>
<td>1177.7</td>
<td>1566.4</td>
</tr>
<tr>
<td>holdings</td>
<td>(5183.1)</td>
<td>(6098.9)</td>
<td>(3891.6)</td>
<td>(4971.1)</td>
</tr>
<tr>
<td>Change in value of option</td>
<td>967.4</td>
<td>1932.3</td>
<td>884.6</td>
<td>1100.9</td>
</tr>
<tr>
<td>holdings</td>
<td>(5102.0)</td>
<td>(7282.5)</td>
<td>(4599.5)</td>
<td>(5383.8)</td>
</tr>
<tr>
<td>Other compensation</td>
<td>1840.2</td>
<td>2091.5</td>
<td>1297.9</td>
<td>1701.6</td>
</tr>
<tr>
<td></td>
<td>(2640.0)</td>
<td>(3134.9)</td>
<td>(1977.2)</td>
<td>(2550.1)</td>
</tr>
<tr>
<td>Total Compensation</td>
<td>9083.0</td>
<td>11026.5</td>
<td>7558.5</td>
<td>8899.9</td>
</tr>
<tr>
<td></td>
<td>(10718.3)</td>
<td>(12975.3)</td>
<td>(9245.9)</td>
<td>(10730.4)</td>
</tr>
<tr>
<td>Observations</td>
<td>373</td>
<td>207</td>
<td>420</td>
<td>1000</td>
</tr>
</tbody>
</table>

Table 1 summarizes the cross-sectional data on components of CEO compensation by sectors in my sample. I break the total compensation into six components: salary and bonus, value of restricted stocks granted, value of options granted, changes in value of stock holdings, changes in value of option holdings, and other compensation including value of retirement and non-equity incentive compensation. Salary, bonus and other compensation account for around 30.31\% of the total compensation, and the other four components collectively account for the remaining 69.69\%. It shows that a large fraction of managerial compensation is linked
to firms’ performance. Moreover, managerial compensation from holding granted financial securities has a very large standard deviation, suggesting that managerial compensation from holding granted financial securities whose value is affected by firms’ performance accounts for most of the variability in the total compensation of CEOs.

### 3.2 Abnormal Returns

I measure the abnormal returns of firms as the residual component of firms’ returns in the financial market that managers can control. In the optimal contract, the compensation should depend on this residual to provide each manager appropriate incentives; however, it should not depend on changes in stochastic factors that originate outside the firm and are not able to be controlled by managers. Specifically, I follow Gayle and Miller (2009) and impute the abnormal returns of firms in two steps by using the monthly stock price data on the 1000 largest companies in the US in 2011. First, I calculate the difference between the financial returns on the individual firm’s stock and the returns on the market portfolio. Second, I then regress this difference on constant and sector-specific dummy variables.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>6.49e-10</td>
<td>0.2364</td>
<td>-0.7145</td>
<td>1.7704</td>
</tr>
<tr>
<td>Primary</td>
<td>3.67e-09</td>
<td>0.2179</td>
<td>-0.6597</td>
<td>0.9581</td>
</tr>
<tr>
<td>Consumer</td>
<td>-5.25e-10</td>
<td>0.2503</td>
<td>-0.6586</td>
<td>1.2620</td>
</tr>
<tr>
<td>Services</td>
<td>-1.45e-09</td>
<td>0.2455</td>
<td>-0.7145</td>
<td>1.7705</td>
</tr>
</tbody>
</table>

Table 2 displays the summary statistics on the imputed abnormal returns for my sample in 2011. The table shows that the mean abnormal returns in the primary sector are positive, and those in the other two sectors are negative. Firms in the consumer sector have the lowest mean of abnormal returns. The dispersion of the abnormal returns is highest in the consumer sector and lowest in the primary sector. This reflects that firms in the consumer sector are exposed to the highest risk relative to the other sectors.
3.3 Firm Characteristics

Firm characteristics affect firms’ risk, the nature of their managers’ responsibilities, and the satisfaction he/she derives from managing the firm. These characteristics are also relevant to the nonpecuniary benefits of managerial effort. Table 3 shows the summary statistics on firms’ characteristics by sectors. They include the firms’ market values, assets, sales, employees and debt-equity ratios. The average market value and asset of the firms are USD $10.58 and $20.94 billion, respectively. Firms in the primary sector have the lowest market value, and those in consumer sector have the highest assets. The average sales for all firms are around USD $9.85 billion, and firms in the consumer sector have the highest sales. Firms in the consumer sector are also relatively more labor-intensive. These statistics provide some ideas about the scope of managerial responsibilities. The table also shows the summary statistics on the debt-equity ratios, which reflect firms’ risk to some extent. Firms in the consumer sector have the highest leverage level, and those in the primary sector have the lowest leverage level.

Table 3: Summary Statistics on Firms’ Characteristics

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Primary</th>
<th>Consumer</th>
<th>Services</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Value</td>
<td>11.13</td>
<td>10.97</td>
<td>9.55</td>
<td>10.58</td>
</tr>
<tr>
<td></td>
<td>(27.39)</td>
<td>(24.33)</td>
<td>(24.70)</td>
<td>(26.01)</td>
</tr>
<tr>
<td>Assets</td>
<td>14.29</td>
<td>36.53</td>
<td>23.09</td>
<td>20.94</td>
</tr>
<tr>
<td></td>
<td>(40.26)</td>
<td>(194.39)</td>
<td>(132.90)</td>
<td>(114.12)</td>
</tr>
<tr>
<td>Sales</td>
<td>10.98</td>
<td>13.28</td>
<td>6.42</td>
<td>9.85</td>
</tr>
<tr>
<td></td>
<td>(28.39)</td>
<td>(40.82)</td>
<td>(16.96)</td>
<td>(27.96)</td>
</tr>
<tr>
<td>Employee</td>
<td>27.37</td>
<td>46.80</td>
<td>14.95</td>
<td>26.48</td>
</tr>
<tr>
<td></td>
<td>(51.36)</td>
<td>(181.80)</td>
<td>(38.66)</td>
<td>(86.16)</td>
</tr>
<tr>
<td>Debt-Equity Ratio</td>
<td>2.18</td>
<td>2.42</td>
<td>2.41</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>(3.68)</td>
<td>(3.05)</td>
<td>(2.85)</td>
<td>(3.31)</td>
</tr>
</tbody>
</table>

The prerequisite for this study is that CEO compensation increases with firms’ size and abnormal returns. Table 4 shows the sample correlations between CEO compensation, firm characteristics, and firms’ abnormal return. The correlation between firms’ market value
and CEO compensation is 0.3630, which is the largest among all characteristics of firms. It suggests that firms’ market value has the most explanatory power on CEO compensation, and this justifies my use of market value as the basis for estimating firms’ size. Table 4 also shows that the correlation between CEO compensation and firms’ abnormal return is very high. This is consistent with my model prediction that CEO compensation increases with firms’ performance as measured by the abnormal returns of firms.

<table>
<thead>
<tr>
<th>Table 4: VARIABLE CORRELATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEO compensation 1.0000</td>
</tr>
<tr>
<td>Market value 0.3630 1.0000</td>
</tr>
<tr>
<td>Assets 0.0635 0.4092 1.0000</td>
</tr>
<tr>
<td>Sales 0.2648 0.7908 0.3905 1.0000</td>
</tr>
<tr>
<td>Employee 0.1997 0.5013 0.2574 0.6737 1.0000</td>
</tr>
<tr>
<td>Abnormal returns 0.4324 0.0748 -0.0777 -0.0110 -0.0317 1.0000</td>
</tr>
</tbody>
</table>

4 Identification and Estimation

4.1 Identification

I estimate my model from cross-sectional data on firms’ financial returns, CEO compensation, and CEO and firm characteristics. The measure of abnormal returns is constructed by regressing the difference between the financial returns of the individual firm’s stock and the returns of the market portfolio on constant and sector-specific dummy variables. The measure of compensation is assumed to have a measurement error, \( \varepsilon_i = \tilde{w}_i - w_i \), which is orthogonal to the other variables of interests with zero mean. In our empirical application, I will include the firm’s information on industry, economic conditions, and characteristics to account for the heterogeneity across firms.

In the model, in principle, firms can choose to accept the contract of implementing their matched managers to work or shirk, depending on which one to maximize their expected
profits. However, in the data, I observe that all managers are paid on firms’ performance and we also know that managers would be paid a flat wage. Therefore, I assume that all firms accept a contract that requires managers to work diligently. The probability density function of abnormal returns under working, $f_2(x)$, is then nonparametrically identified from abnormal returns alone for the set of managers whose compensation varies with abnormal returns (in our sample, all of them). Similarly, the optimal compensation contract, $w(x)$, for diligent work is identified from compensation $w$ of all managers and their consistently estimated abnormal return $x$.

After identifying $f_2(x)$ and $w(x)$, we proceed as they are known for identifying other parameters in the moral hazard model: the risk-aversion parameter $\rho$; nonpecuniary benefits for shirking versus working, $\alpha_1/\alpha_2$; nonpecuniary benefits for outside options versus working, $\alpha_0/\alpha_2$; and signaling function $g(x)$. Appendix 2 shows a detailed result for identifying these parameters for the moral hazard model with CRRA utility function. Similar results for identifying a more generalized hybrid model with CARA utility function can be founded in Gayle et al. (2015).

The first step is to prove that given the risk-aversion parameter, $\rho$, is known, the remaining parameters are nonparametrically identified. First, $g(x)$ is nonparametrically identified by tracing out the slope of the optimal contract as a function of abnormal returns $x$ and by using a boundary condition, $g(\infty) = 0$. Recalling the definition of the signal function, $g(x) \equiv f_1(x)/f_2(x)$, I can then identify $f_1(x)$ using $f_1(x) = f_2(x)g(x)$. Given $f_1(x)$ and $\rho$, the two relative nonpecuniary benefits, $\alpha_0/\alpha_2$ and $\alpha_1/\alpha_2$, are identified by using the participation and incentive compatibility constraints, (4) and (5), respectively, with equality in equilibrium.\(^{17}\)

The risk-aversion parameter $\rho$ is identified by exploiting the participation constraints (4). To do so, I follow Gayle and Miller (2009) and identify it by assuming that there are at least two states, that is, two sectors, where the nonpecuniary benefits and outside options are the

\(^{17}\)See Appendix 2 for details.
same. Let $w^\kappa(x)$ and $w^\tau(x)$ denote managers’ compensation for the two sectors, and $f^\kappa_2(x)$ and $f^\tau_2(x)$ denote the probability density functions of abnormal returns under working in the two states. Managers’ participation constraints are equal in both datasets, giving,

$$\int \frac{[w^\kappa(x)]^{1-\rho}}{1-\rho} f^\kappa_2(x) dx = \int \frac{[w^\tau(x)]^{1-\rho}}{1-\rho} f^\tau_2(x) dx$$  \hspace{1cm} (16)$$

If there is at least one solution satisfying this equation (16), $\rho$ is then identified. In this study, I divide the sample into three sectors with similar firms and managers to identify $\rho$.

After identifying the above parameters in the moral hazard model, I finally identify the distributions of firms’ size and managers’ talent. The basic idea is derived from literature on labor that states that the productive types distributions can be recovered from the income distributions observed in the competitive market (see e.g., Sattinger (1979, 1993); Terviö (2008)). By using it, I identify the quantile functions of firms’ size and managers’ talent, $S(\cdot)$ and $T(\cdot)$, from the assignment conditions in equilibrium, which provide us a one-to-one mapping between the distributions of firms’ effective size and managers’ talent, and the distributions of firms’ profits and managerial compensation.

By using the slope of the effective incomes of firms and managers given in the proof of Proposition 2, $v'(r) = \tilde{S}(r)T'(r)$ and $\pi'(r) = \tilde{S}'(r)T(r)$, and the fact that the expected effective output is divided by managers and firms, $\tilde{S}(r)T(r) = \pi(r) + v(r)$, I obtain the relative quantile functions of firms’ effective size and managers’ talent:\footnote{Derivation is available upon request.}

$$\frac{\tilde{S}(r)}{S(0)} = \exp \left[ \int_0^r \frac{\pi'(i)}{v(i) + \pi(i)} di \right],$$

$$\frac{T(r)}{T(0)} = \exp \left[ \int_0^r \frac{v'(i)}{v(i) + \pi(i)} di \right].$$  \hspace{1cm} (17) and (18)$$

Equations (17) and (18) show that the relative quantile functions, $\tilde{S}(r)/S(0)$ and $T(r)/T(0)$, are identified from the effective profits and effective compensations: $\pi(r)$ and $v(r)$ for all $r \in [0, 1]$ are known. $v(r)$ is obtained by ordering the relative nonpecuniary benefits of
outside options over working, \( v(r) = \alpha_0/\alpha_2 \). \( \pi(r) \) is recovered by using \( \pi(r) = e^{-\chi_n} \Pi_n \), where \( \Pi_n \) is the expected profits of firm \( n \), and \( e^{-\chi_n} \) is obtained by using (9) from the identified parameters. With \( \tilde{S}(r)/\tilde{S}(0) \) in hand, the relative quantile function of firms’ actual size \( S(n)/S(0) \) can then be identified by using the definition of firms’ effective size, \( \tilde{S}(r) \equiv S(n)e^{-\chi_n} \).

4.2 Estimation

The probability density function of abnormal returns under working \( f_{n2}(x) \) is first estimated by using the constructed firms’ abnormal returns. I follow Gayle and Miller (2009) to adopt a parametric approach for estimating \( f_{n2}(x) \). I assume that the probability density functions of abnormal returns \( x \) under shirking \( (e = 1) \) and working \( (e = 2) \) are both truncated normal with support bounded below by \( \psi \),

\[
 f_{ne}(x) = \left[ \Phi \left( \frac{\mu_{ne} - \psi}{\sigma_n} \right) \sigma_n \sqrt{2\pi} \right]^{-1} \exp \left[ -\frac{[x - \mu_{ne}]^2}{2\sigma_n^2} \right],
\]

where \( \Phi(\cdot) \) is the standard normal distribution function and \( (\mu_{ne}, \sigma_n^2) \) denote the mean and variance, respectively, of the corresponding parent normal distributions for firm \( n \). In this specification, the probability density functions of abnormal returns for all firms share the same functional form; however, their means and variances are different. The probability density functions under shirking and working have different means but the same variance, implying that managerial effort does not affect the risk of firms in my static model.

The model restricts the expected abnormal returns conditional on working to be zero for all firms, \( E_n[x|e = 2] = 0 \). Moreover, my data support the fact that mean of abnormal returns is zero. Thus, I use this restriction in the estimation of \( f_{n2}(x) \). By using the truncated normal specification of \( f_{n2}(x) \), the implicit function for \( \mu_{n2} \) is then given by

\[
 0 = E_n[x|e = 2] = \mu_{n2} + \frac{\sigma_n \varphi \left( [\mu_{n2} - \psi]/\sigma_n \right)}{\Phi \left( [\mu_{n2} - \psi]/\sigma_n \right)},
\]

26
where $\varphi(\cdot)$ denotes the standard normal probability density function.

To introduce firms’ heterogeneity, I specify the mean and variance of the probability density functions, $(\mu_{n2}, \sigma^2_{n2})$, respectively, as functions of observed firms’ characteristics, including number of employees, debt-equity ratio, and sector dummy variables. I denote the firms’ covariates by $z_{n1}$, and I follow Gayle and Miller (2009) to specify variance $\sigma^2_{n2}$ as an exponential function:

$$
\sigma^2_{n2} = \exp[\beta'z_{n1}],
$$

where $\beta$ is a vector of parameters for the observed firms’ covariates $z_{n1}$. $\mu_{n2}$ will also be a function of $\beta$, defined by (20). The estimation of $f_{n2}(x)$ is completed by estimating $(\psi, \beta)$ using a Maximum Likelihood Estimator (MLE) through the probability density function of abnormal returns under working subject to the restriction that the expected value of abnormal returns is zero when managers work (20).

Second, I estimate the probability density functions of abnormal returns under shirking, $f_{n1}(x)$, the two relative nonpecuniary benefits, $\alpha_{m0}/\alpha_{n2}$ and $\alpha_{n1}/\alpha_{n2}$, and the risk-aversion parameter, $\rho$. Under the truncated normal distribution specification, the parameters that characterize $f_{n1}(x)$ are the mean $\mu_{n1}$ and variance $\sigma^2_{n1}$. $f_{n1}(x)$ is assumed to share the same variance as $f_{n2}(x)$. Thus, the estimation of $f_{n1}(x)$ is then completed by estimating its mean, $\mu_{n1}$. To capture firms’ heterogeneity, I follow Gayle and Miller (2009) to specify the mean $\mu_{n1}$ as a linear function of the observed firms’ covariates, $\mu_{n1} = u'_1z_{n1}$, where I use the same observed firms’ covariates as that in the specification of variance $\sigma^2_{n2}$ because $\mu_{n2}$ is an implicit function of the covariates.

The relative nonpecuniary benefit of managerial effort from outside options versus working, $\alpha_{m0}/\alpha_{n2}$, is determined by the demand for management service in the labor market and by managers’ satisfaction from working for firms. Thus, $\alpha_{m0}/\alpha_{n2}$ would ideally be specified as a function of both firms’ and managers’ characteristics. However, I specify $\alpha_{m0}/\alpha_{n2}$ only
as a function of firms’ characteristics because I do not have information on managers’ characteristics, and heterogeneous firms are endogenously matched with heterogeneous managers and the characteristics of firms and managers are sorted to some extent in equilibrium. In particular, I specify \( \alpha_{m0}/\alpha_{n2} \) as a linear function of firms’ characteristics, \( \alpha_{m0}/\alpha_{n2} = a'_0 z_{n2} \), where \( z_{n2} \) is a vector of firms’ characteristics, including firms’ assets and number of employees. The relative nonpecuniary benefit of managerial effort from shirking versus working, \( \alpha_{n1}/\alpha_{n2} \), reflects the costs that firms bring to their hired managers. I also specify it as a linear function of firms’ characteristics, \( \alpha_{n1}/\alpha_{n2} = a'_1 z_{n2} \). To sum up, in this step, the parameters that need be estimated are \( (\mu_1, \alpha_0, \alpha_1, \rho) \). I employ a Minimal Distance Estimator (MDE) constructed from the difference between the observed compensation and model predicted optimal contracting compensation.

I finally use equations (17) and (18) to estimate the relative quantile functions of firms’ effective size, managers’ talent, and actual size, \( \tilde{S}(\cdot)/\tilde{S}(0) \), and \( S(\cdot)/S(0) \), respectively, using the definition \( \tilde{S}(r) \equiv S(n)e^{-\chi_n} \). To apply our assignment model to data, three complications must be dealt with. First, economic value created as a result of the interactions between firms and their current CEOs is not directly observed. The economic value is also affected by past CEOs and the expectation of future CEOs. Second, the observed firms’ market values may not reflect the expected profits obtained by firms. Third, the managers’ effective compensations and firms’ effective profits estimated based on observed noisy managerial compensation and firms’ market value may not be perfectly rank-correlated.

I follow Terviö (2008) to deal with these complications. The first two complicated are dealt with by explicitly modeling the impacts of past and future CEOs on the economic value and by also modeling the optimal choice of adjustable capital by firms. The noisy relation of managerial effective compensations and firms’ effective profits is smoothed into a strictly monotonic relation. I perform a Lowess Smoothing of the relation of the levels of managerial effective compensations and firms’ effective profits using the method first proposed by Cleveland (1979). Because the rank of firms’ effective profits is used to order the
observations, there is no need to smooth it; I only need to perform a simple connect-the-dots interpolation to create a continuum. Hereafter, I use the smoothed effective compensation to refer to the actual effective compensations.

After dealing with these complications, equations (17) and (18) can be rewritten as

\begin{align}
\tilde{S}(r) &= \exp \left[ (1 - B) \int_0^r \frac{\tilde{\pi}'(i) - v'(i)/\gamma}{[1 - B] \tilde{\pi}(i) + v(i)} \, di \right], \\
T(r) &= \exp \left[ \frac{\gamma + 1 - B}{\gamma} \int_0^r \frac{v'(i)}{[1 - B] \tilde{\pi}(i) + v(i)} \, di \right],
\end{align}

where \( B \equiv [1 + g]/[1 + r] \) is the growth-adjusted discount factor, with \( g \) representing the growth rate and \( r \), the discount factor; and \( \gamma \) is the fading rate of CEOs’ impact on firms’ economic value. The effective compensations of managers \( v(r) \) is estimated by \( \hat{v}(r) = \hat{a}_0 z_n^2 \).

The effective profits of firms \( \tilde{\pi}(r) \) can be estimated by \( \hat{\tilde{\pi}}(r) = e^{-\hat{\chi}_n} \Pi_n \). \( \Pi_n \) is the expected profits of firms, and it is estimated by

\begin{equation}
\hat{\Pi}_n = \xi \Pi'_n - \frac{[1 - \xi] \hat{v}(r)}{[1 - B] e^{-\hat{\chi}_n}},
\end{equation}

where \( \xi = \frac{1 - \theta + \frac{1 - \theta}{[1 - B]} e^{-\hat{\chi}_n}}{1 - \theta + \frac{1 - \theta}{[1 - B]} e^{-\hat{\chi}_n}} \), with \( \theta \) denoting the share of adjustable capital in production, and \( \Pi'_n \) is the observed market value and CEO compensation of firm \( n \). I will use equations (21) and (22) to estimate the relative quantile functions of firms’ effective size and managers’ talent.

5 Results

I now present the main results on the importance of moral hazard in the assignment of managers to firms, and the estimates of model parameters that drives them. First, I report the

\footnote{The basic idea of this method is to take a weighted moving average of effective compensation along the rank by firms’ effective profits, using larger weights for nearby observations. In principle, I could also smooth firms’ effective profits according to the rank of effective compensations. However, since managerial compensations is more volatile, firms’ market value tends to be more stable.}

\footnote{The derivations of these equations are available upon request.
estimates of the probability distribution of abnormal returns, both under working, \( f_{n2}(x) \), and shirking, \( f_{n1}(x) \). This gives an estimator of their ratio, \( g_n(x) = f_{n1}(x)/f_{n2}(x) \). The estimates of these distributions convey information on how the risk premium paid to managers depends on firms’ characteristics through their risk. Second, I report the estimates of the nonpecuniary benefits of outside options versus working, \( \alpha_{m0}/\alpha_{n2} \), shirking versus working, \( \alpha_{n1}/\alpha_{n2} \), and risk-aversion parameter, \( \rho \). \( \alpha_{m0}/\alpha_{n2} \) provides information on the certainty equivalent for each firm’s manager, and \( \alpha_{n1}/\alpha_{n2} \) provides information on how the risk premium depends on the nonpecuniary benefits for each firm’s manager. The estimates of the risk-aversion parameter reflect the importance of firms’ risk in determining the risk premium paid to managers. Third, I report the estimates of the distributions of firms’ size and managers’ talent. These distributions help us to understand how the variations of executive compensation depend on firms’ size and managers’ talent.

By using the estimates of the model parameters, I conduct counterfactuals to investigate how moral hazard affects the assignment of managers to firms in equilibrium and evaluate its importance. Specifically, I quantify three measures of the efficiency loss due to moral hazard: i) loss arising from the risk-sharing inefficiency, (ii) loss arising from the inefficient equilibrium assignment of managers, and (iii) loss arising from a random assignment of managers. The definitions of these measures will be given later.

5.1 Abnormal Returns from Working and Shirking

The probability density function of abnormal returns under working, \( f_{n2}(x) \), is estimated in the first step. Recall that the estimation of \( f_{n2}(x) \) is completed by estimating the variance of its parent distribution function \( \sigma_n \), lower bound \( \psi \), and mean of its parent distribution function. The first column of Table 5 shows the estimates for the variables in the specification of the variance \( \sigma_n \). The estimates provide information on how firms’ risk depends on their different observed firms’ characteristics. The estimate of the debt-equity ratio is positive, suggesting that the more leveraged firms are riskier, holding all other factors constant. The
number of employees has a negative effect on firms’ risk, implying that firms with more employees are less risky, holding all other factors constant. Firms in the consumer sector tend to have the highest risk, and those firms in the primary sector have the lowest risk. This is consistent with the dispersion of abnormal returns, as shown in Table 2. The estimate of the truncation lower bound $\psi$ is $-0.7145$, which is the lowest abnormal return in the data. The mean for the parent distribution of $f_{n2}(x)$ is estimated by imposing the restriction that the expected abnormal returns are zeros under manager working, $E_n(x|e = 2) = 0$.

Table 5: PROBABILITY DENSITY FUNCTIONS OF ABNORMAL RETURNS $x$

(Standard errors are in the parentheses.)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Variance $\sigma_n^2$</th>
<th>Mean $\mu_{n1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.7267</td>
<td>-0.1365</td>
</tr>
<tr>
<td></td>
<td>(0.2593)</td>
<td>(0.0964)</td>
</tr>
<tr>
<td>Debt-equity ratio</td>
<td>0.0071</td>
<td>-0.0030</td>
</tr>
<tr>
<td></td>
<td>(0.0136)</td>
<td>(0.0024)</td>
</tr>
<tr>
<td>Log No. employees</td>
<td>-0.1523</td>
<td>-0.0136</td>
</tr>
<tr>
<td></td>
<td>(0.0284)</td>
<td>(0.0059)</td>
</tr>
<tr>
<td>Consumer</td>
<td>0.4261</td>
<td>0.0114</td>
</tr>
<tr>
<td></td>
<td>(0.1314)</td>
<td>(0.0303)</td>
</tr>
<tr>
<td>Service</td>
<td>0.1690</td>
<td>-0.0026</td>
</tr>
<tr>
<td></td>
<td>(0.1038)</td>
<td>(0.0290)</td>
</tr>
<tr>
<td>$\psi$</td>
<td></td>
<td>-0.7145</td>
</tr>
</tbody>
</table>

The probability density function under shirking, $f_{n1}(x)$, has the same variance as that for $f_{n2}(x)$. Only its mean $\mu_{n1}$ then needs to be estimated; It is estimated in the second step. The estimates for the parameters in the specification of $\mu_{n1}$ are reported in the second column of Table 5. The mean decreases with firms’ debt-equity ratio and number of employees. This result suggests that the average abnormal returns decrease more for firms with higher debt-equity ratio and more employees if their managers switch from working to shirking. It suggests that managers have higher impacts at firms with higher debt-equity ratio and more employees. The estimate of sector dummies is positive for firms in the consumer sector and negative for firms in the service sector, indicating that managers have larger impacts in firms.
of the service sector.

Table 6: PARAMETERS IN UTILITY FUNCTIONS

(Standard errors are in the parentheses.)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Nonpecuniary Cost</th>
<th>Non-pecuniary Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shirking</td>
<td>Working</td>
</tr>
<tr>
<td></td>
<td>$a'<em>0 z</em>{n2}$</td>
<td>$a'<em>1 z</em>{n2}$</td>
</tr>
<tr>
<td>Constant</td>
<td>5.0660</td>
<td>3.3655</td>
</tr>
<tr>
<td></td>
<td>(0.4919)</td>
<td>(2.6911)</td>
</tr>
<tr>
<td>Log employees</td>
<td>0.1012</td>
<td>-0.0789</td>
</tr>
<tr>
<td></td>
<td>(0.0320)</td>
<td>(0.1396)</td>
</tr>
<tr>
<td>Log assets</td>
<td>0.2011</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0351)</td>
<td>(0.0809)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3940)</td>
<td></td>
</tr>
</tbody>
</table>

5.2 Nonpecuniary Benefits of Managerial Efforts

Table 6 shows the estimates of the parameters in managers’ preferences. They include the parameters characterizing the relative nonpecuniary benefits of outside options versus working, $\alpha_m/\alpha_n$, shirking versus working, $\alpha_{n1}/\alpha_{n2}$, and the risk-aversion attitude of managers $\rho$. The results show that managers working for firms with more assets and employees obtain more nonpecuniary benefits from outside options versus working. This may be because such firms endogenously match with more talented managers whose reservation utilities are higher. The results also show that managers working for firms with more employees obtain less nonpecuniary benefits from shirking. Furthermore, managers working for firms with more assets obtain more nonpecuniary benefits from shirking. This may be because the management responsibilities in firms with more employees provide managers with more satisfaction; in contrast, firms having more assets impose more responsibilities and provide less satisfaction to managers.
5.3 Risk-Aversion Parameter

The risk-aversion parameter $\rho$ is estimated to be 1.18. At this level of risk-aversion, a manager with wealth of USD 8 million (mean compensations of CEOs in my sample) is willing to pay USD 76,200 to avoid a gamble in which he/she has equal probability of losing and winning USD 1 million. Gayle and Miller (2009) estimated absolute risk-aversion parameters using two data samples spanning about 60 years; the values were 0.501 and 0.519, respectively. With these two risk-aversion parameters, managers are willing to pay about USD 240,000 and USD 248,000 to avoid a gamble that has an equal probability of losing or winning USD 1 million. Because I consider CEOs with high levels of wealth, the relative risk aversion coefficient is small among the estimates in the most literature (see also e.g., Friend and Blume (1975); Garcia et al. (2003); Mankiw (1985); Weber (1975)).

5.4 Distributions of Firms’ Size and Managers’ Talent

I use equations (21) and (22) to estimate the relative quantile functions of firms’ effective size and managers’ talent, $\tilde{S}(\cdot)/\tilde{S}(0)$ and $T(\cdot)/T(0)$, and then estimate the relative quantile function of firms’ actual size, $S(\cdot)/S(0)$, by using the definition if firms’ effective size, $\tilde{S}(r) \equiv S(n)e^{-\gamma n}$. For estimating $\tilde{S}(\cdot)/\tilde{S}(0)$ and $T(\cdot)/T(0)$, the values of several parameters must be obtained: managerial talent impact fading rate $\gamma$, discount factor $r$, share of adjustable capital in production $\theta$, and growth rate $g$. Because these parameters are determined outside my structural model, I follow Terviö (2008) and try different combinations of these parameters in the estimation and counterfactuals. Figure 1 shows the estimated quantile functions of firms’ effective size and managers’ talent under the certain values of these parameters (case B in Table 7). It shows that the distribution of firms’ effective size is highly skewed to right, and there is not much difference in the managers’ talent. This

\footnote{Friend and Blume (1975) estimated that the coefficient of relative risk aversion exceeds 1 for consumers demanding for risky assets. Garcia et al. (2003) using a generalization of a Black-Scholes option pricing model to S&P 500 call option prices reported the estimates of relative risk aversion in the range of 0.83 to 3.28. Mankiw (1985) found large estimates in the range of 2.44 to 5.26 using data on consumption. Weber (1975) used expenditure data to estimate relative risk aversion in the range of 1.3 and 1.8.}
is consistent with the findings in literature that most variations on CEO compensation are explained by the differential of firms’ size instead of managers’ talent (see e.g., Terviö (2008) and Jung and Subramanian (2013)).

5.5 Loss from Risk-sharing inefficiency

The first counterfactual quantifies the loss from the risk-sharing inefficiency. It is measured by the difference between the expected compensation that a firm pays to its manager and the counterfactual flat compensation it would pay to its manager if moral hazard is absent. Let $E_n[w_n^o(x)]$ denote the expected compensation in equilibrium, and $v_n = \alpha_{m0}/\alpha_{n2}$ denote the counterfactual flat compensation for firm $n$. The risk-sharing inefficiency for firm $n$ is then

$$L_n^1 = E[w_n^o(x)] - v_n.$$  \hspace{1cm} (24)

The sum of $L_n^1$ for all firms, $\sum_n L_n^1$, gives the total loss of firms from risk-sharing inefficiency.

The first row in the middle of Table 7 shows the estimated total loss from risk-sharing inefficiency over all firms; it is around USD 588.25 million. This number is smaller than the estimates obtained in the literature. By using a different theoretical framework, Edmans and

Table 7: EFFICIENCY LOSS ASSOCIATED WITH MORAL HAZARD

<table>
<thead>
<tr>
<th>Cases</th>
<th>A</th>
<th>A1</th>
<th>A2</th>
<th>B</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount rate (r)</td>
<td>0.08</td>
<td>0.08</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Growth rate (g)</td>
<td>0.02</td>
<td>0.04</td>
<td>0.025</td>
<td>0.025</td>
<td>0.025</td>
</tr>
<tr>
<td>Share of adj. capital (θ)</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Rate of impact fading (λ)</td>
<td>∞</td>
<td>∞</td>
<td>0.1</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>Loss from CEO misallocation (mill.)</td>
<td>603.59</td>
<td>943.88</td>
<td>1291.30</td>
<td>1555.6</td>
<td>1585.2</td>
</tr>
<tr>
<td>Loss from a random CEO allocation (mill.)</td>
<td>184.91</td>
<td>277.69</td>
<td>341.57</td>
<td>416.86</td>
<td>358.92</td>
</tr>
<tr>
<td>Loss from risk-sharing inefficiency (mill.)</td>
<td>588.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total CEOs’ pay (bill.)</td>
<td>5.7216</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total market value (bill.)</td>
<td>10,583</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.6 Loss from Inefficient Equilibrium Assignment of CEOs

The second counterfactual quantifies the loss from the equilibrium assignment of CEOs to firms. I measure the loss by the difference between the expected production output of firms from the counterfactual efficient assignment and that from the inefficient equilibrium assignment of CEOs to firms. Let \( \tilde{T}(n) \) denote the talent of the CEO hired by firm \( n \) under the equilibrium allocation.\(^{22}\) Firm \( n \)'s loss from inefficient assignment in equilibrium is then

\(^{22}\)Since we know the firm \( n \) has effective size \( \hat{S}(r) \), the firm will end up with a talent \( T(r) \) manager. Thus, \( \tilde{T}(n) = T(r) \)
given by

\[
L_n^2 = E_n[S(n)T(n)[1 + x]] - E_n[\bar{S}(n)\bar{T}(n)[1 + x]]
\]

(25)

\[
= S(n)\left[T(n) - \bar{T}(n)\right],
\]

where the second equality uses the normalization \(E_n[x] = 0\). The sum of \(L_n^2\) for all firms, \(\sum_n L_n^2\), gives the gross loss that firms would incur from CEO misallocation in equilibrium.

To quantify the loss, I need to calculate the expected output under the efficient allocation and that under equilibrium allocation of CEOs to firms. They depend on the quantile functions of firms’ actual and effective sizes as well as managers’ talent. Recalling that the quantile functions depend on the values of parameters outside the model, I follow Terviö (2008) and try several combinations of these parameters. The second row in the middle of Table 7 shows that the loss ranges between USD 0.604 million and 1.585 billion for different values of the parameters. This accounts for between 10.5%–27.71% of the total CEO compensation. Furthermore, this value can be more than two times the aggregate loss from risk-sharing inefficiency. Edmans and Gabaix (2011) calibrated a different theoretical framework to obtain USD 7.7 billion as an upper bound of the loss for the top 500 firms in 2005 Execucomp databases.

5.7 Loss from Random Assignment of CEOs to Firms

I finally conduct a counterfactual to quantify the efficiency loss from a random assignment of CEOs to firms. I measure the loss by the difference between the expected production output of firms from a counterfactual random assignment and that from the efficient assignment of CEOs to firms. Let \(\bar{T}(n)\) denote the talent of the CEO hired by firm \(n\) under a random
allocation of CEOs. Firm \( n \)'s loss from a random allocation of CEOs is then given by

\[
L^3_n = E_n \left[ S(n)T(n)[1 + x] \right] - E_n \left[ S(n)\bar{T}(n)[1 + x] \right] \\
= S(n) \left[ T(n) - \bar{T}(n) \right],
\]

(26)

where the second equality uses the normalization \( E_n[x] = 0 \). The sum of \( L^3_n \) over all firms, \( \sum_n L^3_n \), gives the gross loss that all firms would incur from a random allocation of CEOs to firms.

The third row at the bottom of Table 7 shows the results for this loss for different combinations of parameters outside the model. The loss is USD 184.9-416.9 billion. This amount is significant, as it accounts for 1.72%-3.94% of firms’ total market value. When there are more managers available than the number of CEO positions at firms in the market, this loss is a more plausible estimate of the loss from CEO misallocation. This amount is much larger than my estimated loss from CEO misallocation in equilibrium where we assume that the number of managers available and firms in the market for CEOs is the same. This assumption tends to underestimate the loss from the CEOs misallocation in equilibrium. Edmans and Gabaix (2011) calibrated a different theoretical framework to obtain USD 16 billion as a lower bound of the loss from a random allocation of CEOs to firms for the top 500 firms in 2005 Execucomp databases.

The counterfactual results show that the total loss firms incur from the inefficient assignment of CEOs to firms in equilibrium can be more than two times the loss due to the risk-sharing inefficiency when moral hazard is present. Most previous studies focus on the risk-sharing inefficiency, also called agency costs. My results suggest that the studies focusing solely on risk-sharing inefficiency can severely underestimate the efficiency loss due to moral hazard. The sum of the loss from both risk-sharing inefficiency and assignment inefficiency due to moral hazard in equilibrium is the aggregate costs that all firms would

\(^{23}\)To the best of my knowledge, Edmans and Gabaix (2011) is the only exception. My empirical results are consistent with theoretical predictions and calibration results in Edmans and Gabaix (2011).
incur by using the optimal compensation contract to solve the moral hazard problem. Corporate governance is considered a substitute mechanism to reduce the cost of moral hazard. Understanding the aggregate costs of alternative ways to solve moral hazard problem can help in implementing better corporate governance.

6 Conclusion

I have proposed a structural model to illustrate that the presence of moral hazard not only leads to risk-sharing inefficiency but also leads to inefficient assignment of CEOs to firms in equilibrium. The model parameters are estimated using the data on the 1000 largest firms in the US for 2011. The estimation results showed that more leveraged firms and/or firms with more assets tend to end up with less-talented CEOs. Counterfactuals conducted from the estimates of the model parameters suggest that the loss from the inefficient assignment of CEOs in equilibrium is large, and studies ignoring this issue may be misleading.

Literature on labor usually focuses on the agency theory and testing the implication of this theory. This paper suggests that hiring the right employees in the first place is also significantly important, and ignoring this guideline may lead firms to incur large losses. However, we should be aware that the results are based on some simplifying assumptions, among which two are critical. First, I consider a static assignment model incorporating with a static contracting problem, and I ignore more realistic dynamic considerations regarding both assignment and contracting. Second, I assume that all CEOs have no existing wealth when estimating the model parameters.
Appendix

1. Proofs of Propositions

Proposition 1. The cost-minimizing compensation contract that corresponds to working is given by

\[ w_{nm}(x) = \frac{\alpha_{m0}}{\alpha_{n2}} \left[ 1 + \lambda_n \left[ \frac{\alpha_{n1}}{\alpha_{n2}} \right]^{\rho-1} - g_n(x) \right]^{\frac{1}{\rho-1}} \]

where \( \lambda_n \) is the unique positive solution to the equation,

\[ \int \frac{f_{n2}(x)}{1 + \lambda_n \left[ \frac{\alpha_{n1}}{\alpha_{n2}} \right]^{\rho-1} - g_n(x)} dx = 1. \]

Proof. To solve firm \( n \)’s cost-minimizing contract that induces its manager to work, I follow Grossman and Hart (1983) and transfer the cost-minimization problem (3) into a convex programming problem. Defining a function of compensation \( h(w(x)) \equiv \left[ \frac{\alpha_{n2}/\alpha_{m0}}{w(x)} \right]^{1-\rho} \), we have \( \log[h] = [1-\rho] \log \left[ \frac{\alpha_{n2}/\alpha_{m0}}{w(x)} \right] \). Since \( \rho > 1 \), we have \( 1-\rho < 0 \). We can then rewrite the cost-minimization problem (3) as a convex programming problem,

(27) \[ \min_h E[-\log[h]] \]

(28) \[ \text{s.t. } E[h] \leq 1, \]

(29) \[ E[h] \leq E \left[ \frac{\alpha_{n1}/\alpha_{n2}}{1-\rho} \cdot h \cdot g_n(x) \right]. \]

The Lagrange Function corresponding to the above optimization problem is given by

\[ \mathcal{L} = E[\log(h)] + \lambda_{n1} \left[ 1 - E[h] \right] + \lambda_{n2} \left[ E \left[ \frac{\alpha_{n1}/\alpha_{n2}}{1-\rho} \cdot h \cdot g_n(x) \right] - E[h] \right] \]

39
The point-wise first order condition with respect to $h$ is derived as

$$h^{-1} = \lambda_{n1} + \lambda_{n2} \left[1 - \left[\alpha_{n1}/\alpha_{n2}\right]^{1-\rho}g_n(x)\right].$$ \hspace{1cm} (30)

I now show that the participation and incentive compatibility constraints, (28) and (29), are binding, which implies that $\lambda_{n1}$ and $\lambda_{n2}$ are both positive. First, I show that $\lambda_{n2} > 0$ by contradiction. Suppose $\lambda_{n2} = 0$, we have $h^{-1} = \lambda_{n1}$, which implies that $w(x) = \lambda_{n1}^{\frac{1}{\rho}}\alpha_0$. This gives that the compensation is a flat constant. This contradicts with the fact $w(x)$ should depend on $x$. Thus $\lambda_{n2} > 0$. Second, I show that $\lambda_{n1} = 1$. Multiplying the first order condition (30) by $h$ and taking expectation on both sides yields

$$1 = \lambda_{n1}E[h] + \lambda_{n2} \left[E[h] - E \left[\left[\alpha_{n1}/\alpha_{n2}\right]^{1-\rho} \cdot h \cdot g_n(x)\right]\right].$$ \hspace{1cm} (31)

$\lambda_{n2} > 0$ gives incentive compatibility constraint binding, $E[h] - E \left[\left[\alpha_{n1}/\alpha_{n2}\right]^{1-\rho} \cdot h \cdot g(x)\right] = 0$, which implies that the equation (31) reduces to $\lambda_{n1}E[h] = 1$. Thus $\lambda_{n1}$ has to be positive, $\lambda_{n1} > 0$, which gives that participation constraint binding $E[h] = 1$. Substituting it into $\lambda_{n1}E[h] = 1$ gives $\lambda_{n1} = 1$.

Substituting $\lambda_{n1} = 1$ into the first-order condition (30) yields

$$h^{-1} = 1 + \lambda_{n2} \left[1 - \left[\alpha_{n1}/\alpha_{n2}\right]^{1-\rho}g_n(x)\right] = 1 + \lambda_n \left[\left[\alpha_{n1}/\alpha_{n2}\right]^{\rho-1} - g_n(x)\right]$$ \hspace{1cm} (32)

where $\lambda_n \equiv \lambda_{n2}[\alpha_{n1}/\alpha_{n2}]^{1-\rho}$. Substituting the definition of $h$ into it and simplifying it gives the optimal compensation (6) in Proposition 1. Equation (32) and the binding participation constraint give (7) in Proposition 1, determining $\lambda_n$. Since the convex programming problem has a unique solution, we know that there exists unique $h$, $\lambda_{n1}$, and $\lambda_{n2}$ for the convex programming problem (27), which implies that $w_n^q(x)$ and $\lambda_n$ is unique. \hfill \Box

**Proposition 2.** When moral hazard is present, the equilibrium allocation of managers to firms exhibits positive assortative sorting by firms’ effective size $\tilde{S}(r)$ and managerial talent.
\( T(r) \). The profiles of managers' effective compensations and firms' effective profits supporting the equilibrium are given by,

\[
v(r) = v^0 + \int_0^r \tilde{S}(j)T'(j) dj,
\]

and

\[
\pi(r) = \pi^0 + \int_0^r \tilde{S}'(j)T(j) dj.
\]

Proof. The first part of the proposition is shown in the discussion of the main text, which states that when moral hazard is present, firms act according to their effective sizes. The equilibrium allocation of managers to firms would exhibit positive assortative sorting by firms' effective size \( \tilde{S}(r) \) and managerial talent \( T(r) \). This follows directly from effective output is complementary between firms' effective size and managerial talent.

To derive the profiles of effective incomes of firms and managers supporting the equilibrium allocation, I first replace \( m \) with \( r - \epsilon \) in the conditions (11) and divide both sides by \( \epsilon \), yielding

\[
\frac{\tilde{S}(r)T(r) - \tilde{S}(r)T(r - \epsilon)}{\epsilon} \geq \frac{v(r) - v(r - \epsilon)}{\epsilon},
\]

which becomes equality as \( \epsilon \to 0 \). By the definition of derivation, the slope of the managerial effective compensation profile is given by \(^{24}\)

\[
(34) \quad v'(r) = \tilde{S}(r)T'(r).
\]

\(^{24}\)This equation can also be obtained by solving firm’s profit maximization problem. Firm \( n \) chooses a manager to maximize its expected profits and thus it solves the maximization problem,

\[
(33) \quad \max_m \tilde{S}(r)T(m) - v(m).
\]

Taking first derivative with respect to \( m \) and using the positive assortative sorting condition in equilibrium \( (m = r) \) gives us (34).
The managerial effective compensation profile can then be obtained by integrating the slope and using the boundary condition, \( v(0) = v^0 \). I obtain the managerial effective compensation as

\[
v(r) = v^0 + \int_0^r \tilde{S}(j)T'(j)dj.
\]

Recall the effective output \( \tilde{S}(r)T(r) = \pi(r) + v(r) \). Totally differentiating it with respect to \( r \) gives \( \tilde{S}'(r)T(r) + \tilde{S}(r)T'(r) = \pi'(r) + v'(r) \). Substituting (34) into it gives us the slope of firms’ effective profits profile,

\[
(35) \quad \pi'(r) = \tilde{S}'(r)T(r),
\]

which gives the effective profits of firms,

\[
\pi(r) = \pi^0 + \int_0^r \tilde{S}'(j)T(j)dj,
\]

by integrating (35) over \( r \) and using the boundary condition, \( \pi(0) = \pi^0 \). \( \square \)

2. Identification of \( f_1(x), \alpha_0/\alpha_2 \) and \( \alpha_1/\alpha_2 \)

The second step identifies the probability density function of abnormal returns \( x \) under shirking, \( f_1(x) \), and the two relative nonpecuniary benefits of managerial effort, \( \alpha_0/\alpha_2 \) and \( \alpha_1/\alpha_2 \). They are identified by using the optimal compensation equation (6), the participation and incentive compatibility constraints (4) and (5) with equality, given the risk-aversion parameter \( \rho \) is known. Following the identification results in Gayle et al. (2015), under Assumption 3 and the fact that \( E[g(x)] = 1 \), I derive \( g(x) \), \( \alpha_0/\alpha_2 \), and \( \alpha_1/\alpha_2 \) as functions of
$f_2(x), w^o(x), \rho$:

\begin{align*}
g(x) &= \frac{\bar{w}^{\rho-1} - w^o(x)^{\rho}}{\bar{w}^{\rho-1} - E[w^o(x)]^{\rho-1}}, \\
\alpha_0/\alpha_2 &= \left[ E[w^o(x)]^{1-\rho} \right]^{\frac{1}{\rho-1}}, \\
\alpha_1/\alpha_2 &= \left( \frac{\bar{w}^{\rho-1} - 1/E[w^o(x)]^{\rho-1}}{\bar{w}^{\rho-1} - E[w^o(x)]^{\rho-1}} \right)^{\frac{1}{\rho-1}},
\end{align*}

(36) (37) (38)

where $\bar{w}$ is the maximum compensation that all managers can receive. Equations (36)-(38) show that $g(x)$, $\alpha_0/\alpha_2$, and $\alpha_1/\alpha_2$ can be identified from $\bar{w}$, $f_2(x)$, $w^o(x)$ and $\rho$, which are identified or known by assumption. More specifically, $\bar{w}$ is identified from the maximum compensation observed in the sample. $f_2(x)$ and $w^o(x)$ are identified in step one and $\rho$ is known by assumption. Once $g(x)$ is known, the probability density function of abnormal returns under shirking is then identified from the definition, $g(x) \equiv f_1(x)/f_2(x)$.

The equations (36)-(38) can be derived by exploiting the optimal compensation contract (6), participation and incentive compatibility constraints (4) and (5) holding with equality, the boundary Assumption 3, and the fact that $E[g(x)] = 1$ as follows. First, equation (36), determining $g(x)$, is derived from the optimal compensation contract (6), Assumption 3, and the fact $E[g(x)] = 1$. Recall the optimal compensation equation (6) after suppressing firm index,

\begin{equation}
w^o(x) = \frac{\alpha_0}{\alpha_2} \left[ 1 + \lambda \left[ \frac{\alpha_1}{\alpha_2} \right]^{\rho-1} - g(x) \right]^{\frac{1}{\rho-1}}.
\end{equation}

(39)

Defining $\eta(x) \equiv [\alpha_2/\alpha_0]w^o(x)]^{\rho-1}$, equation (39) can be then rewritten as

\begin{equation}
\eta(x) = \left[ 1 + \lambda \left[ \frac{\alpha_1}{\alpha_2} \right]^{\rho-1} - g(x) \right].
\end{equation}

(40)
Assumption 3 tells that \( \lim_{x \to \infty} g(x) = 0 \). Taking limit on both sides of (40) yields

\[
\bar{\eta} \equiv \lim_{x \to \infty} \eta(x) = 1 + \lambda \left[ \frac{\alpha_1}{\alpha_2} \right]^{\rho-1}.
\]

Note the fact that \( E[g(x)] = 0 \). Taking expectation on both sides of (40) gives

\[
\bar{\eta} \equiv E[\eta(x)] = 1 + \lambda \left[ \frac{\alpha_1}{\alpha_2} \right]^{\rho-1} - 1.
\]

Subtracting both sides of equation (41) to equation (40) gives us

\[
\bar{\eta} - \eta(x) = \lambda g(x).
\]

Similarly subtracting both sides of equation (41) to equation (42) gives us

\[
\bar{\eta} - \bar{\eta} = \lambda.
\]

Dividing equation (43) over equation (44) yields

\[
g(x) = \frac{\bar{\eta} - \eta(x)}{\bar{\eta} - \bar{\eta}}.
\]

On the other hand, recall the definition \( \eta(x) \equiv \left[ \frac{\alpha_2 w^o(x)}{\alpha_0} \right]^{\rho-1} \). Taking the limit of it gives that

\[
\bar{\eta} \equiv \lim_{x \to \infty} \eta(x) = \left[ \frac{\alpha_2}{\alpha_0} \left[ \lim_{x \to \infty} w^o(x) \right] \right]^{\rho-1} \equiv \left[ \frac{\alpha_2}{\alpha_0} \bar{w} \right]^{\rho-1},
\]

where \( \bar{w} \) is the maximum compensation that managers can receive. Taking the expectation
of \( \eta(x) \) yields

\[
\eta = E[\eta(x)] = E \left[ \left( \frac{\alpha_2}{\alpha_0} w_0(x) \right)^{\rho-1} \right].
\] (47)

Substituting \( \eta(x) \) in definition, \( \bar{\eta} \) given by equation (46), and \( \bar{\eta} \) given by equation (47), into equation (45) gives equation (36) determining \( g(x) \).

The equation (37), determining the relative nonpecuniary benefits of outside options versus working \( \alpha_0/\alpha_2 \), can be derived directly from the participation constraint (4) holding with equality. Rearranging it gives (37). To derive (38), determining the relative nonpecuniary benefits of managerial effort \( \alpha_1/\alpha_2 \), rearranging (41) gives that

\[
\alpha_1/\alpha_2 = \left[ \left( \frac{\bar{\eta} - 1}{\lambda} \right)^{\frac{1}{1-\rho}} \right] = \left[ \frac{\bar{\eta} - 1}{\eta - \bar{\eta}} \right]^{\frac{1}{1-\rho}}.
\] (48)

Substituting \( \bar{\eta} \) and \( \eta \) into it gives

\[
\alpha_1/\alpha_2 = \left[ \frac{\bar{w}^{\rho-1} - \left[ \alpha_0/\alpha_2 \right]^{\rho-1}}{w^{\rho-1} - E \left[ w_0(x) \right]^{\rho-1}} \right]^{\frac{1}{\rho-1}}.
\] (49)

Finally I substitute \( \alpha_0/\alpha_2 \) given by (37) into it gives us (38).\(^{25}\)

\(^{25}\)Please refer to Gayle et al. (2015) for the intuition on the identification in this step.
References


