Managerial Talent Misallocation and the Cost of Moral Hazard

Liang Chen*

University of Calgary

October 16, 2013

Job Market Paper

Abstract

This paper identifies the sources and the magnitude of inefficiency in the allocation of CEOs to firms in the market for talent. I develop an estimable model which illustrates that the presence of moral hazard not only leads to inefficiency caused by risk sharing across the two parties, but also creates inefficiency due to a talent misallocation. The talent misallocation is novel to the literature. A new empirical method is proposed to identify the separate surplus of both firms and CEOs in a matching market with moral hazard. An application of this method to the U.S market for CEOs shows that the aggregate efficiency loss due to talent misallocation is $12.64 billion. This is more than *four times* as large as the loss stemming from risk-sharing between firms and CEOs. The findings suggest that studies focusing solely on risk-sharing can severely underestimate efficiency losses due to moral hazard.

Key Words: Talent Misallocation, Moral Hazard, Allocation Inefficiency, Nonparametric Identification, the Market for CEOs

JEL Codes: J21 J30 J31 L11

---

*E-mail: *liachen@ucalgary.ca.* I am very grateful to Eugene Choo (advisor), Arvind Magesan and Joanne Roberts for their invaluable support in the whole process of writing this paper. I am indebted to Alexander David, Curtis Eaton, Yao Luo, Atsuko Tanaka, Trevor Tombe, Kunio Tsuyuhara and Matthew Webb for extensive discussions. I thank John Boyce, Maksim Isakin and Victor Yang Song for helpful comments. All remaining errors are my own.

1 Introduction

The inefficient allocation of CEOs (also referred to as managers hereafter) to firms can cause a large amount of economic loss. Evidence shows that managers and firms differ in their productive ability, and they are complementary in production.[1] In these circumstances, efficiency requires positive assortative sorting by the types of managers and firms. In the absence of information asymmetries, the competitive equilibrium of the market for managers exhibits this sorting.[2] However, if the output of firms is stochastic and depends on unobserved managerial action, a hidden action moral hazard problem arises and efficient sorting may not be achieved in equilibrium. While this moral hazard problem has been well studied, with few exceptions the literature has focused on how to motivate managers. The way in which moral hazard affects the equilibrium allocation remains unclear. More importantly, little is known about the magnitude of inefficiency in the equilibrium allocation of managers to firms in the presence of moral hazard.

This paper has two main contributions. First, it develops a new method for recovering the preferences (model primitives) of agents, who are matched only once in a two-sided matching market with moral hazard. These primitives are important for evaluating the efficiency of equilibrium allocation and assessing the role of policy designed in affecting equilibrium allocation outcomes. The existing literature focuses on estimating the aggregate surplus of matched partners in matching markets without asymmetric information. My method has two advantages. First, I provide a method to estimate the *separate* preferences between matched partners, while previous papers focus on estimating the aggregate surplus. Second, I estimate the primitives of a matching model with asymmetric information, which has not been attempted before. Therefore, the method developed here may be applicable to study other matching markets with asymmetric information.[3]

Second, this method is applied to estimate preferences in the U.S. market for CEOs Using the estimates, I empirically quantify the magnitude of inefficiency in the allocation of CEOs to firms caused by moral hazard. Moral hazard brings inefficiency due to the need for the shareholders to share risk with CEOs, which is well understood in the literature. However, the magnitude of the inefficiency in the allocation of CEOs to firms also stemming from moral hazard has not been empirically quantified.[4] I estimate a model in which moral

[1] Pan (2012) shows three productive complementarities between executive and firm characteristics: firm size and managerial talent, the degree of diversification of the firm and the cross-industry experience of the manager, the R&D intensity of the firm and the innovation propensity of the manager.

[2] Becker (1973) shows that in a perfect market, when the production technology is complementary, positive assortative matching is achieved in all equilibria.

[3] Examples include venture investment, book selling and many other high-skilled labor markets.

[4] Actually both traditional risk-sharing inefficiency and allocation inefficiency are caused by the need of

hazard can lead to an inefficient allocation of CEOs to firms in equilibrium. I then quantify the magnitude of this inefficiency using the estimates of the model primitives. I find that the efficiency loss of CEO misallocation is large: namely this is more than *four times* as large as the loss induced by risk-sharing inefficiency. The findings suggest that studies focusing solely on risk-sharing can severely underestimate efficiency losses and impacts of moral hazard on allocation efficiency should be considered in future work.

Why would moral hazard possibly matter for the allocation of CEOs to firms? To explain this, consider first a competitive market without moral hazard. Suppose the production of firms and CEOs is complementary in their types. Since higher type CEOs generate more value at higher type firms, the equilibrium allocation of CEOs to firms exhibits positive assortative sorting. This allocation is efficient in the sense of maximizing the total surplus. Now introduce moral hazard into this market. If output of firms is stochastic and depends on unobserved managerial action, firms have to pay performance-based wages to CEOs in order to motivate them. Thus, firms need to pay a risk premium to risk-averse CEOs and then riskier firms have to pay higher expected wages. In this context, for a very risky high type firm, although a higher type CEO can generate more value at this firm, this firm also needs to pay a very high expected wage to motivate him. As a result, this firm may not get a high type CEO because it cannot afford to compensate him for the risk.

I develop a model to illustrate this insight. The model extends the matching model in Tervio (2008) and incorporates moral hazard following Gayle and Miller (2009). In a competitive market, risk-averse CEOs differing in types (talent) are hired by firms varying in types (size), risk and costs of managerial effort. To focus on moral hazard, I assume that CEOs have private information about their effort, all characteristics are known to firms and CEOs but not to the econometrician.[5] Specifically, firms know managerial talent and CEOs know all firms' size, risk and costs of managerial effort. The allocation of CEOs to firms in equilibrium is driven by the level of CEO compensation, which is determined by the optimal wage contracting in the presence of moral hazard. Since firms differ in multidimensional characteristics, it is challenging to solve the equilibrium allocation of CEOs to firms. Making the following assumptions: managers have constant relative risk aversion preferences with two possible effort levels (shirking and working) to choose from, and the production function is complementary in firms' size and managerial talent, I could collapse multidimensional

shareholders to share risk with CEOs in the presence of moral hazard. While the existing literature focuses only on the traditional risk-sharing inefficiency, I study both risk-sharing and allocation inefficiency.

[5] Motivated by Tervio (2008), the size (actual) of firms should be thought of as the combination of all firm factors that contribute to production, including reputation, market capitalization, growth potential, and so on. Similar logic applies to managerial talent. Therefore, it is reasonable to assume that they are not observed by the econometrician.

characteristics of firms into a one single index, namely effective size. I then show that the equilibrium allocation of CEOs to firms exhibits positive assortative sorting by firms' effective size and managerial talent. This may or may not be efficient because the total surplus is maximized when the way that CEOs is allocated to firms is positive assortative sorting by firms' actual size and managerial talent.

To identify the model, I must confront two facts that differ from the data usually studied from other matching markets. There two facts are (i) the types of firms and CEOs are not observed to the econometrician; (ii) firms and CEOs with different types are matched only once. Instead of a standard revealed preference approach, I identify the model using observed monetary transfers (CEO compensation), profits of firms, and information in the environment with moral hazard. Tervio (2008) shows that the type distributions of firms and CEOs can be inferred from observed CEO compensation and the profits of firms. Following this idea, I recover the type distributions of CEOs and firms from their observed income. With their type distributions in hand, the production function is then identified under complementarity assumption. While the presence of moral hazard brings challenges in solving the model, it provides additional information for identifying the managerial preferences from the model structure. In particular, managerial preferences can be identified by exploiting optimal compensation contracts, managers' participation and incentive compatibility constraints.

Building on the identification strategy, I propose a technique for estimating the primitives of the model. The technique is applied to estimate the model using data on the 1000 largest firms in terms of market value from S&P Compustat databases for 2011. The estimation results show that the ordering of firms' effective size differs from that of their actual size, which implies moral hazard does affect the efficient allocation of CEOs to firms. While the efficient allocation requires positive assortative sorting by firms' actual size and managerial talent, the equilibrium allocation exhibits positive assortative sorting by firms' effective size and managerial talent. I then use the estimates to investigate the importance of moral hazard in the allocation of CEOs to firms by quantifying four measures of losses related to moral hazard. The counterfactual results show that (i) the aggregate loss of the 1000 largest firms due to talent misallocation is \$12.64 billion, which is more than *four times* as large as the loss induced by the risk-sharing inefficiency; (ii) maximum aggregate loss from talent misallocation is 15% of the total market value, while the aggregate loss of not motivating CEOs is 19.69% of the total market value.

The empirical method in this paper contributes to the recent literature on estimating matching models with transferable utility. The majority of papers focus on estimating a single aggregate surplus in markets without asymmetric information. The empirical analysis of matching models with transferable utility is initiated by Choo and Siow (2006), and subsequently extended by Galichon and Salanié (2010), Chiappori et al. (2011), and others. Fox (2008) proposes a different approach to estimate matching models of transferable utility, which is applied by Bajari and Fox (2005), Pan (2012), and others. In these papers, the possibility of recovering two *separate* utility functions is limited because data on monetary transfers between matched partners are often not observed. Moreover, the models in these papers do not feature asymmetric information. To the best of my knowledge, this is the first paper that estimates matching models with asymmetric information.

This paper also contributes to an extensive literature on moral hazard and matching in the market for CEOs.[6] The empirical literature analyzes moral hazard and CEO matching separately. Seminal papers by Murphy and Jensen (1990) and Hall and Liebman (1998) study the importance of moral hazard by estimating CEO pay-performance sensitivity. Margiotta and Miller (2001), and Gayle and Miller (2009) analyze principal agent models to estimate economic costs of moral hazard. Gabaix and Landier (2008), and Tervio (2008) employ CEO matching models to investigate the cross sectional differential and time trend of CEO compensation.[7] In these papers the cost of CEO misallocation caused by moral hazard is unable to be quantified because moral hazard and CEO allocation are investigated separately and incorporating them is not an obvious extension. The theory in this paper is closely related to Edmans et al. (2009) and Edmans and Gabaix (2011a), which also incorporates moral hazard into matching models. My model differs from theirs' in its empirical feature that all primitives of the model can be recovered from observables.

The rest of the paper is organized as follows. Section 2 presents the theoretical model on which my empirical analysis is based. The model shows that moral hazard can lead to an inefficient allocation of CEOs to firms. Section 3 proposes a new empirical method to recover the preferences of CEOs and firms in a matching framework with moral hazard. Section 4 introduces data on the U.S. market for CEOs in 2011. While section 5 presents the estimates of the model primitives and counterfactual results about the importance of moral hazard in the allocation of CEOs to firms, section 6 concludes.

[6] There is also an extensive theoretical literature on moral hazard and matching. The theoretical papers on moral hazard mainly focus on deriving the optimal incentive wage contract by solving principal agent models of moral hazard. See, e.g., Mirrlees (1976), Hölmstrom (1979, 1982), Grossman and Hart (1983), Hölmstrom and Milgrom (1987) and Edmans and Gabaix (2011b). The theoretical papers on matching try to understand the distribution of labor income. See, e.g., Tinbergen (1951, 1956), Koopmans and Beckmann (1957), Sattinger (1979, 1993) and Rosen (1982).

[7] See, e.g., Chiappori and Salanié (2003) for a survey of the contracting literature.

2 The Model

This section lays out a matching model of managers and firms featuring moral hazard in the market for CEOs as the theoretical framework of my empirical analysis. The interactions between managers and firms in my static framework are modeled to be in two stages: (i) firms and managers match one to one to produce output in a competitive labor market; (ii) firms offer compensation contracts to their hired managers. I start with the second stage by formulating firms' optimal contracting problem. I then derive firms' optimal compensation contracts in the presence of moral hazard. Taking the optimal contracts into account, I derive the equilibrium allocation of managers and firms. Finally I show how moral hazard can affect the efficient CEO allocation and provide measures of loss to firms generated from moral hazard.

2.1 Managers

In this model there is continuum of managers in a competitive market. They differ in their talent, whose distribution is captured by a quantile function $T(m)$. $T(m)$ is the talent of a m quantile manager.[8] A typical manager's preference is captured by using a constant relative risk aversion (CRRA) utility function. All managers have the same coefficient of relative risk aversion, denoted by ρ. For a manager with talent $T(m)$ serving for firm n, the cost of his effort is captured by the coefficient α_{ne} corresponding to two levels of managerial effort $e \in \{1, 2\}$, where $e = 1$ represents shirking and $e = 2$ represents working.[9] Here α_{ne} is firm specific. The manager's utility can then written as

$$U_n(w(x)) = \frac{(\alpha_{ne} w(x))^{1-\rho}}{1-\rho} \quad \text{for } e = 1 \text{ or } 2. \tag{1}$$

I assume that a typical manager prefers shirking to working, that is $\alpha_{n1} > \alpha_{n2} > 0$. The manager's reservation utility from outside options is assumed to be $\alpha_{m0}^{1-\rho}/(1-\rho)$, which depends on the manager's talent $T(m)$.

With CRRA preference, effort has a multiplicative effect on a manager's utility. This preference captures the idea that a manager's utility from shirking is increasing with their wealth. This specification is plausible if effort is interpreted as forgoing leisure. A day of vacation is more valuable to a richer manager as he has more wealth to enjoy it. CRRA preferences are also commonly used in macroeconomics as it leads to realistic income effects.

[8] Denoting the distribution function of firm size as $F_s(\cdot)$, $T(m)$ is defined by $T(m) = T,\ s.t.\ F_t(T) = m$.

[9] Shirking and working should be thought of as managers pursue their own interests and those of firms, respectively.

Moreover, the specification of CRRA utility function is necessary to achieve multiplicative form of optimal compensation, which is desired to solve the equilibrium CEO allocation outcome in my model.

2.2 Firms

There is also continuum of firms in the competitive market. They are characterized by three dimensional characteristics: size, disutility of effort and risk. The firm size should be thought of as the combination of all firm characteristics that contribute to production, such as assets, reputation, capitalization ability and so on. Its distribution is also captured by a quantile function $S(n)$. $S(n)$ is the size of a n quantile firm.[10] The disutility of managerial effort, which is captured by α_{ne}. It is assumed to be firm specific. For example, a firm in a regulated industry or headquartered in an unattractive location is unpleasant to work for regardless of the effort exerted by the CEO.[11]

The risk of firms is reflected by the volatility of their output, which is produced after a manager is allocated to a firm. The production output is specified as a function of firm's size, managerial talent and effort. Specifically, the production function is assumed to be multiplicatively separable between the firm's and the manager's contribution to output. Consider a firm with size $S(n)$ has hired a manager with talent $T(m)$. Its production function is written as [12]

$$Y[S(n), T(m), e] = S(n) \cdot T(m) \cdot [1 + x(e)]. \tag{2}$$

This is the simpliest form that exhibits complementarity.

$x(e)$ is the idiosyncratic output signal attributed to managerial effort, whose probability density function is conditional on managerial effort e. For firm n, conditioning on manager shirking ($e = 1$), probability density function of x is denoted by $f_{n1}(x)$. Conditioning on manager working ($e = 2$), probability density function of x is denoted by $f_{n2}(x)$. Since x is the only stochastic variable in the output function, the variances of $f_{n1}(x)$ and $f_{n2}(x)$ reflect the risk of firms. Hereafter throughout this paper, the symbol $E_n[\bullet]$ is employed to represent the expectation taken over $f_{n2}(x)$, i.e., $\int \bullet f_{n2}(x)dx$. For the purpose of solving the model, I define likelihood ratio function by $g_n(x) \equiv f_{n1}(x)/f_{n2}(x)$.[13] It is nonnegative for all x and

[10] Denoting the distribution function of firm size as $F_s(\cdot)$, $S(n)$ is defined by $S(n) = S, \;\; s.t. \;\; F_s(S) = n$.

[11] This is the argument in Edmans and Gabaix (2011a)

[12] The production function is motived by Tervio (2008).

[13] Following Gayle and Miller (2009), $g_n(x)$ can be interpreted as the signal firm n receives about managerial effort choice. $g_n(x) = 0$ implies working while $g_n(x) = \infty$ implies shirking. When $g_n(x) = 1$, the signal is useless such that compensation does not depend on x.

$E[g_n(x)] = \int x f_{n1}(x) dx = 1$.

I make three important assumptions on probability density functions of x, which are crucial to my analysis. First, I normalize the expected value of x conditional on manager working to be zero, which is formalized in the next assumption.

Assumption 1. *$E_n(x)$ is normalized to be 0 for any firm n.*

Second, I assume each firm perfers manager working to shirking. That is the expectation of x is larger under manager working than it under manager shirking.

Assumption 2. *$E_n[x] = \int x f_{n2}(x) dx > \int x f_{n1} dx = E_n[x g_n(x)]$ for any firm n.*

Finally, I assume that a very large x is extremely unlikely to be obtained if the manager shirks, which is formalized in next assumption.

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

In other words, this assumption tells that an extraordinary high level x can only be achieved when the manager works.

2.3 Timing

The timing of interactions between managers and firms in my static framework is as follows. First, at the beginning of the period, heterogeneous firms and managers match one to one in a competitive labor market. I assume there is no search friction in this market. All characteristics of firms and managers are observed by firms and managers but not by the econometrician. The continuous distributions of firms and managers rule out match-specific rents and therefore, any need to model bargaining. Second, after managers are matched with firms, each firm proposes a performance-based compensation contract $w(x)$ to its hired manager. Recall x is the output signal whose distribution is conditional on the level of managerial effort. Given the firm's compensation offer, each manager makes his choice of whether to take the offer or not. If the manager rejects the offer, he obtains his reservation wage from outside options. If the manager accepts the offer, he then chooses his effort level. The effort choice is not observed by the shareholders of the firm. This hidden action moral hazard problem is the only friction in my model. Finally, the output signal x is realized and each manager gets paid according to his compensation contract $w(x)$.

2.4 Optimal Contracting

I solve the optimal compensation in two steps.[14] First, I solve the optimal compensation contracts corresponding to both effort levels: shirking and working. Second, the optimal effort level is determined by the firm's profit maximization problem.

Consider the problem of a firm with size $S(n)$ which has hired a manager with talent $T(m)$. I assume firms are risk neutral. Their utility is then measured by their profits, a benefit net of a cost. Let's first consider the optimal compensation contract problem of a firm which wants to induce its manager to work. The firm's cost is then the expected compensation paid to the manager. The firm's benefit is the expected output when its manager works, which is a constant because firm's size $S(n)$ and managerial talent $T(m)$ are fixed and the expected value of x when the manager works, $E_n(x)$, is normalized to be zero from assumption (1).

The firm's problem is now to design an optimal compensation contract to minimizing its cost, the expected compensation $\int w_n(x) f_{n2}(x)dx$, subject to (i) the manager accepts the offer instead of choosing his outside options (participation constraint); (ii) and the manager pursues the interests of the firm instead of his own (incentive compatibility constraint). They can be formally written as

$$\int \frac{[\alpha_{n2} w_n(x)]^{(1-\rho)}}{(1-\rho)} f_{n2}(x)dx \geq \frac{\alpha_{n0}^{(1-\rho)}}{(1-\rho)}, \tag{3}$$

and

$$\int \frac{[\alpha_{n2} w_n(x)]^{(1-\rho)}}{(1-\rho)} f_{n2}(x)dx \geq \int \frac{[\alpha_{n1} w_n(x)]^{(1-\rho)}}{(1-\rho)} f_{n1}(x)dx. \tag{4}$$

Intuitively, (3) requires the expected managerial utility from working is not less than the utility that the manager obtains from outside options. (4) requires the expected utility that the manager obtains from working is not less than the utility from shirking, which gives the manager incentive to work.

Solving the firm's cost minimization problem, I derive the optimal compensation contract that induces the manager to work. The next lemma formalizes the result.

Lemma 1. *The firm n's optimal contract inducing its manager to work is given by*

$$w_n(x) = (\alpha_{n0}/\alpha_{n2})\{\theta_0 + \theta_1[1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x)]\}^{1/\rho}, \tag{5}$$

[14] The optimal compensation contracts are first solved under a principal-agent framework by Grossman and Hart (1983). Here I follow their procedure,

where θ_0 and θ_1 is the unique positive solutions to the following system of equations,

$$E\{[\theta_0 + \theta_1(1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x))]^{(1-\rho)/\rho}\} = 1, \tag{6}$$

$$E\{[\theta_0 + \theta_1(1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x))]^{(1-\rho)/\rho} g_n(x)\}(\alpha_{n1}/\alpha_{n2})^{1-\rho} = 1. \tag{7}$$

All proofs are in the appendix. Optimal compensation for working is the multiplication of a constant term and a risk primium term. If moral hazard is not a problem because managerial action could be perfectly monitored, the manager would be paid a flat compensation, α_{n0}/α_{n2}. The second term in the optimal compensation determines how it varies with output signal x through the likelihood ratio function $g_n(x)$. If a firm is very risky or generate high costs for managers making effort, variance of x or α_{n0}/α_{n2} is large, the expected value of this second term will be large. Thus this firm need to pay a higher expected wage.

If the firm wants to induce the manager to shirk, its optimal compensation is a flat wage α_{n0}/α_{n1}, which ensures the following participation constraint for shirking to hold with equality,

$$\int \frac{[\alpha_{n1} w_n(x)]^{(1-\rho)}}{(1-\rho)} f_{n1}(x)dx \geq \frac{\alpha_{n0}^{(1-\rho)}}{(1-\rho)}.$$

Profit maximization determines whether firms should offer the optimal contracts that induce managers to work or to shirk. Hereafter I assume that inducing managers to work gives firms more profits, and as such focus only on the case where all firms offer optimal compensation contracts that induc their manangers 5o work.[15]

2.5 Equilibrium Allocation of Managers to Firms

I now incorporate the above moral hazard problem into an matching model of managers and firms. The matching setup follows closely the model of Tervio (2008). Under the complementary production, efficiency requires positive assortative sorting by managers and firms. In the absence of moral hazard, the competitive equilibrium exhibits this sorting.[16] In the presence of moral hazard, this positive assortative sorting may not be achieved in the competitive equilibrium. In the following I show how moral hazard can affect this positive assortative sorting.

I begin with the expected utility of managers. Using the optimal compensation derived

[15] I observe that all the firms provide compensation based on their performance in the data and there is no firm paying flat wage.

[16] See, eg., Sattinger (1993), Gabaix and Landier (2008) and Tervio (2008)).

in Lemma 1, the expected utility of firm n's manager can be written as[17]

$$EU_n = \frac{\{\alpha_{n2} E[w_n(x)] e^{-\chi_n}\}^{1-\rho}}{1-\rho}, \tag{8}$$

where $E[w_n(x)]$ is the expected compensation and χ_n is defined by

$$\chi_n \equiv \ln E\{[\theta_0 + \theta_1(1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g(x))]^{\frac{1}{\rho}}\}. \tag{9}$$

χ_n is the risk premium that firm n pay to its manager in the sense that $\chi_n = \ln\{E[w_n(x)]\} - \ln U^{-1}(EU_n)$. Riskier and higher disutility firms need to pay a high risk premium in terms of χ_n. From the perspective of the manager, χ_n is the total loss he/she suffered from working and sharing risk with the firm.

After adjusting for the loss, the certainty equivalent wage (named effective wage) of firm n's manager is defined by

$$v_n = E[w_n(x)] e^{-\chi_n}, \tag{10}$$

which is fixed and gives the manager the same expected utility as that under the optimal compensation contract. From the proof of Lemma 1, the participation constraint (3) is met with equality and hence we have $v_n = \alpha_{n0}/\alpha_{n2}$, which is nonpecuniary benefits the manager obtains from outside options versus working. The outside options for the manager should be thought of as the opportunities to work for firms whose sizes are next to firm n. In this sense the effective wages of managers are endogenously determined by the competition for managers among firms, meaning more talented managers obtain higher effective wages in equilibrium.

In the competitive equilibrium, a manager with ability $T(m)$ should receive an effective wage of $v(m)$ whose quantile is m. If firm n wishes to hire him, it has to pay him an effective wage $v(m)$ and thus an expected dollar wage of $E[w_n(x)] = v(m)e^{\chi_n}$. The expected profits that the firm will obtain is the expected output net of expected wage of the manager,

$$\begin{aligned} &E[S(n)T(m)(1+x)] - v(m)e^{\chi_n} \\ &= S(n)T(m) - v(m)e^{\chi_n} \\ &= e^{\chi_n}[S(n)e^{-\chi_n}T(m) - v(m)], \end{aligned}$$

where the first equality is from the normalization $E_n(x) = 0$. The firm chooses a manager to maximize above expected surplus, which tells that firm n with actual size $S(n)$ now acts

[17] The derivation of the expected utility is in the appendix.

as a firm with size $S(n)e^{-\chi_n}$, namely "effective size". Riskier and higher disutility firms have lower value of $e^{-\chi_n}$ and thus smaller effective size given actual size. Hereafter I order firms using the effective size; and thus suppose firm n's effective size has quantile r, denoted by $\tilde{S}(r) \equiv S(n)e^{-\chi_n}$.

In absence of moral hazard, firms act according to their actual sizes. The competitive equilibrium is positive assortative sorting by firm actual size and manager talent, which is efficient. In the presence of moral hazard, efficieny still requires positive assortative sorting. However, in this circumstance, firms act according to their effective sizes, the competitive equilibrium involves positive assortative sorting by firm's effective size and managerial talent[18]. This may or may not be efficient because efficiency involves postitive assortative sorting by firm's actual size and managerial talent. If the ranking of firms' effective size is different from that of firms' actual size, the equilibrium outcome is not efficent in the present of moral hazard. The main purpose of this paper is to study the importance of moral hazard on the allocation of CEOs to firms by quantifying the magnitude of this inefficiency.

In the presence of moral hazard, the profiles of manager effective wage and firm effective profits must support the allocation that involves perfect sorting by firm effective size and managerial talent. The following two types of conditions must be satisfied,

$$\tilde{S}(r)T(r) - v(r) \geq \tilde{S}(r)T(m) - v(m) \qquad \forall r, m \in [0,1] \tag{11}$$

$$\tilde{S}(r)T(r) - v(r) \geq \pi^0 \qquad \forall r \in [0,1] \tag{12}$$

$$v(r) \geq v^0 \qquad \forall r \in [0,1] \tag{13}$$

$\tilde{S}(r)T(r)$ is effective output, which is divided between the managers and firms. (π^0, v^0) is the effective wage and effective profits that managers and firms could obtain from the job opportunities outside the market. The lowest active firm-manager pair ($r = 0$) is the one that just breaks even with the alternative opportunity outside the market, $\tilde{S}(0)T(0) = v^0 + \pi^0$. First type conditions (11) guarantee each firm must prefer hiring its manager to hiring any other managers at their equilibrium effective wages. Second type conditions (12) and (13) guarantee all firms and managers are active in the market.

Replacing m with $r - \epsilon$ in sorting constraints (11) and dividing both sides by ϵ gives

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

which becomes equility as $\epsilon \to 0$. By definition, the slope of the managerial effective wage

[18] The interpretion on it is shown in the appendix

profile is given by [19]

$$v'(r) = \tilde{S}(r)T'(r). \tag{15}$$

The effective wage profile then can be obtained by integrating the slope and using $v(0) = v^0$:

$$v(r) = v^0 + \int_0^r S(j)T'(j)dj. \tag{16}$$

Analogously by the fact that $\pi(r) = \tilde{S}(r)T(r) - v(r)$, the profile of firm effective profits satisfies

$$\pi'(r) = \tilde{S}'(r)T(r), \tag{17}$$

which gives

$$\pi(r) = \pi^0 + \int_0^r S'(j)T(j)dj. \tag{18}$$

From (16) and (18), we see that all inframarginal manager-firm pairs produce an effective output over the sum of their opportunities outside the market, and the division of this effective output depends on the distributions of firm effective size and managerial talent. The effective profits and wage of firms and managers increase with their quantiles. The following definition summarizes the conditions that a competitive assignment equilibrium should satisfy when taking moral hazard into account.

Definition 1. *When moral hazard is present, a competitive allocation equilibrium is defined by the allocation of managers to firms and a profile of firm effective profits and managerial effective wage, which satisfy the following conditions:*
(i) the assignment satisfies the conditions presented in (11)-(13);
(ii) the effective wages and effective profits are given by distribution functions of firms' size and managerial talent in (16) and (18).

[19] This equation can also be obtained by solving firm's profit maximization problem. The firm n chooses a CEO to maximize its expected profits and thus it solves the maximization problem,

$$\max_m \quad \tilde{S}(r)T(m) - v(m). \tag{14}$$

Taking first derivative with respet to m and using positive assortative matching in equilibrium ($m = r$) gives us (14).

2.6 Measuring the Losses due to Moral Hazard

The importance of moral hazard on the allocation of CEOs to firms is assessed by quantifying four measures of losses related to moral hazard. First, in the presence of moral hazard, firms offer performance-based compensation contracts that lead CEOs to share risk with the firms. Consequently firms incur losses by paying risk premium to risk-averse CEOs. Second, firms incur production losses arising from talent misallocation. The third measure is the maximum losses that firms would incur from talent misallocation caused by moral hazard. The final measure is the loss that firms would incur from ignoring moral hazard problem.

First, I provide a measure to the loss from risk-sharing inefficiency, which are the costs that firms pay to motive CEOs when moral hazard is present. If moral hazard is not a problem because of perfect monitoring, the firm with quantile n pays a fixed ceretainty equivalent wage $\alpha_{n0}/\alpha_{n2} = v_n$. In the presence of moral hazard, firm n offers a wage contract $w_n(x)$ and expected wage $E[w_n(x)]$, which is worth to the fixed wage v_n. The loss of firm n from risk-sharing inefficiency is the difference between the expected wage $E[w_n(x)]$ and the fixed wage v_n. Denoting it by L_{n1}, it is given by

$$L_{n1} = E[w_n(x)] - v_n = E[w_n(x)][1 - e^{-\chi_n}], \tag{19}$$

where the second equality comes from the definition $v_n = E[w_n(x)]e^{-\chi_n}$. The sum of L_{n1} over all firms gives the gross loss firms would incur from risk-sharing inefficiency.

Second, I provide a measure to the loss from talent misallocation associated with moral hazard. It is measured by the difference between firm n's output from efficient talent allocation and that from actual talent allocation. Let $\tilde{T}(n)$ denote the talent of the manager assigned to firm n in equilibrium[20]. Denoting the loss by L_{n2}, it is then given by

$$\begin{aligned} L_{n2} &= E[S(n)T(n)(1+x)] - E[S(n)\tilde{T}(n)(1+x)] \\ &= S(n)[T(n) - \tilde{T}(n)], \end{aligned} \tag{20}$$

where the second equality exploits the normalization $E(x) = 0$, that is the expected value of x is zero when the manager pursues the interests of the firm. The sum of L_{n2} over all firms give the gross loss firms would incur from matching inefficiency. The sum of L_{1n} and L_{2n} over all firms are the total loss that firms incur to solve moral hazard problem by optimal contracting.

Third, I provide a measure to the maximum loss that firms would incur from talent

[20] Since we know the firm n has effective size $\tilde{S}(r)$, the firm will end up with talent $T(r)$ manager. Thus, $\tilde{T}(n) = T(r)$

misallocation. This loss the difference between the output of efficient talent allocation and that of the most inefficient talent allocation. The most inefficient talent allocation is negative assortative sorting, that is the least talented manager is allocated to largest firm.

Finally, I provide a measure to the loss that firms would incur from ignoring moral hazard. This is the total benefits of firms from motivating the CEOs. If a firm ignores moral hazard problem, its manager will pursue the interests of his own. Let L_{3n} denote the loss of firm n from ignoring moral hazard. The loss is measured by the difference between the expected output from the manager pursuing the firm's interests versus his own, which is given by

$$L_{3n} = \int S(n)\tilde{T}(n)(1+x)f_{2n}(x)dx - \int S(n)T(n)(1+x)f_{1n}(x)dx. \tag{21}$$

The talent allocation will be efficient if all firms ignore moral hazard since the firm will pay a fixed wage to the manager. Thus the second term at the right side of above equation gives expected output when the firm ignores moral hazard.

2.7 A Numerical Illustration

I now provide a numerical example to illustrate the intuition of the model. In this example, I focus on how the efficient allocation is distorted and the losses that arise from this distortion. Consider a market for CEOs with a continuum of firms and managers in the market for CEOs. The distributions of firm's actual size and managerial talent are represented by quantile functions, which are assumed to be $S(n) = 100n$ and $T(m) = 10m$ for $n, m \in [0, 1]$. For a typical firm n, the probability density functions of x under manager working and shirking are assumed to be normal distributions,

$$x_n \sim \begin{cases} N(0, \sigma_n^2) & \text{if working,} \\ N(-1, \sigma_n^2) & \text{if shirking.} \end{cases}$$

Here σ_n^2 is the variance of the probability density functions, which reflects firms' risk. Here I assume managerial effort does not affect the risk of firms. The firms' disutility is assumed to be the same across firms as I focus on the effect of firms' risk. α_{n1}/α_{n2} is then set to be 4. Finally I set the risk aversion parameter $\rho = 1/2$.

Under above setup, I first derive the θ_0 and θ_1 for each firm following Lemma 1. The solutions are given by

$$\theta_{n1} = 1 + \frac{1}{2(\exp(1/\sigma_n^2) - 1)},$$
$$\theta_{n2} = \frac{1}{2(\exp(1/\sigma_n^2) - 1)}.$$

With θ_{n0} and θ_{n1} in hand, risk premium χ_n can be calculated by using (9). Figure (1) displays how it changes with firms' risk. It shows that riskier firms need pay more risk premium, which is consistent with our model intuition. Using the risk premium, I can derive the firms' effective size from its definition $\tilde{S}(r) \equiv S(n)e^{-\chi_n}$. The equilibrium allocation involves perfect sorting by firms' effective size and managerial talent.

I first randomly assign variances to firms from $\sigma_n^2 \in [1, 11]$. Table (1) shows the equilibrium allocation. The matching is distorted in the sense that some small firms hire relatively high talented managers and some large firms hire relatively low talented managers. The loss from allocation inefficiency due to moral hazard is measured by using (20). Figure (2) shows the efficiency loss. The total value of the loss is 2575, which is about 7.61% of total output under the efficient allocation. When the variances of x are assigned to be positively correlated with firms' actual sizes, total efficiency loss is at its maximum level, which is 10520. This is about 31.1% of total output under the efficient allocation.

3 Identification and Estimation

3.1 Identification

After the model is defined and solved, I identify and estimate the model primitives. I first define the structural model primitives and observables. The model primitives consist of probability density functions of firm abnormal returns x under both shirking and working, $f_{n1}(x)$ and $f_{n2}(x)$, nonpecuniary benefits of managers from outside options versus working α_{n0}/α_{n2}, nonpecuniary benefits of managers from shirking versus working α_{n1}/α_{n2}, managers' risk aversion parameter ρ, and quantile functions of firm size and managerial talent, $S(n)$ and $T(n)$. Among these primitives, $f_{n1}(x)$, $f_{n2}(x)$, $S(n)$ and $T(n)$ characterize the production function of firms. α_{n0}/α_{n2}, α_{n1}/α_{n2}, and ρ chracterize the utility function of managers.

These primitives are identified from observed monetary transfer between CEOs and firms, profit and performance of firms, and the final matches. The monetary transfer between CEOs and firms is CEO compensation. The profit of firms is measured by their market value. The performance of firms is measured by their abnormal returns. Thus the observables are comprised of firm abnormal returns x, firm market value V, and managerial compensation

w. The data on these observables are assumed to be available for a sample of N observations generated in equilibrium. The observables for the sample are then denoted by $\{x_i, V_i, w_i\}_{i=1}^N$.

Heterogeneity in firms' risk and disutility, reflected by the variances of abnormal returns distributions and nonpecuniary benefits from shirking and working, are the key in this analysis. The heterogeneity is introduced by using the observable covariates of firms. In other words, conditional on firms' observable covariates, firms have the same distributions of abnormal returns under both shirking and working, and nonpecuniary benefits from shirking and working. For the purpose of interpreting identification, I discuss the identification of model primitives after controlling the observed covariates of firms. After suppressing firm index n, these primitives are $f_1(x)$, $f_2(x)$, α_0/α_2, α_1/α_2, $S(n)$ and $T(n)$.

Since we observe all firms tie their managerial compensation to their performances in the data, the identification of model primitives focuses on the case in which working is induced and compensation w_n depends nontrivially on abnormal returns x_n. In this case, probability density function for working $f_2(x)$ is nonparametrically identified from abnormal returns x_n and the optimal compensation function $w(x)$ can be nonparametrically identified from x_n and w_n. $f_1(x)$ is not directly identified from the observed abnormal returns x_n. However, it can be recovered by first identifying the likelihood ratio $g(x)$. If $g(x)$ is identified, $f_1(x)$ is identified by $f_1(x) = g(x)f_2(x)$ as $g(x)$ and $f_2(x)$ have been identified. This leaves to identify $g(x)$, α_0/α_2, α_1/α_2, ρ, $S(n)$ and $T(n)$ from $f_2(x)$, $w(x)$ and V.

The identification on these remaining model primitives is established in three steps. First, I show that if the risk aversion parameterρ is known, the likelihood ratio $g(x)$, nonpecuniary benefits α_0/α_2 and α_1/α_2 are identified by exploiting optimal compensation equation, participation and incentive compatibility constraints with equality. Second, I discuss that the risk aversion parameter ρ is identified under the assumption that there are data on at least two states where the nonpecuniary benefits of managers from outside options versus working are the same. Finally, I show that firms' effective size and managerial talent quantile functions, $\tilde{S}(n)$ and $T(n)$, are identified by using the equilibrium matching conditions in equilibrium from observables and identified variables. The actual size is recovered from identified effective size $\tilde{S}(n)$ by exploiting the definition of effective size.

IDENTIFICATION OF $g(x)$, α_0/α_2 AND α_1/α_2 GIVEN ρ

Suppose the risk aversion parameter ρ is known, it can be shown that $g(x)$, α_0/α_2 and α_1/α_2 can be expressed as functions of identified functions, $f_2(x)$ and $w(x)$. The next lemma formalizes this result.

Lemma 2. *Under assumption (3) and the fact $E[g(x)] = 1$, if risk aversion parameter ρ is* *known, $g(x)$, α_0/α_2 and α_1/α_2 are given by*

$$g(x) = \frac{\overline{w}^\rho - w(x)^\rho}{\overline{w}^\rho - E[w(x)^\rho]}, \tag{22}$$

$$\alpha_0/\alpha_2 = \{E[w(x)^{1-\rho}]\}^{1/(1-\rho)}, \tag{23}$$

$$\alpha_1/\alpha_2 = \Big\{\frac{\overline{w}^\rho - E[w(x)^\rho]}{\overline{w}^\rho - E[w(x)]/E[w(x)^{1-\rho}]}\Big\}^{1/(1-\rho)}, \tag{24}$$

where $\overline{w}$ is the maximum compensation managers can receive, which can be identified from the maximum compensation observed in the data.

The basic ideas for the proof are to exploit the optimal wage equation (5), participation constraint (3) and incentive compatibility constraint (4) with equality. Rearranging the optimal wage equation (5) and differentiating it with respect to x yields

$$g'(x) = -(\rho\theta_1)^{-1}(\alpha_1/\alpha_2)^{1/(1-\rho)}(\alpha_0/\alpha_2)^{-1/\rho}w(x)^{1-\rho/\rho}\partial w(x)/\partial x,$$

which implies that this slope is defined up to one normalization. A second normalization is needed to determine the level of $g(x)$ from its slope. Assumption (3) provides us the first normalization and the fact that $E[g(x)] = 1$ provides us the second. The proof of Lemma (1) shows that the participation constraint (3) holds with equality at optimal compensation, which gives us the equation (23) for α_0/α_1. The incentive compatibility constraint (4) with equality gives

$$E[w(x)^{1-\rho}] = (\alpha_1/\alpha_2)^{1-\rho}E[w(x)^{1-\rho}g(x)].$$

Substituting $g(x)$ from (22) on the right side and rearranging it yield (24) for α_1/α_2.

I now show that $g(x)$, α_0/α_2 and α_1/α_2 given by (22)-(24) can be served as primitives for the model presented in section (2). First, let us consider $g(x)$. Intergrating (22) over x yields that $E[g(x)] = 1$ for all $\rho > 0$. $\overline{w} \geq w(x)$ guarantees both denominator and numerator in (22) are nonnegative, which gives us $g(x) \geq 0$ for all $\rho > 0$. From the proof of Lemma (2), we have $\lim_{x\to\infty} w(x) = \overline{w}$, which implies $\lim_{x\to\infty} g(x) = 0$ from (22). These prove that $g(x)$ can be served as a likelihood ratio for the model. Second, let us consider α_0/α_2 and α_1/α_2. From (23), $\alpha_0/\alpha_2 > 0$ because $w(x)^{1-\rho} > 0$. Similarly $\alpha_1/\alpha_2 > 0$ as $\overline{w}^\rho > E[w(x)^\rho]$ and $\overline{w}^\rho - E[w(x)]/E[w(x)^{1-\rho} = \theta_1 > 0$ from the proof of Lemma 2. Moreover, it is easy to show that $\alpha_1 > \alpha_2$ for any positive ρ by Jensen Inequality.[21] It is therefore that α_0/α_2 and

[21] If $0 < \rho < 1$, $w(x)^\rho$ and $w(x)^{1-\rho}$ are both concave. From Jensen Inequality, I have $E[w(x)^\rho] < \{E[w(x)]\}^\rho$ and $E[w(x)^{1-\rho}] < \{E[w(X)]\}^{1-\rho}$, which together gives us $E[w(x)]/E[w(x)^{1-\rho} > E[w(x)^\rho$. Thus, I have $\alpha_1 > \alpha_2$. Similarly if $\rho > 1$, I can also show $\alpha_1 > \alpha_2$.

α_1/α_2 can both be primitives for the model.

The above discussion shows that for any positive ρ, the model primitives $g(x)$, α_0/α_2 and α_1/α_2 can be recovered from the probability density function $f_2(x)$ and compensation function $w(x)$, which have already been identified and estimated from observables (x_n, w_n). The next proposition summarizes this result.

Proposition 1. *Under assumption (3) and the fact $E[g(x)] = 1$, if the risk aversion parameter ρ is known, the model primitives $g(x)$, α_0/α_2 and α_1/α_2 are identified using (22), (23) and (24), respectively.*

IDENTIFICATION OF ρ

Most previous literature uses data from laboratory experiments (See, e.g., Holt and Laury (2002, 2005) and Harrison et al. (2007)), field experiments (See, e.g., Andersen et al. (2008) and Dohmen et al. (2010)), and individual behavior in actual markets (See, e.g., Chetty (2006) and Cohen and Einav (2007)) to identify risk preferences. Since I do not access to any type of above data on managers, I follow Gayle and Miller (2009) and Gayle and Miller (2012) to study the identification of risk preferences parameter by exploiting the participation constraint (3).

Rearranging the participation constraint (3) with equality yields

$$\alpha_0/\alpha_2 = \{E[w(x)^{1-\rho}]\}^{\frac{1}{1-\rho}}, \tag{25}$$

which shows that the participation constraint equation is satisfied by different combinations of ρ and α_0/α_2. The right side of equation (25) is the generalized mean of wage, which decreases with ρ [22]. As ρ increases, managers become more risk averse and the generalized mean of wage $w(x)$ declines, but this is just offset by nonpecuniary benefits from outside options versus working. Consequently a researcher with cross sectional data on managerial compensation can not distinguish between a sample of managers with high risk aversion and more nonpecuniary benefits from work versus a sample of managers with low risk aversion and less nonpecuniary benefits from work.

To identify ρ, I suppose there is data on at least two states, that is data from two sectors where the nonpecuniary benefits from outside options versus are the same. The risk aversion parameter ρ can be identified by using the participation constraints in the two states. Formally, let $w_\kappa(x)$ and $w_\tau(x)$ denote managerial compensation schedules for states κ and τ, $f_{\kappa 2}(x)$ and $f_{\tau 2}(x)$ denote the probability density functions of abnormal returns under working in the two states. because the participation contraints hold in both states, we could

[22] This can be showed by using Jensen Inequality.

in principle solve the following equation

$$\int w_\kappa(x)^{1-\rho} f_{\kappa 2}(x)dx = \int w_\tau(x)^{1-\rho} f_{\tau 2}(x)dx \tag{26}$$

in ρ. If there is at least one solution in above equation, ρ is identified.

Intuitively, a person's risk preference cannot be identified from playing one single lottery if there are unobserved benefits from not playing the lottery. However, if he plays two lotteries with different risk characteristics but the same unobserved benefits from not playing the lotteries, his risk preferences are partially revealed by the pecuniary compensating differential between them, which equals his expected utility from playing one versus the other.

IDENTIFICATION OF QUANTILE FUNCTIONS $S(n)$ AND $T(n)$

The remaining model primitives to be identified are the quantile functions of firm size and managerial talent, $S(n)$ and $T(n)$. The basic ideas to identify them exploit the differential equations for effective managerial wage (15) and firm's profit (17) derived from the optimal assignment problem. Using the two differential equations, the relative quantile functions of firm effective size and managerial talent can then be expressed as functions of effective wage and effective profit. The next lemma formalizes this result.

Lemma 3. *The relative quantile functions of firm effective size and managerial talent can be written as*

$$\frac{\tilde{S}(r)}{\tilde{S}(0)} = \exp\left(\int_0^r \frac{\pi'(i)}{v(i)+\pi(i)} di\right), \tag{27}$$

$$\frac{T(r)}{T(0)} = \exp\left(\int_0^r \frac{v'(i)}{v(i)+\pi(i)} di\right). \tag{28}$$

Similarly as in Tervio (2008), the indeterminacy of the lowest talent $T(0)$ and smallest effective size $\tilde{S}(0)$ is because there is no information to infer the relative contributions of managerial talent and firm size to the effective output created at and below the smallest firm in the sample.

If the effective wage and effective profits for each firm are known, the relative quantile functions of firm effective size and managerial talent can be identified directly by using equations (27) and (28). However, these effective terms cannot be observed directly from the data. They are recovered by using two conditions. First, each firm's total expected output equals to the sum of the expected managerial wage $E[w(x)]$ and expected firm profits, which is measured by using firms' market value V_n. Since firms' market value reflects also people's expection on firms' market in the future, the economic value of current CEOs' talent is the aggregate value current CEOs generate for firms in current and future periods. Second, the each firm's effective output equals to the sum of the effective managerial wage and effective firm profits. Using these two conditions, we have $\pi(r) = e^{-\chi_n} V_n$ and $v(r) = e^{-\chi_n} E[w_n(x)]$,[23] where $e^{-\chi_n}$ is given by (9).

With the recovered effective wage $v(r)$ and effective profits $\pi(r)$ in hand, the relative quantile functions $\frac{\tilde{S}(r)}{\tilde{S}(0)}$ and $\frac{T(r)}{T(0)}$ are identified by using equations (27) and (28) in Lemma (3). Using the identified relative quantile function of firm effective size, the relative quantile function of firm actual size $\frac{S(n)}{S(0)}$ can then be identified by using the definition of effective size, $\tilde{S}(r) \equiv S(n)e^{-\chi_n}$.

3.2 Estimation

In view of the identification results, I propose a three-step procedure to estimate the model primitives using the data on managerial compensation w, firms' abnormal returns x and market value V, and the observed firms' characteristics. The observed firms' characteristics are used to introduce firms' heterogeneity in the model primitives capturing firms' risk and disutility. In the first step, I show the estimation of the probability density function under working, $f_{n2}(x)$, from the observed firms' abnormal returns x. Second, I present the estimation of the probability density function of abnormal returns for shirking $f_{n1}(x)$, risk aversion parameter ρ, nonpecuniary benefits of outside options versus working, α_{n0}/α_{n2}, and nonpecuniary benefits of shirking versus diligent work α_{n1}/α_{n2}. They are estimated by using a Minimum Distance Estimator (MDE) from observed managerial compensation w and the observed firms' characteristics. Finally, The estimation of quantile functions of firm's size and managerial talent, $S(n)$ and $T(n)$, can be achieved by directly following the identification of them.

ESTIMATION OF $f_{n2}(x)$

In principal, the probability density function of abnormal returns for working $f_{n2}(x)$ could be estimated nonparametrically using the data on x. However, firms' risk and disutility are heterogeneous. Recall that firms' risk is reflected by the variance of probability density functions of abnormal returns. To capture firm heterogeneity in firms' risk, I exploit the observed firms' characteristics as the controlling covariates in the estimation of $f_{n2}(x)$. Since it is intractable to undertake nonparameteric estimation with many covariates, I will adopt a parametric estimation method to estimate $f_{n2}(x)$.

[23] Since total expected output equals to the sum of the expected managerial wage and expected firm profits, I have $S(n)T(r) = E[w_n(x)]+V_n$. Multiplying $e^{-\chi_n}$ on both sides of above equation yields $\tilde{S}(r)T(r) = e^{-\chi_n}E[w_n(x)] + e^{-\chi_n}V_n$. On the other side, the fact that effective output consists of effective wage and effective profits gives $\tilde{S}(r)T(r) = v(r) + \pi(r)$. Comparing these two equations gives $\pi(r) = e^{-\chi_n}V_n$ and $v(r) = e^{-\chi_n}E[w_n(x)]$.

More specifically, I follow Gayle and Miller (2009) to assume that the probability density functions of abnormal returns under shirking ($e = 1$) and working ($e = 2$) are both truncated normal with support bounded below by ψ,

$$f_{ne}(x) = [\Phi(\frac{\mu_{ne} - \psi}{\sigma_n})\sigma_n\sqrt{2\pi}]^{-1} \exp[\frac{-(x - \mu_{ne})^2}{2\sigma_n^2}], \tag{29}$$

where Φ is the standard normal distribution function and (μ_{ne}, σ_n^2) denotes the mean and variance of the corresponding parent normal distributions for firm n. In this specification, I assume that probability density functions of abnormal returns for all firms have the same functional form, but different values of mean and variance. The density functions under shirking and working have different means, but share the same variance.

My model restricts that the expected abnormal returns conditional on manager working are zeros. Moreover, in the data I fail to reject that the mean of abnormal returns is zero. Thus I will use the restriction in the estimation of $f_{n2}(x)$. Using the truncated normal specification of $f_{n2}(x)$, the implicit function for μ_{n2} is given by

$$0 = E(x|e = 2) = \mu_{n2} + \frac{\sigma_n\varphi[(\mu_{n2} - \psi)/\sigma_n]}{\Phi[(\mu_{n2} - \psi)/\sigma_n)}, \tag{30}$$

where φ denotes standard normal probability density function.

To introduce firms' heterogeneity, the mean and variance (μ_{n2}, σ_n^2) are specified as functions of the observed firms' characteristics, including number of employees, debt-to-equity ratio and sector dummies. Denoting above observed firms' covariates by z_{n1}, varance σ_n^2 is then specified as the following exponential function,

$$\sigma_n^2 = \exp(\beta' z_{n1}),$$

where β is a parameter vector for the firms' observed covariates. μ_{n2} will be also a function of β, defined by (30). The estimation of $f_{n2}(x)$ is completed by estimating (ψ, β). ψ is consistently estimated by using the lowest value of abnormal return x in the data. β can be estimated by using a Maximum Likelihood Estimator (MLE) through the probability density function of x under working. The maximum likelihood estimator $\hat{\beta}$ is then found by choosing β to minimize the following negative sum of the log-likelihood functions,

$$L_N(\beta) = \sum_{n=1}^{N} \times \{\ln \sigma_n(\beta) + \ln \Phi[\frac{\mu_{n2} - \psi}{\sigma_n(\beta)}] + \frac{[x_n - \mu_{2n}]^2}{2\sigma_n(\beta_2)^2}\}, \tag{31}$$

subject to the restriction that the expected value of abnormal returns is zero when managers work (30).

ESTIMATION OF $f_{n1}(x)$, α_{n0}/α_{n2}, α_{n1}/α_{n2}, ρ

Under the truncated normal distribution specification, the parameters characterize the probability density function of abnormal returns under shirking $f_{n1}(x)$ are the mean μ_{n1} and variance σ_n^2 of its parent distribution. Recall that $f_{n1}(x)$ is assumed to share the same variance as $f_{n2}(x)$. The estimation of $f_{n1}(x)$ will be completed if the mean μ_{n1} is estimated. To capature firm heterogeneity, similarly I specify the mean μ_{n1} as a linear function of the observed firms' covariates,

$$\mu_{n1} = u_1' z_{n1}.$$

Here I use the same observed firm covariates as in the specification of variance σ_n^2 because μ_{n2} is also an implicit function of the covariates.

Nonpecuniary benefits from outside options versus working α_{n0}/α_{n2} are determined by demand for management service and managerial satisfication on working for the firm. Thus it is ideally to specify α_{n0}/α_{n2} as functions of both firm's and managerial characteristics. Since I do not have the information on managerial characteristics, in this study I specify α_{n0}/α_{n2} and α_{n1}/α_{n2} only as functions of firm's characteristics. It is plausible in the sense that heterogeneous firms are matched with heterogeneous managers endogenously in equilibrium. In particular, I specify it as a linear function of firms' characteristics,

$$\alpha_{n0}/\alpha_{n2} = a_0' z_{n2},$$

where z_{n2} is a vector of firm's characteristics, including firms' assets and number of employees. Nonpecuniary benefits from shirking versus working α_{n1}/α_{n2} reflect disutility that the firm bring to its working manager. Thus, I also specify it as a linear function of firms' characteristics,

$$\alpha_{n1}/\alpha_{n2} = a_1' z_{n2}.$$

The estimation task in this step now becomes to estimate the parameters set $\Omega \equiv (u_1, a_0, a_1, \rho)$. It is estimated by using a Minimum Distance Estimator (MDE) through exploiting the optimal wage equation (5), participation constraint (3) and incentive compatibility constraint (4). Denote the true value of Ω by Ω_0. Let $\overline{w}_n$ denote the observed wage of firm n's manager,

$$\overline{w}_n = w_n^\star(\Omega_0, x), \quad n = 1, 2, \cdots, N.$$

Then the parameter set Ω can be estimated by choosing Ω to minimize the distance of observed wage and model generated wage. Equivalently I estimate Ω by choosing Ω to minimize the distance of log observed wage and log model generated wage, which is given by

$$\sum_{i=1}^{N}[\ln(\overline{w}_i) - \ln w_i^{\star}(\Omega, x)]^2.$$

Using the optimal wage equation and the specifications of the parameters, Ω is estimated by

$$\hat{\Omega} = \underset{\Omega}{\text{argmin}} \sum_{i=1}^{N}[\ln(\overline{w}_i) - \ln(a_0' z_{n2}) + \frac{1}{\rho}\ln\{[\theta_0 + \theta_1(1 - \frac{f_{n1}(x, u_1)}{f_{n2}(x)}(a_1' z_{n2})^{1-\rho})]\}]^2$$

subject to the equation systems (6) and (7) determined by the participation constraint (3) and incentive compatibility constraint (4). Note that θ_0 and θ_1 are the solution of a fixed-point problem, which must be solved for each value of the parameter vector Ω, to evaluate the econometric criterion function. This is an example of a nested fixed-point algorithm, first proposed by (Rust (1987)) in the empirical industrial organization literature. In the literature on managerial compensation, Ferrall and Shearer (1999), Margiotta and Miller (2001) and Gayle and Miller (2009) also use a nested fixed-point algorithm to obtain their estimates.

ESTIMATION OF $S(n)$ AND $T(n)$

The estimation of quantile functions of firms' size $S(n)$ and managerial talent $T(n)$ directly follows the identification. As noted in the identification section, I first need estimate the firm effective profit and managerial effective wage. The effective profit $\pi(r)$ can be estimated from $\hat{\pi}(r) = V_n e^{-\hat{\chi}_n}$, where V_n is the observed firm market value and $\hat{\chi}_n$ is the estimated value of χ_n by using the estimated variables,

$$\hat{\chi}_n = \ln E\{[\hat{\theta}_0 + \hat{\theta}_1(1 - (\hat{a}_1' z_{n2})^{1-\hat{\rho}}\hat{g}_n(x)]^{1/\hat{\rho}}\}.$$

The effective wage $v(r)$ is estimated directly from $\hat{v}(r) = \hat{a}_0' z_{n2}$, where the equality is from the definition of effective wage $v(r) \equiv E[w_n(x)]e^{-\chi_n} = \alpha_{n0}/\alpha_{n2}$.

As noted by Tervio (2008), the prerequisite of the assignment models to make sense is that the incomes that firms and managers obtain need to exhibit perfect positive rank correlation. In this study firms' effective profit and managerial effective wage need exhibit perfect positive rank correlation. However, since in practice managerial wage is affected by many factors, including some stochastic factors, we could not expect that the effective wage and effective profit are perfect rank correlated. Thus the noisy relation of managerial effective wage and effective profit needs to be smoothed into a strictly monotonic relation.

The smooth can be done in many ways. Here I follow Tervio (2008) to perform a Lowess Smoothing of the relation of the levels of managerial effective wage and firm effective profit. The Lowess Smoothing is first proposed by Cleveland (1979). The basic idea of the method is to take a weighted moving average of effective wage along the rank by firms' effective profit, using higher weights for nearby observations [24]. Hereafter I use the smoothed effective wage to refer to the actual effective wage. Since the rank of effective profit is used to order the observations, there is no need to smooth it. I only need to do a simple connect-the-dots interpolation to create a continuous distribution for it.

Using the effective profit and smoothed effective wage, the relative quantile functions of firm effective size $\frac{\tilde{S}(r)}{\tilde{S}(0)}$ and managerial talent $\frac{T(r)}{T(0)}$ can be estimated by exploiting Lemma (3). More explicitly, they are estimated by

$$\frac{\tilde{S}(r)}{\tilde{S}(0)} = \exp(\int_0^r \frac{\hat{\pi}'(i)}{\hat{v}(i)+\hat{\pi}(i)} di),$$
$$\frac{T(r)}{T(0)} = \exp(\int_0^r \frac{\hat{v}'(i)}{\hat{v}(i)+\hat{\pi}(i)} di),$$

Finally using the estimated relative quantile function of firm effective size, and estimated $\hat{\chi}_n$, the quantile function of firm actual size $S(n)$ can be estimated by using the definition, $\tilde{S}(r) \equiv S(n)e^{-\chi_n}$.

4 Data

My sample is comprised of the 1000 largest publicly traded firms in market value and their CEOs in the S&P Compustat database for 2011. Data on executive compensation is collected from the S&P Compustat Execucomp database. I extract only the information on the compensation of chief executive officers for this study. The compensation data are supplemented by firm information from the S&P Compustat North America database and monthly stock price data from the Center for Research in Security Prices (CRSP) database. The firm characteristics data are exploited to introduce heterogeneity in firms' risk and disutility. The monthly stock prices data are used to construct the abnormal returns of firms.

Industrial level factors may affect risk of firms and thus mangarial compensation. To consider the these industrial level effects but not complicate the analysis too much, I follow Gayle and Miller (2012) to specify firms in the sample as three industrial sectors according

[24] In principle, I could also smooth firms' effective profit according the rank of effective wage. However, since managerial wage is more volatile, the firms' market value tends to be better.

to GICS code. The first is called primary sector, including firms in energy (GICS: 1010), materials (1510), industrials (2010, 2020, 2030), and utility (5510). Sector 2, called consumer goods, comprises firms from consumer discretionary (2510, 2520, 2530, 2540, 2550) and consumer staples (3010, 3020, 3030). Finally firms in health care (3510, 3520), financial service (4010, 4020, 4030, 4040), information technology and telecommunication services (410, 4520, 4030, 4040, 5010) comprise Sector 3, called services.

4.1 CEOs Compensation

I measure a CEO's compensation as the the sum of his salary and bonus, the value of restricted stocks and options granted, and the value of retirement and long-term compensation schemes. It is the costs to shareholders of employing a CEO and the total compensation a CEO obtains associated with employment. I use this approach to measure CEO compensation for two reasons. First, I consider a static model. When firms are matching with managers, firms care about the costs of employing a manager and managers care about total compensation associated with employment. Second, the CEO compensation measured by this approach is alway positive, which satisfies the model restriction.

Table (2) summarizes the cross-sectional information on components of CEO compensation by sectors in our data. The total compensation is broken out into four components: salary and bonus, the value of options granted, the value of restricted stock granted, and other compensation, where other compensation includes the value of retirement and nonequity incentive compensation. Salary, bonus and other compensation account for about 44.4% of the total compensation while other three components collectively account for about 55.6% of the total compensation. It shows that a large fraction of managerial compensation is linked to firm performance. Moreover, managerial income from holding granted financial securities has very large standard deviation, which suggests that managerial income from holding granted financial securities whose value is affected by the firm's performance accounts for most variability of total compensation.

4.2 Abnormal Returns

I follow Gayle and Miller (2009) to define the abnormal returns x of a typical firm as the residual component of returns that its manager is able to control. In the optimal contract, the compensation should depend on this residual in order to provide the manager appropriate incentives, but it should not depend on changes in stochastic factors that originate outside the firm and are not able to be controlled by the manager. More specifically, following Gayle and Miller (2009), I impute x, the abnormal returns to the firm, using the monthly stock price data on the 1000 largest companies from 1998 to 2011 in two steps. First, I calculate the difference between the financial return on the individual firm stock and the return on the market portofolio. Second, I then regress this difference on a sector-specific constant and the time-varying factors, including GDP.

Table (3) displays the summary description on the residual for the sample in 2011. All the estimated coefficients in the regression used to measure the abnormal returns are proven to be significant. The table shows that the means of abnormal returns are all negative in all three sectors, The mean is highest in services sector and lowest in primary sector. The dispersion of the abnormal returns is highest in consumer sector and lowest in primary sector.

4.3 Firm Characteristics

Firm characteristics affect firm risk, the nature of its manager's responsibilities and the satisfaction he derives from managing the firm. These characteristics are also relevant to the nonpecuniary benefits of managers from pursuing his own interests within firms. Table (4) summarizes the cross sectional information on firms charactersitics by sectors. It gives summary statistics on assets, market value, sales and employees. These characteristics give us some idea on the scope of managerial responsibilities. It also shows summary statistics on debt equity ratio, which reflects firms' risk to some extent. The firms in the consumer sector are most highly leveraged while those in primary are the least leveraged. The fact suggests that averagely firms in consumer sectors may be rikier, which is consist with the dispersion of abnormal returns shown in Table (3).

The prerequisite of the study to make sense is that the compensation increases with firm size and abnormal returns. Figure (3) displays the relation of CEO compensation and firm rank by market value in 2011. The sample correlations between CEO compensation, firm characteristics and abnormal return are displayed in Table (5). The correlation between market value and CEO compensation is 0.5251, which is the largest. It suggests that market value has the most explanation power on CEO compensation. The correlation between CEO compensation and abnormal return is also positive.

5 Estimation Results

The estimation approach is applied to the above data on CEO compensation, firms' abnormal returns, market value and other observed characteristics. In view of the estimation, I first estimate the means and variances of probability density functions under manager working, $f_{n2}(x)$. The estimates for the variables in the specification of variance capture how firms' heterogeneous risk depends on the observed firm covariates. I then estimate the means of probability density functions under manager shirking, $f_{n1}(x)$, nonpecuniary benefits from outside options versus working, α_{n0}/α_{n2}, nonpecuniary benefits from shirking versus working, α_{n1}/α_{n2}, and the relative risk aversion parameter, ρ. Finally I estimate the quantile functions of firm size $S(n)$ and managerial talent $T(n)$.

With the model primitive estimates in hand, I then conduct counterfactuals to assess the importance of managerial talent misallocation as a result of moral hazard on the aggregate production of firms. More specifically, I conduct counterfactuals to quantify the four measures of loss generated from moral hazard presented in Section (2), including the loss from risk-sharing inefficiency, the actual loss from talent misallocation, the maximal loss from talent misallocation, and the loss from ignoring moral hazard problem. Quantifying each of the losses requires me to conduct a counterfactual. In the section I first report the estimates of the model primitives. I then report the counterfactual results on the losses related to moral hazard.

5.1 Estimated Model Primitives

The probability density function under working $f_{n2}(x)$ is estimated in the first step. Recall $f_{n2}(x)$ is parameterized to be a truncated normal distribution. The estimation is completed by estimating its lower bound, the mean and variance of its parent distribution function. On the top of Table (6), I report the estimates for the variables in the specification of variance of its parent normal distribution. The estimates convey information on risk of firms with different observed characteristics. The estimate on debt equity ratio is positive, which suggests that the more leveraged firms measured by debt equity ratio are riskier holding other factors constant. Number of employees has a negative effect on firms' risk, which means firm with more employees are less risky holding other factors constant. The firms in the services sector tend to have the highest risk while those in primary sectors have the lowest. The mean for the parent distribution of $f_{n2}(x)$ is estimated by using the restriction $E(x|e = 2) = 0$, which implies it is not greater than zero for all firms. On the bottom of Table (6), it shows that consistent estimate of truncation lower bound is $\psi = -0.744$, which is the lowest abnormal return in the data.

The proability density function under shirking $f_{n1}(x)$ shares the same variance as $f_{n2}(x)$. This leave its mean μ_1 to be estimated. The parameter estimates for the variables in its specification of μ_1 are reported in the middle of Table (6). The abnormal returns for the firms with higher debt equity ratio and more employees tend to have lower means if their managers pursue their own interests. This implies that managers have more impacts on these firms. The estimates on dummy variables indicating consumer and services sectors is negative. Managers have more impacts on firms in the consumer and services sectors. This might be because firms in those sectors face more competition.

The top of Table (7) reports the estimation results for the nonpecuniary benefits from outside options versus working, α_{n0}/α_{n2}. The results show that managers serving for firms with more assets and employees obtain more nonpecuniary benefits from their outside options versus working. The reason is that those firms endogenously end up with more talented managers, whose reservation utility are higher. The middle of Table (7) reports that the estimation results for the nonpecuniary benefits from shirking versus working, α_{n1}/α_{n2}. These results show that managers serving for firms with more assets or employees obtain more nonpecuniary benefits from shirking than those from working. It might be because that managing these firms needs more responsibilities and gives less satisfaction to managers. Moreover, the estimates of α_{n1}/α_{n2} for all the firms are greater than 1, which is consistent with our model restriction. The risk aversion parameter for all managers is estimated to be 1.4565. With such a risk aversion level, a manager having \$2 millions is willing to pay \$0.382 millions to avoid a gamble that he has equal probability losing \$1 millions and winning \$1 millions.

The remaining model primitives are the distributions of firms' size and managerial talent. The relative quantile function of firms' effective size and managerial talent, $\frac{\tilde{S}(r)}{\tilde{S}(0)}$ and $\frac{T(r)}{T(0)}$, are estimated according to Lemma (3). Using the estimated $\frac{\tilde{S}(r)}{\tilde{S}(0)}$, the relative quantile function of firms' actual size $\frac{S(n)}{S(0)}$ is estimated by exploiting the definition of effective size, $\tilde{S}(r) \equiv S(n)e^{-\chi_n}$. Figures (4) and (5) display the estimated distribution of firms' effective size and managerial talent, respectively. The distribution of firms' effective size is highly skewed to right and there is no much difference in managerial talent. This is consistent with the finding by Tervio (2008) and Jung and Subramanian (2013) that most variation on CEO compensation is explained by differential of firm size.

Figure (6) displays actual, efficient, and least efficient allocation of CEOs to firms. The grey 45° dot represents the efficient allocation of CEOs to firms. The blue dot represents the actual allocation of CEOs to firms. From the comparison between the efficient and actual allocations, we can see that many firms that end up with much less talented CEOs in equilibrium than they should in the efficient allocation. The reason is that these firms are very risky and pay higher risk premium to CEOs. Figure 6 displays the risk premium that firms need pay by their actual size. This graph is consistent with that Figure (6). The blue green line represents the worst allocation of CEOs to firms. The counterfactuals quantify the losses of all firms from actual and worst allocation of CEOs to firms by comparing the difference of output under these allocation between the output under efficient allocation.

5.2 Counterfactuals

I conduct four counterfactuals to assess the importance of managerial talent misallocation as a result of moral hazard using the estimated model primitives. The importance is evaluated by quantifying the four measures of loss presented in Section (2.6). Quantifying each of the four measures requires a counterfactual. The remaining of this section presents the counterfactual details and discusses the results.

LOSS FROM RISK-SHARING INEFFICIENCY

The first counterfactual concerns on quantifying the loss of firms from risk-sharing inefficiency. The loss is measured by the difference between the expected wage the firm pays and flat wage it would pay if moral hazard is not a problem, defined by L_{n1} in (19). The first row of Table (8) gives the summary statistics on this loss, The total estimated aggregate loss over all firms is 2.9 billion dollars. This value is plausible if I compare it with the literature on this measure. In Edmans and Gabaix (2011a), calibrating a different theoretical model, this loss is calibrated to be $2 billion over the top 500 firms in 2005 Execucomp database. Assuming absolute risk aversion perferences of managers, Gayle and Miller (2009) estimates the aggregate costs from risk sharing inefficiency are about 15.39 billion 2000 year US dollars over 3026 firms from 1992 to 2004.

ACTUAL LOSS FROM MATCHING INEFFICIENCY

In the second counterfactual, I quantify the loss to firms from talent misallocation. The loss is measured by the difference between the expected output of firms from efficient matching with that from actual matching, which is defined by L_{n2} in (20). I first calculate the expected output from a counterfactual in which the matching between managers and firms is efficient, which is given by

$$E[S(n)T(n)(1+x)] = S(n)T(n).$$

From the estimated $\frac{\tilde{S}(r)}{\tilde{S}(0)}$ and χ_n, I can recover $\frac{S(n)}{\tilde{S}(0)} = \frac{\tilde{S}(r)e^{\chi_n}}{\tilde{S}(0)}$ by exploiting the definition effective size. I can then derive $\frac{S(n)T(n)}{\tilde{S}(0)T(0)} \equiv \Pi_n$, which gives $S(n)T(n) = \Pi_n \cdot [\tilde{S}(0)T(0)] = \Pi_n \cdot [\hat{v}(0) + \hat{\pi}(0)]$, where $\hat{v}(0) + \hat{\pi}(0)$ is the estimated total effective output for the lowest firm-CEO pair.

The total production loss from actual inefficient matching for the firm n is the firm's total output in the data, which is given by

$$E[S(n)T(r)(1+x)] = S(n)T(r) = E[w_n(x)] + V_n,$$

where $E[w_n(x)]$ is the expected wage the firm's manager obtain and V_n is the firm's market value. The difference between the total expected output from efficient matching and that from actual inefficient matching is the production loss from talent misallocation. The second row of Table (8) displays the summary statistics on the production loss for all firms. The total production loss is estimated to be about 12.64 billion dollars from talent misallocation. To the best of my knowledge, there is only one other paper which calibrates the loss from matching inefficiency due to moral hazard. Edmans and Gabaix (2011a) provide a calibrated upper bound ($7.7 billion) for this loss for the top 500 firms in 2005 Execucomp database.

MAXIMUM LOSS FROM MATCHING INEFFICIENCY

Third, I conduct a counterfactual to quantify the maximal loss that firms in the market would incur from managerial talent misallocation due to moral hazard. This loss is measured by the difference between the total expected output from the efficient matching and that from the worst matching, that is the best firm matches with the least talented manager. The summary statistics on this loss is displayed on the third row of Table (8). The total maximal loss that firms could incur from talent misallocation could reach to $1,815.17$ billion dollars, which is approximately 10.25% of the market capitalization of the 1000 largest firms.

LOSS FROM IGNORING MORAL HAZARD

Finally, I conduct a counterfactual to quantify the loss that all firms would incur from ignoring moral hazard from. It is measured by the difference between the expected output from managers pursuing the firms' interests versus that from managers pursuing their own interests. The summary statistics on this loss is displayed on the bottom of Table (8). The total loss that firms would incur from ignoring moral hazard is about $2,382.49$ billion dollars, which is about 13.45% of the market capitalization of the 1000 largest firms. This is the benefit of all firms from motivating their CEOs. This is consistent with the finding in Gayle and Miller (2009). The estimated total loss of firms from ignoring moral hazard in their paper is about 673.97 billion dollars for for 3026 firms from 1992 to 2004. This value accounts for 19.32% of the total market capitalization of those 3026 firms.

DISCUSSION

The above counterfactuals show that the aggregate loss firms incur as a result of talent misallocation is more than four times as large as the loss due to the standard risk-sharing inefficiency when moral hazard is present. Most previous studies on moral hazard focuses only on the loss of firms from risk-sharing inefficiency, which is commonly considered as the agency costs of moral hazard for firms. The results suggest that the studies on agency costs of moral hazard may severely underestimate efficiency loss if focusing only on risk-sharing inefficiency and ignoring matching inefficiency. The sum of loss from both risk-sharing and matching inefficiecny is the aggregate costs that firms would incur by using optimal contracting to solve moral hazard problem. Corporate governance, such as boarding monitoring, is considered as a substitute mechanism to reduce costs of moral hazard. Understanding the aggregate costs associated with moral hazard gives us some guidance on implementing corporate governance.

On the other side, the sum of loss from both risk-sharing and matching inefficiencies is the total costs all firms would incur by using optimal contracting to solve the moral hazard problem. The total costs of both risk-sharing and matching inefficiencies associated with moral hazard is very small compared with the substantial benefits from motivating the managers to pursue the interests of shareholders. This suggests that aligning the managers to pursue the objective of shareholders instead of their own is extremely beneficial to firms.

6 Conclusion

This paper quantified the magnitude of inefficiency in the equilibrium allocation of CEOs to firms. I developed an estimable model which illustrate the presence of moral hazard could lead to an inefficient allocation of CEOs to firms in equilibrium. An new empirical method was proposed to estimate the model primitives in a matching framework with asymmetric information. The method was applied to estimate the model using data the U.S. market for CEOs in 2011. Using the estimates, I quantified the magnitude of inefficiency in the equilibrium allocation of CEOs to firms caused by moral hazard. I found that the inefficiency is more than *four time* as large as the inefficiency loss from risk-sharing due to moral hazard. The findings suggest that the studies focusing solely on risk-sharing can severely underestimate the inefficiency loss and more work should consider the allocation inefficiency caused by moral hazard.

The methodology developed in this paper has several potential extensions. The first extension would be to consider incorporating a more realistic dynamic contracting problem into the matching model of CEOs and firms, where the dynamic contracts offered to the CEOs consist of a sequence of wage and effort. Second, the methodology developed could also apply to other markets with moral hazard, including labor markets in which employers hire employees whose actions are not perfectly observed, and capital markets in which venture capitalists invest in entrepreneurial companies whose management cannot be perfectly monitored.

References

S. Andersen, G. W. Harrison, M. I. Lau, and E. E. Rutström. Eliciting risk and time preferences. *Econometrica*, 76(3):583–618, 2008.

P. Bajari and J. T. Fox. Measuring the efficiency of an fcc spectrum auction. Technical report, National Bureau of Economic Research, 2005.

G. S. Becker. A theory of marriage: Part i. *The Journal of Political Economy*, pages 813–846, 1973.

R. Chetty. A new method of estimating risk aversion. *The American Economic Review*, 96 (5):1821–1834, 2006.

P. A. Chiappori and B. Salanié. Testing contract theory: A survey of some recent work. Technical report, Econometric Society Monographs, Eighth World Congress, 2003.

P. A. Chiappori, B. Salanié, and Y. Weiss. Partner choice and the marital college premium. 2011.

E. Choo and A. Siow. Who marries whom and why. *Journal of Political Economy*, 114(1): 175–201, 2006.

W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368):829–836, 1979.

A. Cohen and L. Einav. Estimating risk preferences from deductible choice. *The American Economic Review*, pages 745–788, 2007.

T. Dohmen, A. Falk, D. Huffman, and U. Sunde. Are risk aversion and impatience related to cognitive ability? *The American Economic Review*, 100(3):1238–1260, 2010.

A. Edmans and X. Gabaix. The effect of risk on the ceo market. *Review of Financial Studies*, 24(8):2822–2863, 2011a.

A. Edmans and X. Gabaix. Tractability in incentive contracting. *Review of Financial Studies*, 24(9):2865–2894, 2011b.

A. Edmans, X. Gabaix, and A. Landier. A multiplicative model of optimal ceo incentives in market equilibrium. *Review of Financial Studies*, 22(12):4881–4917, 2009.

C. Ferrall and B. Shearer. Incentives and transactions costs within the firm: estimating an agency model using payroll records. *The Review of Economic Studies*, 66(2):309–338, 1999.

J. T. Fox. Estimating matching games with transfers. Technical report, National Bureau of Economic Research, 2008.

X. Gabaix and A. Landier. Why has ceo pay increased so much? *The Quarterly Journal of Economics*, 123(1):49–100, 2008.

A. Galichon and B. Salanié. Matching with trade-offs: revealed preferences over competing characteristics. 2010.

G.L. Gayle and R.A. Miller. Has moral hazard become a more important factor in managerial compensation? *The American Economic Review*, 99(5):1740–1769, 2009.

G.L. Gayle and R.A. Miller. Identifying and testing generalized moral hazard models of managerial compensation. *Unpublished Manuscript, Carnegie Mellon University*, 2012.

S.J. Grossman and O.D. Hart. An analysis of the principal-agent problem. *Econometrica: Journal of the Econometric Society*, pages 7–45, 1983.

B. J. Hall and J. B. Liebman. Are ceos really paid like bureaucrats? *The Quarterly Journal of Economics*, 113(3):653–691, 1998.

G. W. Harrison, J. A. List, and C. Towe. Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion. *Econometrica*, 75(2):433–458, 2007.

B. Hölmstrom. Moral hazard and observability. *The Bell Journal of Economics*, pages 74–91, 1979.

B. Hölmstrom. Moral hazard in teams. *The Bell Journal of Economics*, pages 324–340, 1982.

B. Hölmstrom and P. Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328, 1987.

C. A. Holt and S. K. Laury. Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655, 2002.

C. A. Holt and S. K. Laury. Risk aversion and incentive effects: New data without order effects. *The American economic review*, 95(3):902–904, 2005.

H. W. H. Jung and A. Subramanian. Ceo talent, ceo compensation and product market competition. In *Allied Social Science Association Meeting*, 2013.

T.C. Koopmans and M. Beckmann. Assignment problems and the location of economic activities. *Econometrica: Journal of the Econometric Society*, pages 53–76, 1957.

M.M. Margiotta and R. A. Miller. Managerial compensation and the cost of moral hazard. *International Economic Review*, 41(3):669–719, 2001.

J. A. Mirrlees. The optimal structure of incentives and authority within an organization. *The Bell Journal of Economics*, pages 105–131, 1976.

K. Murphy and M. Jensen. Performance pay and top management incentives. 1990.

Y. H. Pan. The determinants and impact of executive-firm matches. *Unpublished working paper. University of Utah*, 2012.

S. Rosen. Authority, control, and the distribution of earnings. *The Bell Journal of Economics*, pages 311–323, 1982.

J. Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.

M. Sattinger. Differential rents and the distribution of earnings. *Oxford Economic Papers*, 31(1):60–71, 1979.

M. Sattinger. Assignment models of the distribution of earnings. *Journal of economic literature*, 31(2):831–880, 1993.

M. Tervio. The difference that ceos make: An assignment model approach. *The American Economic Review*, 98(3):642–668, 2008.

J. Tinbergen. Some remarks on the distribution of labour incomes. *International Economic Papers*, 1:195–207, 1951.

J. Tinbergen. On the theory of income distribution. *Weltwirtschaftliches Archiv*, pages 155–175, 1956.

Appendix A: Proofs

Proof of Lemma 1:

Proof. First define $\nu_n(x) \equiv [\frac{\alpha_{n2}}{\alpha_{n0}} w_n(x)]^{1-\rho}$, and then the participation constraint (3) can be rewritten as

$$E[\nu_n(x)] \geq 1. \tag{32}$$

Similarly the incentive compatibility constraint (4) for working can be rewritten as

$$E[\nu_n(x)] \geq (\alpha_{n1}/\alpha_{n2})^{(1-\rho)} E[\nu_n(x) g_n(x)]. \tag{33}$$

Consequently, minimizing expected compensation subject to (3) and (4) is equivalent to minimizing $E[\nu_n(x)]^{1/(1-\rho)}$ subject to (32) and (33). To solve the optimal problem, I choose $\nu_n(x)$ to maximize the following Lagrangian,

$$-E[\nu_n(x)]^{1/(1-\rho)} + \theta_0 E[\nu_n(x) - 1] + \theta_1 E[\nu_n(x) - (\alpha_{n1}/\alpha_{n2})^{1-\rho} \nu_n(x) g_n(x)],$$

where θ_0 and θ_1 are Lagrangian Multipliers for participation and incentive compatibility constraints. The first order condition is then given by

$$\nu_n(x)^{\rho/(1-\rho)} = \theta_0 + \theta_1 [1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x)]. \tag{34}$$

From the definition of $\nu_n(x)$, we know $w_n(x) = (\alpha_{n0}/\alpha_{n2}) \nu_n(x)^{1/(1-\rho)}$. Substituting it back into the first order condition, we obtain

$$w_n^\star(x) = (\alpha_{n0}/\alpha_{n2})[1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x)]^{1/\rho}.$$

which is optimal compensation equation (5).

The two equations determining the Lagrangian Multipliers θ_0 and θ_1 in Lemma 1 are derived by exploiting the participation and incentive compatibility constraints. First multiplying $\nu_n(x)$ and taking expectations on both sides of (34) yields

$$E[\nu_n(x)^{1/(1-\rho)}] = \theta_0 E[\nu_n(x)], \tag{35}$$

which implies $\theta_0 > 0$ and the participation constraint (32) holds with equality. Solving (34) for $\nu_n(x)$ and substituting it into binding (32) gives (6). The incentive compatibility constraint also holds with equality by a contradition. If we set $\theta_1 = 0$ in (34), we will obtain $\nu_n(x)^{\rho/(1-\rho)}$ is fixed and then we have a fixed optimal wage. This contradicts with that the optimal compensation contract should be tied with x. Similarly solving (34) for $\nu_n(x)$ and substituting it into binding (33) gives (7). □

Derivation of the Expected Utility:

From Lemma 1, the optimal compensation of firm n's manager is given by

$$w_n^\star(x) = (\alpha_{n0}/\alpha_{n2}) \{\theta_0 + \theta_1 [1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x)]\}^{1/\rho}. \tag{36}$$

Taking the log on both sides of optimal equation (36) yields

$$\ln[w_n^\star(x)] = \ln(\alpha_{n0}/\alpha_{n2}) + \frac{1}{\rho}\ln\{\theta_0 + \theta_1[1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x)]\}. \tag{37}$$

Taking the expectation on both sides of optimal equation (36) gives

$$E[w_n^\star(x)] = (\alpha_{n0}/\alpha_{n2})E\{[\theta_0 + \theta_1(1 - (\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x))]^{1/\rho}\}. \tag{38}$$

The expected utility of firm n's manager is written as

$$\begin{aligned} EU_n &= \frac{\alpha_{n2}^{1-\rho}}{1-\rho} E\{[w^\star(x)]^{1-\rho}\} \\ &= \frac{\alpha_{n2}^{1-\rho}}{1-\rho} E\{e^{(1-\rho)ln[w^\star(x)]}\} \\ &= \frac{\alpha_{n2}^{1-\rho}}{1-\rho} E\{e^{(1-\rho)[\ln(\alpha_{n0}/\alpha_{n2}) + \frac{1}{\rho}\ln\{\theta_0+\theta_1[1-(\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x)]\}]}\} \\ &= \frac{\alpha_{n2}^{1-\rho}}{1-\rho} e^{(1-\rho)\ln(\alpha_{n0}/\alpha_{n2})} E\{e^{\frac{1-\rho}{\rho}\ln\{\theta_0+\theta_1[1-(\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x)]\}}\}. \end{aligned}$$

Letting $\chi_n = \ln\{E\{[\theta_0+\theta_1(1-(\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x))]^{1/\rho}\}\}$, and using equation (6), the expected utility can be rewritten as

$$\begin{aligned} EU_n &= \frac{\alpha_{n2}^{1-\rho}}{1-\rho} e^{(1-\rho)[ln(\alpha_{n0}/\alpha_{n2})+\chi_n-\chi_n]} \\ &= \frac{\alpha_{n2}^{1-\rho}}{1-\rho} e^{(1-\rho)\{\ln[(\alpha_{n0}/\alpha_{n2})E\{[\theta_0+\theta_1(1-(\alpha_{n1}/\alpha_{n2})^{1-\rho} g_n(x))]^{1/\rho}\}]-\chi_n\}} \\ &= \frac{\alpha_{n2}^{1-\rho}}{1-\rho} e^{(1-\rho)((\ln\{E[w_n^\star(x)]\}-\chi_n)} \\ &= \frac{\{\alpha_{n2}E[w_n^\star(x)]e^{-\chi_n}\}^{1-\rho}}{1-\rho}. \end{aligned}$$

Interpretation of the Sorting Conditions:

Consider there are two firms and two managers in the market for CEOs. Assume that firm 1 has larger actual size than firm 2, $S_1 > S_2$, and manager 1 is more talented than manager 2, $T_1 > T_2$. If we observe that firm 1 matches with manager 1 and firm 2 matches with manager 2, we must have that the profits firm 1 obtains from hiring manager 1 are not less than those from hiring manager 2 and the similar condition holds for firm 2. Formally, the following conditions are written as,

$$\begin{aligned} S_1T_1 - E(w_1^1) \geq S_1T_2 - E(w_2^1), \\ S_2T_2 - E(w_2^2) \geq S_2T_1 - E(w_1^2), \end{aligned}$$

where $E(w_m^n)$ is the expected wage firm n would pay to managers m. Using the definition of effective wage, we have $E(w_m^n) = v_m e^{\chi_n}$. Thus above two inequalities can be rewritten as follows,

$$S_1T_1 - v_1e^{\chi_1} \geq S_1T_2 - v_2e^{\chi_1},$$
$$S_2T_2 - v_2e^{\chi_2} \geq S_2T_1 - v_1e^{\chi_2}.$$

Rearranging them yields

$$S_1e^{-\chi_1}(T_1 - T_2) \geq v_1 - v_2,$$
$$v_1 - v_2 \geq S_2e^{-\chi_2}(T_1 - T_2).$$

Since $T_1 > T_2$, we have

$$S_1e^{-\chi_1}(T_1 - T_2) \geq v_1 - v_2 \geq S_2e^{-\chi_2}(T_1 - T_2),$$

which implies that firm 1's effective size $S_1e^{-\chi_1}$ should be no less than firm 2's effective size $S_2e^{-\chi_2}$. We also know that in equilibrium firm 1 matches with manager 1 and firm 2 matches with manager 2. Thus the firm with larger effective size matches with the more talented manager in equilibrium. The matching is perfect sorting by firm's effective size and managerial talent.

Proof of Lemma 2:

Proof. The equation (22), determining $g(x)$, is derived by using the optimal wage equation (5), assumption (3), and the fact $E[g(x)] = 1$. Suppressing the firm index n, the optimal wage equation (5) can be rewritten as

$$w(x) = (\alpha_0/\alpha_2)\{\theta_0 + \theta_1[1 - (\alpha_1/\alpha_2)^{1-\rho}g(x)]\}^{\frac{1}{\rho}}.$$

Defining $\lambda(x) \equiv [\frac{w(x)}{\alpha_0/\alpha_2}]^{\rho}$, above equation is written as

$$\lambda(x) = \theta_0 + \theta_1[1 - (\alpha_1/\alpha_2)^{1-\rho}g(x)]. \tag{39}$$

From assumption (A3), we know that $\lim_{x\to\infty} g(x) = 0$. Taking limit on both sides of (39) yields

$$\overline{\lambda} \equiv \lim_{x\to\infty} \lambda(x) = \theta_0 + \theta_1. \tag{40}$$

Note the fact that $E[g(x)] = 0$. Taking expectation on both sides of (39) gives

$$\underline{\lambda} \equiv E[\lambda(x)] = \theta_0 + \theta_1[1 - (\alpha_1/\alpha_2)^{1-\rho}]. \tag{41}$$

Substracting both sides of equation (40) to equation (39) gives us

$$\overline{\lambda} - \lambda(x) = \theta_1(\alpha_1/\alpha_2)^{1-\rho}g(x). \tag{42}$$

Similarly substracting both sides of equation (40) to equation (41) gives us

$$\overline{\lambda} - \underline{\lambda} = \theta_1(\alpha_1/\alpha_2)^{1-\rho}. \tag{43}$$

Dividing equation (42) over equation (43) yields

$$g(x) = \frac{\overline{\lambda} - \lambda(x)}{\overline{\lambda} - \underline{\lambda}}. \tag{44}$$

On the other hand, from the definition $\lambda(x) \equiv [\frac{w(x)}{\alpha_0/\alpha_2}]^\rho$, we know that $\overline{\lambda} = [\frac{\overline{w}}{\alpha_0/\alpha_2}]^\rho$ and $\underline{\lambda} = E\{[\frac{w(x)}{\alpha_0/\alpha_2}]^\rho\}$, where $\overline{w}$ is the maximum wage. Substituting them into equation (44) gives us the equation (22).

The equation (23), determining the nonpecuniary benefit of outside options versus workingα_0/α_1, can be derived directly from the participation contraint (3) with equality. Rearranging it gives

$$\alpha_0/\alpha_2 = \{E[w(x)^{1-\rho}]\}^{\frac{1}{1-\rho}},$$

which is exactly the equation (23). To derive (24), rearranging equation (43) and using equation (42) give us

$$\alpha_1/\alpha_2 = \Big\{\frac{\overline{\lambda} - \underline{\lambda}}{\theta_1}\Big\}^{\frac{1}{1-\rho}} = \Big\{\frac{\overline{\lambda} - \underline{\lambda}}{\overline{\lambda} - \theta_0}\Big\}^{\frac{1}{1-\rho}}. \tag{45}$$

From equation (35) in the proof of lemma (1), we have

$$\theta_0 = \frac{E[w(x)]}{E[w(x)^{1-\rho}](\alpha_0/\alpha_2)^\rho}$$

Plugging it into (45) gives equation (24). □

Proof of Lemma 3:

Proof. From the fact that each firm's effective output is divided by the manager and the firm, I know that the firm's effective output equals to the manager's effective wage and the firm's effective profits. I also know that the firm whose effective size has quantile r will match with the manager whose talent has quantile r, which gives

$$v(r) + \pi(r) = \tilde{S}(r)T(r). \tag{46}$$

Recall that the slope of managerial effective wage is

$$v'(r) = \tilde{S}(r)T'(r), \tag{47}$$

and the slope of firm's effective profit is

$$\pi'(r) = \tilde{S}'(r)T(r). \tag{48}$$

Dividing (47) to (46) , I have

$$\frac{T'(r)}{T(r)} = \frac{v'(r)}{v(r) + \pi(r)}. \tag{49}$$

Intergrating it over quantile 0 to any quantile i yields (28). Similarly dividing (48) to (46) , I have

$$\frac{\tilde{S}'(r)}{\tilde{S}(r)} = \frac{\pi'(r)}{v(r) + \pi(r)}. \tag{50}$$

Intergrating it over quantile 0 to any quantile i yields (27). □

Appendix B: Tables and Graphs

Table 1: Numerical example 2: matching pattern 1

	0%-20%	20%-40%	40%-60%	60%-80%	80%-100%
0%-20%	0.14	0.06	0	0	0
20%-40%	0.04	0.08	0.08	0	0
40%-60 %	0	0.02	0.03	0.11	0.04
60%-80%	0.01	0.02	0.04	0.04	0.09
80%-100%	0.01	0.02	0.05	0.05	0.07

Table 2: Cross-sectional information on components of compensation by sectors
(In millions of US $ (2011); standard deviations in parentheses)

Variable	Primary	Consumer	Services	All
Observations	358	225	417	1000
Salary and bonus	1.12	1.37	1.02	1.16
	(0.92)	(1.49)	(0.70)	(1.01)
Value of options granted	1.25	1.64	1.20	1.32
	(1.77)	(3.43)	(1.85)	(2.28)
Value of restricted	2.49	2.75	3.60	3.01
Stock granted	(2.50)	(3.90)	(18.56)	(12.22)
Other compensation	2.32	3.09	1.88	2.31
	(3.27)	(4.16)	(2.08)	(2.73)
Total compensation	7.23	8.85	7.71	7.80
	(5.17)	(7.70)	(18.86)	(13.10)

Table 3: Summary description on abnormal returns in 2011

Variable	Sector	Mean	Std	Min	Max
Abnormal returns	All	-0.0542	0.2450	-0.7438	1.7412
	Primary	-0.0748	0.2297	-0.7381	0.8797
	Consumer	-0.0538	0.2631	-0.7366	1.1840
	Services	-0.0367	0.2468	-0.7438	1.7412

Table 4: Cross-sectional information on firm characteristics by sectors
(Assets, market value and sales in billions of US $ (2011),
employees in thansands, standard deviations in parentheses)

Variable	Primary	Consumer	Services	All
Assets	15.45	12.80	33.63	22.44
	(45.27)	(27.11)	(171.04)	(114.76)
Market Value	10.89	13.18	12.56	12.10
	(28.25)	(26.70)	(32.65)	(29.83)
Sales	10.85	9.60	5.52	6.94
	(21.38)	(26.86)	(14.69)	(17.56)
Employee	20.54	58.57	18.69	28.32
	(39.33)	(161.89)	(44.20)	(86.67)
Debt equity Ratio	2.30	5.65	2.87	3.29
	(3.92)	(39.72)	(7.66)	(19.63)

Table 5: Sample correlations

CEO compensation	1.0000					
Market value	0.5251	1.0000				
Assets	0.1099	0.3755	1.0000			
Sales	0.2730	0.7397	0.3900	1.0000		
Employee	0.1414	0.4476	0.2568	0.6698	1.0000	
Abnormal returns	0.0379	0.0879	-0.0693	-0.0202	-0.0396	1.0000

Table 6: Parameter estimates of the returns distributions

Parameter	Description	Variable	Estimate	Standard Error
σ^2	Percent Variance	Constant	-0.4316	1.5622
	of Return	Debt equity ratio	0.0010	0.1246
		Log Employees	-0.7149	0.3782
		Consumer dummy	0.5804	0.6932
		Services dummy	0.6954	0.9448
μ_1	Percent mean return	Constant	-0.0946	0.0098
	from shirking	Debt equity ratio	-0.0107	0.0017
		Log Employees	-0.0121	0.0006
		Consumer dummy	-0.1097	0.0065
		Services dummy	-0.1095	0.0154
ψ		Lower bound of return	-0.744	

Table 7: Nonpecuniary benefits relative to diligence

Parameter	Description	Variable	Estimate	Standard Error
α_{n0}/α_{n2}	Nonpecuniary benefits from outside option	Constant	1.1223	0.1135
		Log Assets	0.3908	0.2890
		Log Employees	0.5787	0.0694
α_{n1}/α_{n2}	Nonpecuniary benefits from shirking	Constant	0.4600	0.0797
		Log Assets	0.3983	0.0592
		Log Employee	0.4756	0.0897
ρ	Risk aversion parameter		1.4565	0.0440

Table 8: Losses of talent misallocation related to moral hazard (billions)

Losses	Mean	Standard error	Total
Actual loss from risk-sharing inefficiency	0.0029	0.0241	2.90
Actual loss from matching inefficiency	0.0126	0.2455	12.64
Maximal loss from matching inefficiency	1.815	6.0879	1815.17
Loss from ignoring moral hazard	2.382	8.056	2382.49

Figure 1: The numerical example: risk premium χ_n

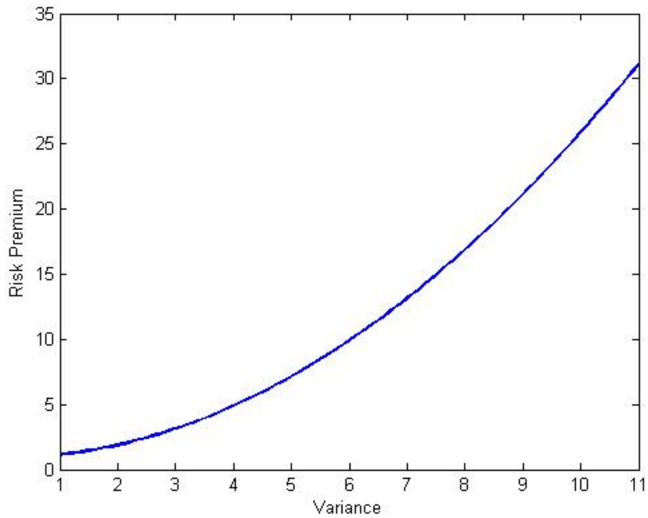

Figure 2: The numerical example 2: loss from mismatch

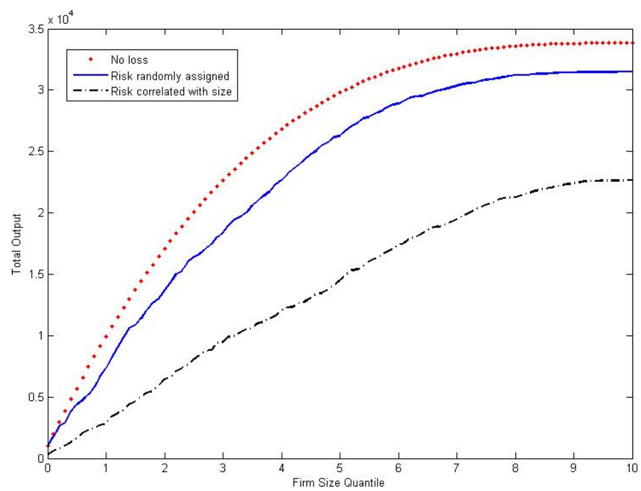

Figure 3: Relation of CEO compensation and firm rank by market value in 2011

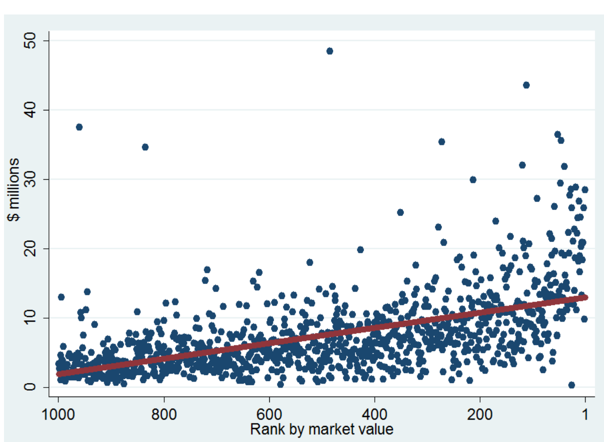

Figure 4: Estimated distribution density of firms' effective size $\frac{\tilde{S}(r)}{\tilde{S}(0)}$

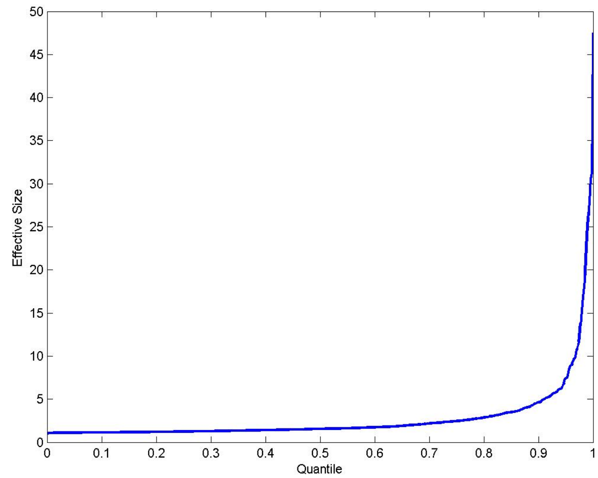

Figure 5: Estimated distribution density of managerial talent $\frac{T(r)}{T(0)}$

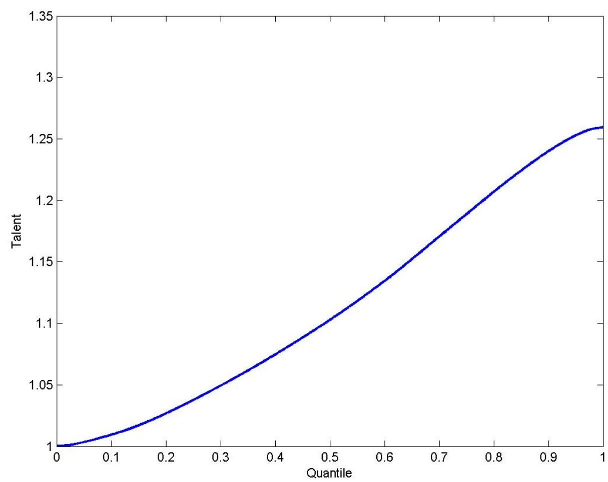

Figure 6: Relation of ranks by firm actual size and managerial talent

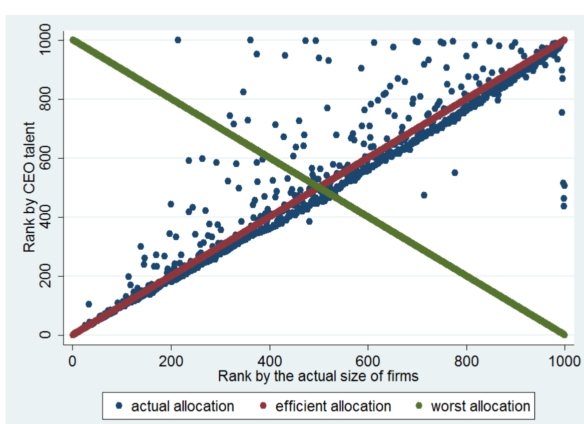

Figure 7: Relation of ranks by firm actual size and the risk premium χ_n