

NBER WORKING PAPER SERIES

RISK AND THE CEO MARKET:
WHY DO SOME LARGE FIRMS HIRE HIGHLY-PAID, LOW-TALENT CEOS?

Alex Edmans
Xavier Gabaix

Working Paper 15987
<http://www.nber.org/papers/w15987>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2010

We thank Chris Armstrong, Karthik Balakrishnan, Philip Bond, Ingolf Dittmann, Vincent Glode, Todd Gormley, Holger Mueller, Derek Neal, Greg Nini, Gustav Sigurdsson, Luke Taylor and seminar participants at Chicago, LBS, NYU, Penn State, Wharton and the Theory Conference on Corporate Finance and Financial Markets for valued input, and Qi Liu for research assistance. We are grateful to Ingolf Dittmann for generously sharing his data on CEO non-firm wealth. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2010 by Alex Edmans and Xavier Gabaix. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Risk and the CEO Market: Why Do Some Large Firms Hire Highly-Paid, Low-Talent CEOs?

Alex Edmans and Xavier Gabaix

NBER Working Paper No. 15987

May 2010

JEL No. G3,J01

ABSTRACT

This paper presents a market equilibrium model of CEO assignment, pay and incentives under risk aversion and heterogeneous moral hazard. Each of the three outcomes can be summarized by a single closed-form equation. In assignment models without moral hazard, allocation depends only on firm size and the equilibrium is efficient. Here, talent assignment is distorted by the agency problem as firms involving higher risk or disutility choose less talented CEOs. Such firms also pay higher salaries in the cross-section, but economy-wide increases in risk or the disutility of being a CEO (e.g. due to regulation) do not affect pay. The strength of incentives depends only on the disutility of effort and is independent of risk and risk aversion. If the CEO affects the volatility as well as mean of firm returns, incentives rise and are increasing in risk and risk aversion. We calibrate the efficiency losses from various forms of poor corporate governance, such as failures in monitoring and inefficiencies in CEO assignment. The losses from misallocation of talent are orders of magnitude higher than from inefficient risk-sharing.

Alex Edmans

The Wharton School

University of Pennsylvania

2433 Steinberg Hall-Dietrich Hall

3620 Locust Walk

Philadelphia, PA 19104

aedmans@wharton.upenn.edu

Xavier Gabaix

New York University

Finance Department

Stern School of Business

44 West 4th Street, 9th floor

New York, NY 10012

and NBER

xgabaix@stern.nyu.edu

1 Introduction

This paper presents a market equilibrium model of CEO assignment, pay and incentives. Risk-averse managers of different talents are hired in a competitive market by heterogeneous firms, which vary in their size, risk and level of effort required. The level of pay drives the assignment of talent to firms, and the strength of incentives induces optimal effort.

Despite the potential complexity caused by combining a talent assignment model with an agency problem under risk aversion, the equilibrium can be summarized by three simple, closed-form equations, one for each of assignment, pay and incentives. The model's tractability allows its economic forces to be transparent, yields clear empirical predictions for which factors do and do not matter for the three outcomes, and allows analysis of welfare consequences. Combining these three questions within a unifying framework generates a number of new implications unattainable from piecing together the results of individual models of each issue in isolation.

We start with assignment. As is standard, we model talent as affecting the maximum firm value that can be achieved in the absence of an agency problem. In a model without moral hazard, more talented CEOs work at larger firms to allow their talent to have greatest impact. We show that this allocation is distorted in the presence of an agency problem. A talented manager is a mixed blessing for two reasons. First, if utility is multiplicative in cash and effort, exerting effort is more costly to a talented and thus wealthy manager – for example, a day of leisure is particularly valuable to a rich CEO as he can enjoy his wealth in leisure time. Thus, the firm must pay a rich CEO a greater premium for disutility. Second, a manager who is already wealthy is less motivated by incentive pay and more willing to sacrifice it for leisure. The firm must therefore provide him with stronger incentives, which in turn requires paying a higher premium for risk. Thus, firms involving greater risk or disutility must pay particularly high premiums to hire talented managers, and so may prefer to appoint a “poor-and-hungry” CEO rather than a “rich-and-contented” alternative. Some talented managers are hired by small firms, where their talent has less effect, because such firms involve lower risk or disutility. Risk aversion thus not only leads to inefficient risk-sharing, but also distortions in real productive activity. If firms also differ in their sensitivity to talent, we have the natural additional prediction that talented managers are assigned to firms with high growth opportunities.

We obtain closed-form solutions for the losses due to inefficient risk-sharing and misallocation. The former depends on the average level of risk in the economy; somewhat less automatically, the latter depends on the cross-sectional variance of risk and not its mean. If risk is high but constant across firms, it has no effect on a CEO's choice of employer and so the assignment is not distorted. The losses from misallocating managers are also increasing in the dispersion of managerial ability, as is intuitive. More surprisingly, they are decreasing in the dispersion of firm size and the size elasticity of talent. When size is more dispersed, or talent has a particularly strong impact on large firms, size becomes more important than risk in determining the equilibrium matching. Thus, assignment becomes closer to the efficient positive

assortative matching on size. The sum of both inefficiencies is a measure of the losses from the failure by boards to control moral hazard through direct monitoring – even if it can be fully solved by contracts, such contracts create distortions. While it is well-known that incentive pay causes inefficient risk-sharing, we show that in a market equilibrium it also distorts real production. Thus, direct monitoring and incentives are not perfect substitutes as governance mechanisms. These losses are moderate: while the allocation is first-best inefficient compared to a world with perfect monitoring, it is second-best efficient given the existence of a moral hazard problem – a social planner would not be able to improve on the allocation. By contrast, if board failures instead lead to CEOs being randomly assigned to firms, the losses are greater and now increasing in the dispersion of firm size and the size elasticity of talent. Our model thus allows analysis of the losses from various manifestations of poor corporate governance.

Turning to the expected level of pay, it is increasing in firm size as in a pure assignment model. The addition of an agency problem means that pay also depends on a firm’s disutility and risk, as the CEO requires additional salary as compensation. Thus, firms with high risk or disutility not only hire less talented CEOs, but also pay their CEOs highly (relative to their skill level) as compensation. Cross-sectionally, riskier firms pay more as found by Garen (1994); greater disutility of effort has the same effect. Gayle and Miller (2009) show theoretically and empirically that, along the cross-section, firms which are more complex to manage or have greater agency problems (and thus stronger required incentives) pay their executives more.

However, what matters is not the absolute level of these parameters, but their magnitudes compared to other firms in the economy. Thus, aggregate changes in risk or the disutility of being a CEO (e.g. due to regulation, stronger board monitoring or activist shareholders) do not affect pay – while working for one’s current firm becomes less attractive, so do the outside options. This conclusion differs from the partial equilibrium model of Hermalin (2005), which argues that the recent strengthening in corporate governance increases the level of effort the CEO must exert and the risk of dismissal, and thus may explain the rise in pay over time. We show that in a market equilibrium, such economy-wide changes have no effect. Indeed, Peters and Wagner (2009) find that the effect on pay of dismissal risk is around eight times as high along the cross section as over the time series. The dependence of pay on economy-wide factors also highlights the importance of controlling for aggregate conditions (or at least time trends) in empirical analyses of the determinants of pay.

Third, the strength of incentives is measured by the percentage change in CEO pay for a percentage firm return. Thus, when considering the contract in terms of the dollar change in pay for a percentage firm return, the strength of incentives reflects the convexity of the contract. Hence, a single parameter controls both the slope of the contract (in percent-percent terms) and its convexity (in dollar-percent terms). This parameter depends only on the disutility of effort and is independent of both risk and risk aversion.

The above core model is presented in Section 2. In Section 3, we extend the model to allow the CEO’s actions to affect firm risk as well as the average return – for example, undertak-

ing a risky, positive-NPV project augments both the mean and volatility. While diversified shareholders do not care about idiosyncratic risk, a risk-averse CEO has private incentives to inefficiently forgo such a project. Therefore, if the CEO is at least a risk-averse as a log utility agent, the optimal contract becomes more convex to give the CEO a benefit from risk to offset his risk aversion; since the incentive measure represents the convexity of the contract, this in turn involves stronger incentives. This result contrasts the argument that powerful incentives induce the CEO to take excessive risk, and thus if the CEO has control over risk, incentives should be weaker. Moreover, incentives are now *increasing* in risk and risk aversion, contrary to traditional models which assume exogenous risk and predict a negative relationship. When the CEO is more risk-averse or the firm is riskier, it is necessary to give him even more convexity (and thus more incentives) to induce him to undertake a risky project. Indeed, Demsetz and Lehn (1985), Core and Guay (1999) and Oyer and Schaefer (2004) find a positive relationship between incentives and risk. For the same reason, incentives are increasing in the marginal increase in risk caused by value-enhancing actions. If the CEO mainly affects firm value by consuming perks, these actions have low effect on risk and so incentives are little changed, but if the CEO creates value by choosing risky projects, incentives must rise. The link between incentives and the effect of value-enhancing actions on risk has both cross-sectional and time-series implications. Along the cross-section, “new economy” firms have little tangible capital and so enhancing firm value involves greater risk – for example, investing in R&D has a zero payoff if the R&D fails, whereas investing in an old economy plant has liquidation value in the downside case. Indeed, Ittner, Lambert and Larcker (2003) and Murphy (2003) find stronger incentives in new economy firms. Over time, as industries mature and competition intensifies due to globalization, “sure-fire” projects which generate value with little risk become scarce, and enhancing firm value increasingly requires taking on risky projects. This may account for the rise in incentives, and in particular options, over time (see, e.g., Jensen and Murphy (2004).)

Our final theoretical extension allows for an elastic supply of CEO talent. We introduce a second labor market involving non-CEO jobs (e.g. hedge funds, entrepreneurship, or consulting), which we call the non-corporate sector. This market provides both a secondary source from which corporate firms can hire, and an outside option for CEOs. An aggregate increase in the disutility of being a CEO (while holding constant the disutility of working in the non-corporate sector) now augments CEO pay, as firms must compensate CEOs to deter them from leaving to the non-corporate sector. Since the additional disutility is particularly costly for talented CEOs, corporate firms hire less skilled managers, reducing the value created by the corporate sector. The magnitude of the rise in pay, downgrade in talent and value loss are all increasing in the size of the non-corporate sector, as this represents the extent of CEOs’ outside options. The value loss is also increasing in the aggregate salary paid to all corporate CEOs. This is intuitive: in general, the economic importance of a distortion to a factor of production is proportional to its marginal product; for a CEO, this is his salary.

The two-sector model can also be used to analyze the effect of trends in a specific industry.

For example, one “sector” could represent the financial industry, and the second all alternative jobs for such CEOs. An increase in regulation of the financial industry (e.g. in response to the recent crisis) may cause talented CEOs to leave. Since the outside options for financial CEOs are extensive (hedge funds and private equity houses in addition to executive positions at non-financial corporations), the value loss to the financial industry may be substantial.

Finally, our model’s closed form solutions allow a calibration of the inefficiencies from various forms of poor corporate governance. Aggregating over the 500 largest firms in Execucomp, if monitoring failures mean that agency problems must be solved by contracting, we estimate losses at from inefficient risk-sharing at \$1.7 billion per year, and misallocation at \$7.4 billion; the latter is an upper bound. The total inefficiency of \$9 billion is approximately twice the aggregate CEO salary; however, it is moderate since assignment is second-best efficient and contracting is optimal. By contrast, if board failures manifest in random assignments of CEOs across firms while retaining optimal contracting, the losses are approximately \$16 billion per year as a *lower* bound. Naturally, all of these losses would be significantly higher when considering all top executives rather than just CEOs. While recent critics of governance focus on inefficiencies in contracting (see, e.g., Bebchuk and Fried (2004)), we show that losses from misallocation of talent can be substantial, even if contracting is perfect.

In addition to the above specific results, our paper makes two methodological contributions. One is solving an assignment problem where firms differ in the severity of moral hazard as well as size. In existing assignment models (e.g. Sattinger (1993), Gabaix and Landier (2008), Terviö (2008)), both firms and workers differ in a single dimension (size and talent, respectively) and thus can be unambiguously ranked.¹ This allows for a relatively simple solution to the assignment problem – positive assortative matching between the ranks. Assignment models are typically complex to solve if one or both sides vary across multiple dimensions, because this makes ranking difficult. We show that risk and disutility can be combined with size into a single dimension which we call “effective” size, which we can use to unambiguously rank firms and thus achieve a tractable solution to a multidimensional allocation problem.

A second methodological contribution is achieving a closed-form solution to a model in which the agent affects the volatility as well as mean of firm returns. Antecedents include Sung (1995) and Ou-Yang (2003), who use the Holmstrom and Milgrom (1987) framework that requires exponential utility, a financial cost of effort, continuous time and Gaussian noise, and Dittmann and Yu (2009) who assume separable preferences and Gaussian noise. We allow for general noise distributions and non-separable utility.

¹In Eisfeldt and Kuhnen (2009), workers (not firms) differ on multiple characteristics; the model specifies that productivity is a weighted average of these characteristics, thus effectively representing a single dimension. Antràs, Garicano and Rossi-Hansberg (2006) consider the allocation of workers to tasks, where both differ along a single dimension (skill and complexity, respectively). Kihlstrom and Laffont (1979) study the allocation of agents to jobs (either worker or entrepreneur) according to a single dimension, risk aversion. Galichon and Salanie (2009) do consider matching where both parties vary according to multiple dimensions, but require utility to be transferable across the matching parties and are unable to obtain closed-form solutions.

This paper is related to a number of models of executive compensation. Himmelberg and Hubbard (2000) is an early attempt to jointly model pay and incentives, but the level of pay is not an equilibrium and the absence of closed-form solutions renders drawing implications difficult. Gabaix and Landier (2008, “GL”) and Terviö (2008) present competitive assignment models of the managerial labor market, absent an agency problem. Edmans, Gabaix and Landier (2009), Axelson and Bond (2009), Baranchuk, Macdonald and Yang (2009) and Dicks (2009) add moral hazard but assume risk-neutrality and thus cannot investigate the effect of risk, risk aversion or risk-taking. Adding risk-aversion is typically a non-trivial extension which leads to very complex contracts. Holmstrom and Milgrom (1987) derive simple contracts under the assumption of exponential utility, a financial cost of effort, Gaussian noise and continuous time. As shown by Edmans et al., a multiplicative non-financial cost of effort is necessary to generate realistic income effects and empirically consistent scalings of incentives with firm size. We thus use the modeling setup of Edmans and Gabaix (2009, “EG”) which yields closed-form contracts without restrictions on the utility function or cost of effort, while retaining the clarity of discrete time. As a result, the equilibrium can be summarized by three closed-form equations. By embedding the EG contracting framework in a market equilibrium, we obtain many new results unattainable in either a partial equilibrium agency model, or a market equilibrium framework under risk neutrality – such as the effects of both cross-sectional and market-wide changes in risk and disutility on CEO assignment and pay, and a calibration of the losses from corporate governance imperfections. Tsuyuhara (2009) considers a market equilibrium with risk aversion, where both firms and workers are ex ante homogeneous. Plehn-Dujowich and Subrahmaniam (2009) allow for both heterogeneity and risk aversion, with output restricted to two possible levels. Acharya, Gabarro and Volpin (2010) extend the standard assignment model (where assignment depends only on firm size) to incorporate heterogeneity in corporate governance and show the allocation depends on governance as well as size.

2 The Model

2.1 Incentive Pay in Partial Equilibrium

We commence with a one-period model featuring a single firm and a single CEO (also referred to as the manager). This section is similar to EG; the main results come in Section 2.2 where we extend the model to a market equilibrium with multiple firms and CEOs. Appendix A provides proofs not given in the main text. The firm’s end-of-period stock price is given by

$$P_1 = se^{a-\bar{a}+\eta}/E[e^\eta] \tag{1}$$

where s represents baseline firm size and $a \in [\underline{a}, \bar{a}]$ is the CEO’s action (“effort”). The action a refers to any decision that improves the stock price but is costly to the manager, such as

exerting effort, forgoing private benefits, or choosing not to consume perks. Since there is a limit to the number of productive activities the agent can undertake to benefit the principal, we specify the firm’s end-of-period fundamental value as

$$V_1 = se^{\min(a,\bar{a})+\eta-\bar{a}}/E[e^\eta]. \quad (2)$$

\bar{a} is the maximum productive effort level. For example, \bar{a} reflects zero stealing in a cash flow diversion model, taking all positive-NPV projects (while rejecting negative-NPV ones) in a project selection model, or a limit to the number of hours a day the CEO can work while remaining productive in an effort model. Actions $a > \bar{a}$ do not benefit the principal but improve the stock price, such as manipulation. We allow for the maximum feasible action $\bar{\bar{a}}$ to exceed the maximum productive action \bar{a} purely for technical reasons – when \bar{a} is an interior action, the incentive compatibility (IC) constraint to implement \bar{a} becomes an equality, which substantially simplifies the proofs. (We conjecture that the results will continue to hold with $\bar{a} = \bar{\bar{a}}$.) Shareholders maximize expected fundamental value net of CEO pay. Appendix A proves that, if firm size s is sufficiently high, maximum productive effort \bar{a} is optimal for the firm because the benefits of effort (which are multiplicative in s) outweigh the costs (which are multiplicative in the CEO’s wage).

The variable η is mean-zero noise with standard deviation σ and bounded interval support. The normalization by $E[e^\eta]$ in (1) and (2) ensures that expected firm value does not depend on the noise distribution. The CEO privately observes η before choosing a . EG show that this assumption leads to closed-form contracts in discrete time, as well as consistency with the optimal contract in continuous-time, where noise and actions are simultaneous.² Note that the CEO remains exposed to risk, since he does not observe η until after signing the contract – as we will see, risk affects virtually all of our results.

On the equilibrium path where $a = \bar{a}$ is exerted, the initial stock price is $P_0 = e^{-\delta}E[P_1]$, where δ is the continuously compounded discount rate. Thus, the firm’s log stock return is:

$$r = \ln \frac{P_1}{P_0} = a + \eta + \mu, \quad (3)$$

with $\mu = \delta - \bar{a} - \ln E[e^\eta]$.

²This timing assumption is also featured in models in which the agent sees total output before deciding how much to divert (e.g. Lacker and Weinberg (1989), DeMarzo and Fishman (2007), Biais et al. (2007)), or observes the “state of nature” before choosing effort (e.g. Harris and Raviv (1979), Sappington (1983), Baker (1992), and Prendergast (2002)). As in most of these papers, to focus on a single source of imperfection (unobservability of effort) we abstract from commitment problems and assume that the CEO cannot quit after η is realized. Quits can be prevented by raising the fixed component of pay (see Appendix F of EG.)

The CEO has no pre-existing wealth, and his utility is given by:

$$U(c, a) = \frac{(ce^{-g(a)})^{1-\Gamma}}{1-\Gamma} \text{ for } \Gamma \neq 1 \quad (4)$$

$$= \ln c - g(a) \text{ for } \Gamma = 1.$$

c is the CEO's monetary compensation. $g(a)$ captures the disutility of effort and is increasing and convex; in Section 2.2 we allow the cost function $g(\cdot)$ to depend on the firm that the CEO is working for, i.e. it is a firm rather than CEO characteristic.³ $\Gamma \geq 0$ denotes relative risk aversion. The CEO's reservation utility is \underline{u} , which is exogenous in this section.

As in Edmans, Gabaix and Landier (2009), effort has a multiplicative effect on both CEO utility (equation (4)) and firm value (equation (2)). When effort has a percentage effect on firm value, the dollar benefits of working are higher for larger firms. Most CEO actions can be “rolled out” across the entire firm and thus have a greater effect in a larger company.⁴ Multiplicative preferences consider private benefits as a normal good, i.e. the utility they provide is increasing in consumption. This is consistent with the treatment of most goods and services in consumer theory; they are also commonly used in macroeconomics (see, e.g., Cooley and Prescott (1995)). This specification is also plausible under the literal interpretation of effort as forgoing leisure: a day of vacation is more valuable to a richer CEO as he has wealth to enjoy during it. Thus, the CEO's expenditure on leisure and private benefits rises in proportion to his wealth – just as with CRRA preferences, an investor's allocation to risky assets rises in proportion to his wealth. Indeed, it is multiplicative preferences which generate the CRRA utility function (4).⁵ Thus, just as CRRA is typically favored over CARA in asset pricing and macroeconomics because it leads to realistic income effects, the same considerations motivate the use of a multiplicative rather than financial cost of effort here. In addition, Edmans et al. show that multiplicative preferences and production functions are necessary to deliver empirically consistent predictions for the scaling of various incentive measures with firm size.

We take an optimal contracting approach that does not restrict the contract to specific functional forms.⁶ The optimal contract is a general function $c(r)$ that implements $a = \bar{a}$,

³More formally, the utility function is $\frac{(ce^{-G})^{1-\Gamma}}{1-\Gamma}$, where G is the disutility that working in the firm imposes on the CEO. Exerting effort a in firm n entails disutility $G = g_n(a)$, where g_n can be a function denoting the cost of effort from working in firm n .

⁴See Bennedsen, Perez-Gonzalez and Wolfenzon (2009) for empirical evidence that CEOs have the same percentage effect on firm value, regardless of firm size.

⁵Consider the general utility function $U(c, a) = \frac{e^{(1-\Gamma)(v(c)-g(a))}}{1-\Gamma}$. Our utility function (4) is a special case of this with $v(c) = \ln c$ (multiplicative preferences), which leads to CRRA. By contrast, Holmstrom and Milgrom (1987) assume that the cost of effort is financial, i.e. $v(c) = c$ and so the utility function becomes $e^{(1-\Gamma)(c-g(a))} / (1-\Gamma)$, which is CARA.

⁶Even though this is a hidden information model (the CEO learns η before choosing a), the optimal contract does not involve messages, as proven in EG. Intuitively, the reason is that the firm wishes to implement \bar{a} in all cases. Hence, on the equilibrium path, there is a one to one correspondence between the firm's return and the noise, which makes messages redundant.

satisfies the participation constraint $E[U] \geq \underline{u}$, and has the minimum cost $w = E[c]$ to the firm. From Theorem 1 of EG, the optimal contract is as follows:

Proposition 1 (*CEO pay in partial equilibrium*). *The optimal contract pays the CEO an amount c defined by:*

$$\ln c = \Lambda r + K, \tag{5}$$

where $\Lambda = g'(\bar{a})$ and K is a constant that makes the participation constraint bind ($E \left[(ce^{-g(\bar{a})})^{1-\Gamma} / (1-\Gamma) \right] = \underline{u}$).

Proof The full proof is in EG; a heuristic proof is in Appendix A. ■

The contract in Proposition 1 has a simple form. It is attainable in closed form, and its slope depends only on the cost of effort Λ , but not on risk σ nor risk aversion Γ – these only affect the scalar K . The sensitivity Λ represents the percentage change in pay c for a given return r . The contract can thus be implemented by giving the CEO Λw of stock and $(1-\Lambda)w$ of cash.⁷ When considering the contract in terms of the effect of firm returns on dollar pay, Λ reflects the convexity of the contract. Thus, changes in Λ affect both the sensitivity of the contract (in percent terms) and its convexity (in dollar terms).

We parameterize $\eta = \sigma\varepsilon$, where ε has unit variance, and define

$$\bar{\Gamma}(\sigma^2) = 2 \left(\ln E[e^{\sigma\varepsilon}] - \frac{1}{1-\Gamma} \ln E[e^{(1-\Gamma)\sigma\varepsilon}] \right).$$

If ε is a standard Gaussian, then

$$\bar{\Gamma}(\sigma^2) = \Gamma\sigma^2$$

while for any distribution with finite expectations, we have

$$\bar{\Gamma}(\sigma^2) \sim \Gamma\sigma^2$$

for $\sigma \rightarrow 0$. $\bar{\Gamma}(\Lambda^2\sigma^2)/2$ is the risk premium required by a CEO receiving the contract in Proposition 1, in the sense that $\bar{\Gamma}(\Lambda^2\sigma^2)/2 = \ln E[c] - \ln U^{-1}(E[U(c)])$ where $U(c) = \frac{c^{1-\Gamma}}{1-\Gamma}$. This interpretation motivates our notation $\bar{\Gamma}$.

2.2 Incentive Pay in Market Equilibrium

The simplicity of the contract in Proposition 1 allows it to be embedded into a market equilibrium where the expected wage w is endogenously determined. We use the equilibrium model of GL, which we summarize here. There is a continuum of firms of different size and managers with different talent. Firm $n \in [0, N]$ has size $S(n)$ and CEO $m \in [0, N]$ has talent $T(m)$. Low

⁷Since r is a continuously compounded return, the contract must be rebalanced continuously so that the percentage of stock remains constant at Λ .

n denotes a larger firm and low m a more talented CEO: $S'(n) < 0$, $T'(m) < 0$. The CEO's talent increases firm value according to:

$$s = S + CTS^\gamma, \quad (6)$$

where γ parameterizes the size elasticity of the impact of talent and C the productivity of talent, which we later allow to be heterogeneous across firms. Since talented CEOs are more valuable in larger firms, the n th most talented manager is matched with the n th largest firm to allow their talent to have greatest impact. The variable s considered in Section 2.1 thus refers to firm size gross of talent and S refers to net size; going forward, unless otherwise stated, the term “size” will refer to S .

GL assume a Pareto firm size distribution $S(n) = An^{-\alpha}$, and the following asymptotic value for the spacings of the talent distribution: $T'(n) = -Bn^{\beta-1}$. As in GL we consider the limit as $n/N \rightarrow 0$, i.e. the upper tail of the pay distribution. The equilibrium expected pay is:

$$w(n) = D(n_*) S(n_*)^{\beta/\alpha} S(n)^{\gamma-\beta/\alpha}, \quad (7)$$

where $S(n)$ is the size of firm n , n_* is the index of a reference firm (e.g. the median firm in the economy), $S(n_*)$ is the size of that reference firm, and $D(n_*) = -Cn_*T'(n_*) / (\alpha\gamma - \beta)$ is a constant. CEOs at large firms earn more as they are the most talented.

GL do not feature an agency problem and only specify the expected level of pay. We now incorporate the incentive model of Section 2.1 to determine the sensitivity of pay. We index the maximum effort level by \bar{a}_n to allow for heterogeneity in the level of effort required. Firms may also differ in their cost of effort, $g_n(a_n)$ – for example, a firm in a regulated industry or headquartered in an unattractive location is unpleasant to work for regardless of the effort a exerted by the CEO. The marginal cost of effort at the implemented effort level becomes $\Lambda_n = g'_n(\bar{a}_n)$. Risk may also vary and is indexed σ_n . We need not make any assumptions on how these parameters vary with n : since the contract implements $a = \bar{a}_n$, from (2), gross firm value remains at s as in the GL market equilibrium.

The expected utility of firm n 's CEO is given by:

$$U_n = \frac{(w_n e^{-\chi_n})^{1-\Gamma}}{1-\Gamma},$$

where

$$\chi_n = g_n(\bar{a}_n) + \frac{\bar{\Gamma}(\Lambda_n^2 \sigma_n^2)}{2} \quad (8)$$

denotes the “equivalent variation” (“EV”) associated with firm n , i.e. the loss suffered by the manager from disutility (the $g_n(\bar{a}_n)$ term) and risk (the $\bar{\Gamma}(\Lambda_n^2 \sigma_n^2)/2$ term). The latter arises because the CEO has a fraction Λ_n of his pay invested in the firm, and firm returns have

volatility σ_n . After adjusting for the EV, CEO n 's “effective” wage is

$$v_n = w_n e^{-\chi_n}. \quad (9)$$

Define $\bar{\chi}$ as the average of the firms' EVs:

$$e^{-\bar{\chi}} = E \left[e^{-\tilde{\chi}/(\alpha\gamma)} \right]^{\alpha\gamma}. \quad (10)$$

CEO assignment, pay and incentives in market equilibrium are given below:

Theorem 1 (*CEO pay in market equilibrium*). *Let n_* denote the index of a reference firm. In equilibrium, the manager of rank n runs a firm whose “effective size”*

$$\hat{S}_n = S_{n_*} e^{-\chi_n/\gamma} \quad (11)$$

is ranked n , and receives an expected pay:

$$w_n = D(n_*) S(n_*)^{\beta/\alpha} S_n^{\gamma-\beta/\alpha} \exp \left(\frac{\beta}{\alpha\gamma} (\chi_n - \bar{\chi}) \right), \quad (12)$$

where χ_n and $\bar{\chi}$ are defined by (8) and (10), $S(n_)$ is the size of the reference firm, and $D(n_*)$ is a constant independent of firm size. The actual pay c_n is given by:*

$$\ln c_n = \Lambda_n r_n + \ln w_n - \ln E \left[e^{\Lambda_n r_n} \right]. \quad (13)$$

Proof (Sketch). Assume that in market equilibrium, a CEO of talent $T(m)$ receives an effective wage (adjusted for effort and risk) of $v(m)$. If firm n wishes to hire manager m , it must pay him a effective wage $v(m)$ and thus a dollar wage $v(m) e^{\chi_n}$. It solves

$$\max_m E \left[(S(n) + CS(n)^\gamma T(m)) \frac{e^\eta}{E[e^\eta]} - v(m) e^{\chi_n} \right]$$

i.e.

$$\max_m C e^{-\chi_n} S(n)^\gamma T(m) - v(m). \quad (14)$$

Firm n behaves like a firm with “effective” size $(e^{-\chi_n})^{1/\gamma} S(n)$. Appendix A proves that it will pay the effective wage $v_n = D(n_*) (e^{-\bar{\chi}} S(n_*))^{\beta/\alpha} (e^{-\chi_n/\gamma} S)^{\gamma-\beta/\alpha}$. Taking into account the EV, the dollar wage is $w_n = v_n e^{\chi_n}$, which yields (12); (13) flows directly from Proposition 1. ■

Theorem 1 shows that CEO assignment, pay and incentives in competitive market equilibrium can be summarized by three simple closed-form equations, (11)-(13). This tractability allows for clear comparative statics. Starting with managerial assignment, in standard models, firms and CEOs each vary along a single dimension (size and talent, respectively). This allows for a relatively simple solution to the assignment problem – positive assortative matching,

where the CEO with the highest attribute is matched to the firm with the highest attribute. Assignment models are typically difficult to solve where there is firm heterogeneity along multiple dimensions, since it is unclear how to rank the firms and determine which is the “best” firm to be matched with the most talented CEO. The above proof sketch shows that risk σ_n and the marginal cost of effort Λ_n can be combined with size S_n into a single dimension, “effective” size $S_n e^{-\chi_n/\gamma}$, which can be unambiguously ranked and determines the equilibrium matching.

In assignment models without moral hazard, more talented managers are assigned to larger firms; this is efficient because talent has a greater impact in a bigger firm. We show that adding an agency problem distorts this efficient allocation. A firm with a higher cost of effort must pay a greater salary as compensation. Given multiplicative preferences, exerting effort is particularly costly for talented, highly-paid CEOs. For example, a day of vacation yields high utility to a rich CEO as he has income to spend during it. Therefore, the compensation for disutility is proportional to the CEO’s wage. The required compensation for risk is also proportional to the CEO’s wage. The incentive contract (13) pins down the fraction Λ_n of the CEO’s salary that must be paid in stock. CEOs that are already wealthy are less motivated by incentives, and thus must be given a greater dollar amount of stock to induce effort. Therefore, an increase in firm risk has a greater dollar effect on the variability of their pay, and requires the firm to pay them a higher dollar risk premium; indeed, Bandiera et al. (2010) find that managers with steeper contracts are paid more. Given CRRA, the required risk premium is a percentage of the wage. In sum, both disutility and risk force a firm to increase the salary of any manager that it hires by a given *proportional* amount, e^{χ_n} . Since this additional compensation is proportional to the CEO’s salary, it is higher for more talented managers and so skilled managers become relatively more expensive. Therefore, the firm chooses to hire a lower ability manager. Acharya, Gabarro and Volpin (2010) find that firms with weaker governance (i.e. lower disutility) employ high-talent managers.

In sum, managerial talent is a double-edged sword. While a talented manager has the potential to improve firm value to a greater degree, he is also more expensive to incentivize: since he already commands a high salary, he is willing to forgo incentive pay to enjoy leisure. Indeed, Malmendier and Tate (2009) find that winning awards (which may lead to an upward revision of the market’s perception of the CEO’s talent) leads to CEOs pursuing outside opportunities such as writing books and assuming board seats. This incentive problem is particularly severe if the firm involves high effort or risk. Thus, start-ups in particular may prefer to hire a “poor-and-hungry” CEO rather than a “rich-and-contented” alternative.⁸

Turning to expected pay, (12) shows that the wage depends not only on firm size S_n as

⁸Note that CEOs in our model have the same utility function, since it is not possible to solve an assignment model tractably when both sides exhibit heterogeneity on multiple dimensions. Thus, it is not that a “poor-and-hungry” CEO has a different cost of effort or risk aversion coefficient. CEOs differ only in their talent and thus reservation wage. Owing to multiplicative preferences, differences in the reservation wage translate into differences in the tendency to shirk, even though the utility function is not CEO-specific.

in GL, but also on how the firm’s cost of effort and risk (χ_n) compare to other firms in the economy ($\bar{\chi}$). Holding $\bar{\chi}$ constant, an increase in χ_n augments the wage as compensation for risk and disutility. Therefore, in the cross-section, firms with high EVs pay more. Indeed, Garen (1994) finds empirically that CEOs of riskier firms command higher pay. However, it is only the *relative* EV, $(\chi_n - \bar{\chi})$, that matters. Thus, disutility and risk only matter in the cross-section but not in the aggregate. If there was an economy-wide increase in risk or the disutility of being a CEO (e.g. due to regulation or activist shareholders), which increases the EV of all firms by the same absolute amount ξ , both χ_n and $\bar{\chi}$ increase by ξ ; $(\chi_n - \bar{\chi})$ and thus wages are unaffected – even though working for one’s present firm becomes less attractive, outside options also become less attractive.⁹ Regarding the own-firm prediction, Peters and Wagner (2009) investigate the link between CEO pay and the risk of firing. A one percentage point increase in firing probability augments pay by 4-8% along the cross-section, but only 0.2-1.3% over the time series. Peters (2009) studies the effect on pay of all sources of risk (changes in CEO wealth in addition to dismissal) and finds that it can explain the higher moments of the cross-sectional pay distribution. Regarding the cross-firm prediction, Acharya, Gabarro and Volpin (2010) find that a firm pays higher salaries if its competitors are worse governed (and thus more attractive to work for). More generally, the dependence of pay on the aggregate variable $\bar{\chi}$ highlights the importance of controlling for economy-wide variables such as average risk (or at least time trends) in empirical analyses of the determinants of pay.

The effect of changes in $(\chi_n - \bar{\chi})$ on expected pay is scaled by $\beta/\alpha\gamma$. A higher α raises the dispersion of firm sizes, and a higher γ augments the size elasticity of talent.¹⁰ Both factors increase the importance of size for CEO assignment and pay, and mean that variations in χ_n are relatively unimportant – as can be seen in (11), the effect of χ_n on “effective” size is decreasing in γ . Hence α and γ appear in the denominator of (12). By contrast, a higher β raises the dispersion of CEO talent. When talent is more variable, firms are more willing to pay the required compensation to attract a talented CEO (rather than “trading down” to the next best CEO), and so $(\chi_n - \bar{\chi})$ has a higher effect on the wage.

Moving to the strength of incentives, (13) shows that this depends only on Λ_n , the cost of effort, and is independent of risk and risk aversion. Hence, the Theorem shows which parameters do and do not matter for the different components of the contract. The cost of effort affects the strength of incentives, and increases the level of pay in the cross-section but not in the aggregate. Risk and risk aversion also augment the level of pay in the cross-section, but not in the aggregate. However, they have no effect on the strength of incentives. The familiar

⁹This prediction assumes that a CEO’s only outside option is to become a CEO of another firm. If CEOs can find a job outside of the CEO market, the more general prediction is that the cross-sectional elasticity of the wage to effort and risk is higher than the market-wide elasticity. See Section 3.2 for an extension to jobs outside the CEO market.

¹⁰In more mathematical terms, it is S^γ that matters for assignment, given equation (6). In turn, $S^\gamma(n) = An^{-\alpha\gamma}$: S^γ has a Pareto distribution with exponent $1/(\alpha\gamma)$. The higher $\alpha\gamma$ is, the more dispersed the distribution of elasticity-adjusted sizes S^γ .

trade-off between incentives and risk, which applies to rank-and-file employees, may not apply to CEOs. Since CEOs impact the entire firm, if the firm is sufficiently large, the benefits of effort are sufficiently strong that the firm implements maximum effort regardless of risk or risk aversion.

Overall, in assignment models without moral hazard, there is a positive correlation between managerial talent, firm size, and CEO pay.¹¹ In the presence of an agency problem and risk aversion, this relationship is mediated by other factors – a firm that involves high risk or disutility chooses to hire a less talented CEO, and pays a higher wage than his talent merits as compensation. Indeed, Nguyen and Nielsen (2010) find that, while there is generally a positive correlation between managerial ability and salary, a significant number of low-ability managers are well paid.

Theorem 1 can be extended to allow for firm heterogeneity not only in total disutility \bar{a}_n and risk σ_n , but also the impact of CEO talent. This extension is given in the following Remark.

Remark 1 (*Heterogeneous talent impact*). *Let the effect of talent on firm value (6) be given by:*

$$s_n = S_n + C_n T S_n^\gamma, \quad (15)$$

where C_n parameterizes the productivity of talent in firm n . In equilibrium, the manager of rank n runs a firm whose “effective size”

$$\widehat{S}_n = S_n C_n^{1-\gamma} e^{-\chi_n/\gamma}$$

is ranked n .

A firm with high C_n particularly benefits from a talented manager and thus has a higher effective size. For example, firms with high growth opportunities or in an unregulated industry have significant scope for a talented manager to add value. Since growing firms are also likely to be risky, and risk *reduces* the talent of the CEO hired from Theorem 1), empirical testing of this prediction will have to control for risk. Note also that talent impact C_n is a quite different concept from disutility $g_n(\bar{a}_n)$, and so the differential effects of these variables on the talent of the CEO hired are mutually consistent. $g_n(\bar{a}_n)$ reflects the total disutility the CEO must suffer when working for the firm, e.g. from regulation, being headquartered in an unfavorable location, or having to exert effort or forgo vacation days. These are inconveniences that are not mitigated by talent; in fact, they are particularly severe for talented managers owing to multiplicative preferences. By contrast, C_n reflects the impact that a talented CEO has on firm value if he exerts maximum effort – recall that gross firm value s only becomes (15) if $a_n = \bar{a}_n$,

¹¹Bandiera et al. (2010) find that more talented managers work for larger firms, and Chang et al. (2010) find that more talented managers are better-paid.

so it is C_n not \bar{a}_n that parameterizes the maximum potential value.¹² Thus, C_n reflects the potential for the manager to add value through exploiting growth opportunities, innovating or changing strategy. In sum, Remark 1 predicts that talented managers will be hired by firms with high growth potential, low risk and low disutility.

One might also think that talent might affect the CEO’s productivity of effort, in addition to its effect on maximum firm value as measured by T . Unfortunately, it is very difficult to solve tractably an assignment model in which both sides differ along multiple dimensions; while we are able to go beyond prior literature by allowing for firm heterogeneity across multiple dimensions, CEOs can only differ along a single dimension (the parameter T) and so we cannot introduce a separate manager-specific parameter for the productivity of effort. However, note that T already captures the productivity of effort to a degree: since $V_{1n} = (S_n + C_n T S_n^\gamma) e^{\min(a, \bar{a}) + \eta - \bar{a}} / E[e^\eta]$, the marginal effect of increasing a on firm value is increasing in T . Hence, our single source of manager heterogeneity does incorporate the realistic notion that a given level of effort by a talented manager is more productive – although exerting a given effort level is costlier to a talented manager owing to multiplicative preferences.¹³

We conclude this section by highlighting the features in the model that generate our results. First, the positive *qualitative* relationship between the CEO’s wage and the required compensation for disutility and risk, and thus distortions in allocation, can be generated by other utility functions and do not require multiplicative preferences. (See Appendix A for a proof.) Multiplicative preferences are only necessary to deliver the *quantitative* result that, as the CEO’s wage rises, his dollar stock holdings must increase in direct proportion. Thus, if Λ_n is constant across firms, the fraction of pay that is in stock is independent across firms of different size, as found empirically by Gibbons and Murphy (1992) and Murphy (1999). This empirical consistency is not a new result – Edmans, Gabaix and Landier (2009) already showed that multiplicative preferences are necessary to generate the size-independence of the stock fraction (albeit in a risk-neutral model) – instead, it provides the justification for using multiplicative preferences here. In turn, the direct proportionality (that results from multiplicative preferences as well as CRRA) leads to substantial tractability, as it means that many key variables scale with CEO pay. In particular, the required compensation for risk and disutility is proportional to the wage, so the effective wage is proportional to the actual wage. This is critical for the derivation of the “effective” size variable \hat{S}_n that allows a tractable solution to a multidimensional allocation problem (see the proof sketch of Theorem 1.)

Second, in standard models, the optimal effort level for the firm is a trade-off between the

¹²From equation (2), we have $E[V_1] = s e^{\min(a, \bar{a}) - \bar{a}}$ and so \bar{a} denotes the range of actions the CEO can take to destroy firm value (compared to the maximum effort benchmark) rather than create value. For example, \bar{a}_n is high in firms with free cash flow problems or weak governance.

¹³Even though effort has a higher dollar productivity for a talented manager, its percent productivity (i.e. the effect of effort on firm returns) is independent of talent, and so the incentive contract (13) is independent of T . Intuitively, since the market already knows that the manager is talented, the firm’s stock price is already high – thus, to increase the stock return, he has to work just as hard as an untalented manager.

costs (disutility plus the risk imposed by incentives) and benefits of effort, and is typically very difficult to solve (see, e.g., Grossman and Hart (1983)). Since CEOs can affect the entire firm value, the benefits of effort outweigh the costs and so the firm always implements maximum effort. This removes the need to analyze small trade-offs and leads to a simple optimal contract.

Third, the CEO observes the noise before taking his action. As shown in the heuristic proof and in EG, this leads to simple contracts such as (5). The intuition is that, since the CEO has observed η when taking his action, the IC condition ((33) in Appendix A) must hold state-by-state, i.e. for every possible realization of η . This tightly constrains the set of contracts available to the principal. If η was realized after a , the IC condition would only need to hold on average. Many contracts would satisfy the IC condition, and the problem becomes complex as the principal must solve for the cheapest contract out of this continuum.

2.3 Efficiency Analysis

2.3.1 Losses From Moral Hazard Under Optimal Contracting

In a pure assignment model, the efficient allocation involves positive assortative matching between talent and size. With an effort decision, the first-best allocation (that would occur if effort was observable) now involves assigning CEOs to firms based on their size and disutility. It may be efficient for a talented CEO not to work for a large firm if it involves high disutility, because working is particularly painful for wealthy CEOs. If all CEOs are risk-neutral, then “effective size” becomes $S_n e^{-g_n(\bar{a}_n)/\gamma}$ and is based on size and disutility alone, and so the market equilibrium allocation is first-best efficient, just as in the risk-neutral model of Edmans et al. (2009). Thus, the addition of an effort decision without risk aversion does not lead to distortions – analogously, in a standard effort model, the first-best can be achieved if the agent is risk-neutral and does not face limited liability.

However, when risk aversion is added, the market allocation now depends on risk aversion as well as size and disutility, and is second-best. Large firms that would benefit highly from a talented CEO nevertheless choose to hire a lower-ability CEO if they are risky. Thus, risk aversion leads to two sources of inefficiency. The first is inefficient risk-sharing between firms and CEOs, which also exists in a single-firm moral hazard model and does not affect production. The second, which is specific to a market equilibrium, is distortions in talent assignment that affect real productive activity.¹⁴

We now derive closed-form expressions for both sources of inefficiency to analyze the cost of the moral hazard problem, even when it is fully solved by contracts. If corporate governance were perfect, boards would monitor the manager’s actions directly, achieving first-best. Given imperfect monitoring, moral hazard must be addressed with incentive pay. Even if such contracts are set optimally, the above inefficiencies remain. Direct monitoring and incentives are

¹⁴In the general equilibrium of Kihlstrom and Laffont (1979), risk aversion also leads to inefficient production.

sometimes seen as substitute governance mechanisms; however the former is more efficient as it does not lead to distortions. While recent criticism of corporate governance has centered around inefficiencies in pay-setting (e.g. Bebchuk and Fried (2004)), the losses from misallocation of talent, even if contracting is optimal, can be significant.

Since inefficiency stems solely from risk, not disutility, for simplicity we set $g_n(\bar{a}_n) = 0 \forall n$. The EV thus becomes $\chi'_n = \bar{\Gamma}(\Lambda_n^2 \sigma_n^2)/2$; its mean $\bar{\chi}'$ is defined analogously using (10). Let

$$W = \int w(n) dn \quad (16)$$

denote the total salary received by CEOs, and normalize the wage of the least talented manager, $w(N)$, to 0. Since a wage of w is worth an “effective” wage of $w e^{-\chi'_n}$, the total loss due to inefficient risk-sharing is:

$$L_{RA} = \int [w(n) - w(n) e^{-\chi'_n}] dn.$$

If $\hat{T}(n)$ denotes the talent of the CEO assigned to firm n under the second-best allocation, the loss due to inefficient talent assignment is given by:

$$L_{Alloc} = \int CS(n)^\gamma T(n) dn - \int CS(n)^\gamma \hat{T}(n) dn.$$

The losses are given in the following Proposition.

Proposition 2 *The loss due to inefficient risk-sharing is:*

$$L_{RA} = E [1 - e^{-\chi'_n}] W \quad (17)$$

and the loss due to inefficient talent assignment is:

$$L_{Alloc} = \frac{1}{\beta} E \left[e^{\frac{\beta}{\alpha \gamma} (\chi'_n - \bar{\chi}')} - 1 \right] W. \quad (18)$$

For small distortions, these expressions become:

$$L_{RA} \sim E [\chi'_n] W \quad (19)$$

$$L_{Alloc} \sim \frac{1 + \beta}{2\alpha^2 \gamma^2} \text{var}(\chi'_n) W. \quad (20)$$

Proof See Appendix A. ■

Both sources of inefficiency are proportional to W , the total wage bill. This is intuitive: the economic importance of a distortion to a factor of production is proportional to its marginal product; for a worker this is measured by his wage. From (19), the approximate loss due to inefficient risk-sharing depends on the mean of χ'_n , since this affects the amount of risk the

average CEO has to bear. By contrast, from (20), the approximate loss due to misallocation is proportional to the variance of χ'_n . If $\chi'_n = \bar{\chi}'$, the rankings of effective size $S_n e^{-\chi'_n/\gamma}$ coincide exactly with the rankings of size S and there is no distortion. It is *relative* differences in χ'_n which cause the rankings to differ and the assignment to be affected. In Theorem 1 we showed that, if disutility rose across all firms by the same additive amount ξ , there is no effect on the equilibrium. Here, we can see that a *proportional* change does have an impact: if the marginal cost of effort Λ_n expands by the same ratio across all firms, $var(\chi'_n)$ and thus L_{Alloc} increases. A given proportional increase in disutility equates to a greater absolute increase for a firm with high disutility to begin with. This in turn requires the firm to pay a particularly high dollar premium to talented managers, and causes it to choose a less skilled CEO. One example of such a change is a proportional tax, such as the UK's tax on 2009 banker bonuses. By contrast, a lump-sum tax would have no effect.

Holding W constant, α , β and γ have the same effects on allocational efficiency L_{Alloc} as they do for the dispersion of wages in (12). The intuition is similar: when α and γ are high, distortions due to differences in χ'_n have a small effect. The ranking of effective size is similar to the ranking of unadjusted size and so assignment is little affected. By contrast, a higher β means that talent is more dispersed, and so the losses from misallocation of talent are greater.

Section 3.3 calibrates the magnitude of these losses. Note that we can already draw some conclusions from the analytical expressions in Proposition 2: the inefficiencies will be moderate as they are proportional to the total wage bill W rather than firm size. This is because firms contract efficiently and hire CEOs optimally, given the need to pay a premium for risk and disutility. Indeed, the allocation is second-best efficient: given the existence of a moral hazard problem (the unobservability of effort), a social planner with the same information as firms could not improve on the outcome. A formal proof is in Appendix A; a heuristic argument is that the efficient contract is the one in Proposition 1; given that this is offered, the competitive matching is efficient.

2.3.2 Losses From Random Assignment

For comparison with the above moderate losses, we now conduct the following thought experiment.¹⁵ Assume that poor corporate governance instead manifests in CEOs being randomly allocated, rather than a second-best optimal assignment. Each firm in the top N by size hires a CEO at random from the top MN CEOs by talent, where $M \geq 1$ is a parameter we discuss below. The efficiency loss is

$$L_{Rand} = \int CS(n)^\gamma T(n) - \int CS(n)^\gamma \bar{T} dn,$$

¹⁵Llense (2009) studies the efficiency losses from another thought experiment, the imposition of a pay cap, in a pure assignment model without moral hazard.

where $\bar{T} = \frac{1}{MN} \int_0^{MN} T(n) dn$ denotes the mean talent.

Proposition 3 *If $\alpha\gamma > 1$, the losses from a random allocation of CEOs are infinite, $L_{Rand} = +\infty$. If $\alpha\gamma < 1$,*

$$L_{Rand} = \left[\frac{1 - \alpha\gamma + \beta}{(1 + \beta)(1 - \alpha\gamma)} M^\beta - 1 \right] \frac{W}{\beta}. \quad (21)$$

For $M = 1$, this specializes to

$$L_{Rand} = \frac{\alpha\gamma}{(1 - \alpha\gamma)(1 + \beta)} W. \quad (22)$$

Proof See Appendix A. ■

Equation (20) showed that losses due to misallocation of talent resulting from moral hazard are decreasing in α and γ (holding W constant). By contrast, the losses due to random assignment are increasing in α and γ . In Proposition 2, assignment is second-best optimal. Thus, when effective firm size is more dispersed (α and γ are higher), variation in χ_n has a relatively small effect on the rankings of effective firm size and we remain close to positive assortative matching. The losses from second-best matching are thus lower. In Proposition 3, assignment is random. Thus, when effective firm size is more dispersed, the losses from random matching are higher. Since talent has a multiplicative effect on scaled firm size S^γ (equation (6)), the cost of random assignment of talent is a function of scaled firm size. When $\alpha\gamma > 1$, this mean firm size $E[S^\gamma]$ is infinite and so losses are infinite.

There are two natural choices for M . One is $M = 1$, i.e. the top N firms randomly choose from the top N CEOs, in which case the losses are given by (22). However, this is not an equal comparison with the losses from second-best assignment given in Proposition 2. With $M = 1$, all firms are guaranteed a CEO in the top N . By contrast, in the allocation of Proposition 2, a firm of size rank N hires a CEO of talent rank $Ne^{\chi_n - \bar{\chi}}$. Therefore, the worst manager that can be hired has rank MN , where

$$M = \sup e^{\chi_n - \bar{\chi}}. \quad (23)$$

Thus, a second natural choice for M in Proposition 3 is given by (23), in which case the worst manager that can be hired also has rank $N \sup e^{\chi_n - \bar{\chi}}$, just as in Proposition 2. With this choice of M , the losses under random and second-best allocation can be directly compared. Since $1 < \frac{1 - \alpha\gamma + \beta}{(1 + \beta)(1 - \alpha\gamma)}$, we have

$$\begin{aligned} L_{Alloc} &= E \left[e^{\frac{\beta}{\alpha\gamma}(\chi'_n - \bar{\chi}')} - 1 \right] \frac{W}{\beta} \\ &\leq (M^\beta - 1) \frac{W}{\beta} = \left[\frac{1 - \alpha\gamma + \beta}{(1 + \beta)(1 - \alpha\gamma)} M^\beta - 1 \right] \frac{W}{\beta} = L_{Rand} \end{aligned}$$

Thus, $L_{Alloc} \leq L_{Rand}$ as is intuitive: second-best matching is superior to random matching. The difference is increasing in α and γ , as these variables raise the dispersion of effective firm

size and thus importance of second-best matching.

3 Extensions

3.1 Providing Risk-Taking Incentives

3.1.1 General Theorem

In the core model, the CEO can improve the mean return r without changing risk, which is exogenous at σ . In reality, increasing firm value may require taking on risky, positive-NPV projects – indeed, many commentators argue that a major goal of incentive compensation is to induce managers to take actions that improve firm value even if they augment risk (see, e.g., Core, Guay and Larcker (2003)). In this section, we endogenize σ so that it depends on the mean return chosen by the CEO. In a standard model with effort and risk-taking where noise follows the action (e.g. Dittmann and Yu (2009)), the above choice can be modeled by allowing the CEO to choose a single action a , which affects both the mean and volatility of the return. Since the action affects the distribution of the noise, the noise must follow the action. However, the framework we use to achieve tractability requires no noise to follow the CEO’s final action, so that the IC constraints hold state-by-state. We therefore operationalize the CEO’s risk choice by extending the model to two periods, so that there are two actions and a single noise in between. The first-period action a_1 affects both the mean of the first-period signal r_1 and the volatility of the second-period signal r_2 . (In this subsection, subscripts index time periods rather than the rank of a firm or CEO.) The second-period action a_2 affects the mean of r_2 only, since there is no noise to follow this action. As with earlier, the principal implements \bar{a} in each period. The full timing is as follows:

1. Noise η_1 is privately observed by the CEO.
2. The CEO chooses a_1 .
3. The signal $r_1 = a_1 + \eta_1$ is publicly observed.
4. Noise η_2 is privately observed by the CEO.
5. The CEO chooses a_2 .
6. The signal $r_2 = a_2 + \sigma(a_1)\eta_2 + \mu(a_1)$ is publicly observed.

We have $\sigma'(a_1) \geq 0$, so that actions to improve firm value also entail augmenting risk (e.g. taking on risky, positive-NPV projects). To ensure that $E[e^{r_2} | a_1]$ is independent of a_1 (so that a_1 affects the volatility of firm value proportionally to e^{a_1}), we assume $E[\eta_2] = 0$ and take

$$\mu(a_1) = -\ln E[e^{\sigma(a_1)\eta_2}]. \quad (24)$$

To our knowledge, the contracting problem where the agent affects the volatility as well as the mean has only been solved in specific cases. Sung (1995) and Ou-Yang (2003) study the Holmstrom and Milgrom (1987) case of exponential utility, a financial cost of effort, continuous time and Gaussian noise, and Dittmann and Yu (2009) consider separable preferences and Gaussian noise in a one-period model.¹⁶ Therefore, before specializing to the utility function (4) used in this paper, we first derive the result for the more general utility function:

$$U(c, a_1, a_2) = u[(v(c) - g_1(a_1) - g_2(a_2))], \quad (25)$$

where $u(x) = e^{(1-\Gamma)x}/(1-\Gamma)$ for $\Gamma \neq 1$ and $u(x) = x$ for $\Gamma = 1$. The only assumption we make on v is that it is increasing and weakly concave. The utility function (4) corresponds to $v(c) = \ln c$ and a single action.

Theorem 2 (*Optimal contract, endogenous risk*). *The optimal contract pays the CEO an amount c defined by:*

$$c(r_1, r_2) = v^{-1}(\Lambda_1 r_1 + \Lambda_2 r_2 + K) \quad (26)$$

with

$$\Lambda_1 = \begin{cases} g'_1(\bar{a}) - \Lambda_2 \mu'(\bar{a}) - \frac{1}{1-\Gamma} \frac{d}{da_1} \ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2)}]_{a_1=\bar{a}} & \text{if } \Gamma \neq 1 \\ g'_1(\bar{a}) - \Lambda_2 \mu'(\bar{a}) & \text{if } \Gamma = 1, \end{cases} \quad (27)$$

$$\Lambda_2 = g'_2(\bar{a}). \quad (28)$$

and K is a constant that makes the CEO's participation constraint bind.

For the particular case where η_2 is Gaussian, or the limit of small noises, then $\mu'(\bar{a}) = -\sigma(\bar{a})\sigma'(\bar{a})$, and so

$$\Lambda_1 = g'_1(\bar{a}) + [\Lambda_2 + (\Gamma - 1)(\Lambda_2)^2] \sigma(\bar{a})\sigma'(\bar{a}). \quad (29)$$

Proof See Appendix A. ■

3.1.2 Application to CRRA Preferences

The utility function (4) corresponds to $v(c) = \ln c$. Applying Theorem 2 to this case yields the following result.

Proposition 4 (*CEO pay in partial equilibrium, endogenous risk*). *The optimal contract pays the CEO an amount c defined by:*

$$\ln c = \Lambda_1 r_1 + \Lambda_2 r_2 + K,$$

¹⁶Lambert (1986) considers a model in which the agent takes separate effort and volatility decisions, in a model where output is restricted to three possible levels.

where Λ_1 and Λ_2 are given by (27) and (28). On the equilibrium path this can be rewritten:

$$\ln c = k + \eta$$

where k is a constant that makes the CEO's participation constraint bind, and

$$\eta = \Lambda_1 \eta_1 + \Lambda_2 \sigma(\bar{a}) \eta_2 \tag{30}$$

is the total noise to which the contract exposes the agent.

The market equilibrium allocation and wage are given by equations (11) and (13) in Theorem 1, with the EV now defined by

$$\chi_n = g_{1n}(\bar{a}_n) + g_{2n}(\bar{a}_n) + \frac{\bar{\Gamma}(\eta_n)}{2},$$

where n indexes firm n 's risk and cost of effort, $\eta_n = \Lambda_{n1} \eta_{n1} + \Lambda_{n2} \sigma(\bar{a}_n) \eta_{n2}$, and we define (with a slight abuse of notation):

$$\bar{\Gamma}(\eta_n) = 2 \left(\ln E[e^{\eta_n}] - \frac{1}{1-\Gamma} \ln E[e^{(1-\Gamma)\eta_n}] \right).$$

EG show that, under exogenous risk, $\Lambda_1 = g'_1(\bar{a})$ and $\Lambda_2 = g'_2(\bar{a})$. We compare this with our slope under endogenous risk with small or Gaussian noises, equation (29). The core case is $\Gamma \geq 1$. Λ_1 is higher when the CEO affects firm risk, since the contract must now induce not only effort but also risk-taking. A risk-averse CEO may forgo risky, positive-NPV projects. To induce him to accept such a project, it is necessary to give him a more convex payout so that he benefits from risk. Since the strength of incentives Λ_1 also represents the convexity of dollar pay to firm value, this increased convexity is achieved by raising Λ_1 .

The strength of incentives Λ_1 is increasing in four parameters. First, it is increasing in risk aversion Γ : the more risk-averse the CEO, the greater the convexity needed to overcome his risk aversion. For similar reasons, it is increasing in $\sigma(\bar{a})$ (the level of firm risk) and Λ_2 (the CEO's exposure to the risk induced by a_1 .) The positive relationship between incentives and risk $\sigma(\bar{a})$ contrasts the negative association predicted by standard models, which assume exogenous risk and posit a trade-off between incentives and risk-sharing, but is consistent with the empirical findings of Demsetz and Lehn (1985), Core and Guay (1999) and Oyer and Schaefer (2004). Finally, Λ_1 rises in the marginal increase in risk caused by implementing all positive-NPV projects $\sigma'(\bar{a})$. The intuition is similar: the greater the additional risk imposed by a positive-NPV project, the greater the convexity the CEO must be given to induce him to take it. If the main way in which the CEO affects firm value is by not diverting cash flows, there is no link between risk and return and so $\sigma'(\bar{a}) = 0$ and $\Lambda_1 = g'_1(\bar{a})$. By contrast, if the key CEO action is the choice of risky projects, $\sigma'(\bar{a}) > 0$ and Λ_1 increases. $\sigma'(\bar{a})$ is likely to be high in

new economy firms since they have little tangible capital and so enhancing firm value involves greater risk – investing in R&D has a zero payoff if the R&D fails, whereas investing in an old economy plant generates liquidation value upon failure. Indeed, incentives are stronger in new economy firms (Ittner, Lambert and Larcker (2003), Murphy (2003)) and have risen over time (Jensen and Murphy (2004)).

An interesting benchmark case is that of a risk-neutral CEO. Plugging $\Gamma = 0$ into (29) gives

$$\Lambda_1 = g'_1(\bar{a}) + [\Lambda_2 - \Lambda_2^2] \sigma(\bar{a}) \sigma'(\bar{a}),$$

which is *lower* than $g'_1(\bar{a})$ if and only if $\Lambda_2 > 1$. A risk-neutral CEO only cares about the expected value of his compensation. If $\Lambda_2 > 1$, then his compensation is a convex function of the firm's market value¹⁷, and thus he has incentives to take excessive risk, i.e. choose an a_1 above the maximum productive level \bar{a} . A lower Λ_1 offsets this tendency and induces the CEO to reduce a_1 to the optimal level. In sum, our results contrast the argument (often made by critics of executive pay) that powerful incentives induce the CEO to take excessive risk, and thus if the CEO is able to affect risk as well as the average return, incentives should be weaker. For the core case of $\Gamma \geq 1$, incentives are unambiguously stronger; only if Γ is sufficiently low and Λ_2 is sufficiently high will incentives be shallower.

We note two additional points. First, even if $g_1(a) = 0 \forall a$ (i.e. the risk-increasing action is costless to the CEO), Λ_1 is typically non-zero – incentives are necessary not because the efficient action requires the CEO to exert effort, but because it exposes him to risk. This is consistent with the idea mentioned at the start of this section, that incentives are used to induce risk-taking, rather than solely to induce effort. Second, since a_1 affects $r_2 (= a_2 + \sigma(a_1)\eta_2 + \mu(a_1))$, it may seem that Λ_2 could be used to control the CEO's choice of a_1 . However, Λ_2 is unchanged at $g'_2(\bar{a})$. This is because the time-2 IC condition must hold state-by-state, i.e. for every possible realization of η_2 . In turn, this forces the slope of the contract (i.e. benefits from effort) to equal the marginal cost of effort, $g'_2(\bar{a})$. This is a similar intuition to the contract's tractability, described at the end of Section 2.2 – since the IC conditions must hold state-by-state, the principal has little freedom in designing the contract.

3.2 Outside Options

The core model considered a single labor market (CEOs) in fixed supply. In reality, CEOs may be able to find jobs outside the CEO market, and firms may hire managers currently employed in other sectors. We thus extend the model to allow for an elastic supply of talent. To do so in a tractable way, we assume the existence of an integrated market between the “corporate sector” and the “non-corporate sector.” The former represents the CEO labor market, and the latter

¹⁷The dollar pay received by the CEO as a result of second-period performance is $e^{\Lambda_2 r_2}$. Substituting $r_2 = \ln(P_2/P_1)$ gives $(P_2/P_1)^{\Lambda_2}$, which is convex in P_2 if and only if $\Lambda_2 > 1$.

represents alternative jobs, such as hedge funds, entrepreneurship or consulting. We assume that firms in both sectors initially have identical characteristics, and that the fraction of firms in the corporate and non-corporate sectors are respectively $1 - \pi$ and π . The probability that a firm is in the corporate sector is drawn independently from the distribution of firm sizes.

Theorem 1 showed that, if the disutility of being a CEO at any firm rises from $g_n(\bar{a}_n)$ to $g_n(\bar{a}_n) + \xi$, the level of pay is unchanged – while working for one’s own firm becomes less unattractive, the outside option of being a CEO at another firm also becomes undesirable. We revisit this prediction in the case where the CEO has an additional outside option, the non-corporate sector, in which disutility is unchanged.

Proposition 5 *Suppose that the disutility of working in the corporate sector rises from $g_n(\bar{a}_n)$ to $g_n(\bar{a}_n) + \xi$ for a small ξ . Then:*

- (i) *Log pay in the corporate sector increases by $\frac{\beta\pi}{\alpha\gamma}\xi$.*
- (ii) *The talent rank of a manager hired by firm n in the corporate sector rises from n to $n\left(1 + \frac{\pi}{\alpha\gamma}\xi\right)$, i.e. the sector hires less talented workers.*
- (iii) *The total loss of value creation (aggregate firm value gross of wages) by the corporate sector is $W\frac{\pi}{\alpha\gamma}\xi$, where W is the initial amount paid to CEOs in the corporate sector.*

Proof See Appendix A. ■

Part (i) of Proposition 5 shows that the log wage increases by $\frac{\beta}{\alpha\gamma}\pi\xi$. The intuition behind the effect of β , α and γ is the same as for their effect on the pay equation (12), discussed earlier: when α and γ are large, ξ has a small effect on the distribution of scaled size S^γ ; when β is high, firms are more willing to pay to retain talent. Part (ii) shows that a corporate firm hires a less talented manager. The intuition is similar to the distortion to CEO assignment caused by moral hazard, discussed in Theorem 1. Since corporate firms must pay a premium for the increased disutility of being a CEO, and the premium is multiplicative in the wage and thus greater for more talented workers, corporate firms hire less skilled agents. The intuition behind the effect of α and γ is the same as in equation (12); the dispersion of talent β has no effect since part (ii) refers to the talent rank of a manager. Part (iii) shows that the total loss in value created by the corporate sector is increasing in the aggregate pay of corporate CEOs W , for the same reason as in Proposition 2.

All three outcomes are increasing in π , the size of the non-corporate sector, as this represents the outside option. When outside options are larger, a higher wage premium is required to keep a CEO within the corporate sector (part (i)). (When the corporate sector is the entire economy ($\pi = 0$), pay does not change, since CEOs have no outside option; this is the result from Theorem 1.) This greater premium in turn leads to greater distortions in CEO assignment (part (ii)) and consequently more value loss (part (iii)).

Note that a sector may be defined as a specific industry within the corporate sector, rather than the CEO market as a whole. This interpretation allows us to study the effect of industry-specific trends. For example, the two “sectors” could be the financial industry and all other jobs

that a financial CEO could take. The recent financial crisis has led to increased regulation of the financial industry in particular, with little changes in the other industries. As warned by some commentators, this may lead talented CEOs to leave the financial industry. The magnitude of the value loss is increasing in the extent of financial CEOs' outside options, and so may be particularly large because financial CEOs often have the option of not only becoming the CEO of a non-financial firm, but also moving to a hedge fund or private equity firm.

3.3 A Calibration

We now undertake an approximate calibration of the efficiency losses in Section 2.3. We start with the losses from second-best assignment under moral hazard, given in Proposition 2. As in GL, we take $\alpha = \gamma = 1$ and $\beta = 2/3$ and consider the top 500 firms in Execucomp by aggregate value. For 2006, aggregate flow compensation ("tdc1" in Execucomp, winsorizing at the 5th and 95th percentiles) is $W = \$5$ billion.

We start with the estimation of L_{Alloc} , given by equation (18). The key challenge is to estimate $\chi'_n - \bar{\chi}'$. This depends on the marginal cost of effort, which is inherently unobservable and we are unable of any previous studies that estimate it. However, a key advantage of our unifying framework is that we can infer $\chi'_n - \bar{\chi}'$ by using the wage equation (12). Taking logs of this equation and rearranging yields

$$y_n = \ln w_n - (\gamma - \beta/\alpha) \ln S_n,$$

where $y_n = k + \frac{\beta}{\alpha\gamma} (\chi_n - \bar{\chi})$ and k is a constant. We have

$$E [e^{-y_n/\beta}] = e^{-k/\beta} E [e^{-(\chi_n - \bar{\chi})/\alpha\gamma}] = e^{-k/\beta}$$

by definition of $\bar{\chi}$, so $e^{-k} = E [e^{-y_n/\beta}]^\beta$. Using $E [e^{\frac{\beta}{\alpha\gamma}(\chi_n - \bar{\chi})}] = E [e^{y_n - k}]$, we obtain:

$$E [e^{\frac{\beta}{\alpha\gamma}(\chi_n - \bar{\chi})}] = E [e^{y_n}] E [e^{-y_n/\beta}]^\beta. \quad (31)$$

We use (31) as an estimate of $E [e^{\frac{\beta}{\alpha\gamma}(\chi'_n - \bar{\chi}')}]$ in equation (18). This yields an estimate of L_{Alloc} as \$7.4 billion. We note two potential issues with our approach. The first is that equation (18) contains $\chi'_n - \bar{\chi}'$, but (31) contains $\chi_n - \bar{\chi}$. Recall that $\chi_n = g_n(\bar{a}_n) + \frac{\bar{\Gamma}(\Lambda_n^2 \sigma_n^2)}{2}$ and $\chi'_n = \frac{\bar{\Gamma}(\Lambda_n^2 \sigma_n^2)}{2}$, so we are implicitly assuming that $g_n(\bar{a}_n)$ is the same across firms. Empirically, cross-sectional variation in χ_n from the average $\bar{\chi}$ may stem from variation in $g_n(\bar{a}_n)$, but the above approach attributes it entirely to differences in χ'_n . Since L_{Alloc} is increasing in the variance of χ'_n , this has the potential to overstate L_{Alloc} . One goal of the calibration is to highlight that the losses from random assignment in Proposition 3 are significantly greater than those from second-best assignment in Proposition 2. Thus, by providing an upper bound on

L_{Alloc} , this approach works against us by underestimating the differences in losses.

The second caveat is that y_n may be slightly mis-measured in practice - even though firm size and the CEO's wage are observable, it may be that the CEO's actual wage differs from his market wage, e.g. if he is given deferred compensation. Again, measurement errors will overstate L_{Alloc} ; moreover, we can estimate the likely magnitude of the resulting bias. Let y_n^* denote the true value and y_n the observed value, and assume the classic errors-in-variables structure $y_n = y_n^* + u_n$ where y_n^* and u_n are independent. This yields the decomposition:

$$E[e^{y_n}] E[e^{-y_n/\beta}]^\beta = E[e^{y_n^*}] E[e^{-y_n^*/\beta}]^\beta E[e^{u_n}] E[e^{-u_n/\beta}]^\beta.$$

Using the notation $\Phi(y) = E[e^{y_n}] E[e^{-y_n/\beta}]^\beta$, the decomposition can be rewritten:

$$\Phi(y) = \Phi(y^*) \Phi(u).$$

The measured $\Phi(y)$ overstates $\Phi(y^*)$ by a factor $\Phi(u)$, which exceeds 1 by Jensen's inequality. To estimate the magnitude of the bias, suppose that $u \sim N(0, \sigma_u^2)$. Then,

$$\Phi(u) = e^{\sigma_u^2/2} \left(e^{\sigma_u^2/2\beta^2} \right)^\beta = e^{(1+1/\beta)\sigma_u^2/2} = e^{\frac{5}{4}\sigma_u^2}.$$

If the measurement error is moderate, e.g. $\sigma_u = 0.2$, the bias is $e^{\frac{5}{4}0.2^2} = 1.05$, i.e. only 5%. Indeed, replacing w by a three-year average has little effect on the results.

We now turn to L_{RA} , given by (17). This requires an estimate of χ'_n alone, rather than $\chi'_n - \bar{\chi}'$, and so we cannot use the above method. We thus infer the marginal cost of effort from observed contracts, under the assumption that firms are contracting efficiently. In our one-period model, where incentives stem only from newly-granted stock and options, Λ_n equals the percentage change in pay for a percentage point return. In reality, the bulk of a CEO's incentives stems from previously granted stock and options (see, e.g., Hall and Liebman (1998), Core, Guay, and Verrecchia (2003)). Hence, Edmans, Gabaix and Landier's (2009) measure of incentives is the dollar change in *wealth* for a one percentage point return, scaled by annual pay, which they call B^I .¹⁸ While B^I measures the sensitivity of CEO wealth to the current period return, the CEO bears risk from changes in the stock price during his entire tenure as CEO. To convert B^I into an estimate of Λ , we assume the CEO works for L years and consumes only at the end (we suppress the dependence on n for brevity). Then, the contract becomes:

$$\ln c = \Lambda \sum_{t=1}^L r_t + K.$$

¹⁸This dataset is available at <http://finance.wharton.upenn.edu/~aedmans/data.html>. The data construction is described in Appendix B of Edmans, Gabaix and Landier (2009).

Thus, $\text{var}(\ln c) = \Lambda^2 \sigma^2 L$, and $\chi' = \frac{\Gamma(\Lambda^2 \sigma^2 L)}{2}$. We also have

$$B^I = \frac{d\text{Wealth}/dr}{w} = \Lambda \frac{\text{Wealth}}{\text{Wage}} = \Lambda R$$

where $R = \text{Wealth}/\text{Wage}$, and since $\Lambda = (d\text{Wealth}/\text{Wealth})/dr$ in a multi-period model.

To estimate the CEO's wealth, we start by taking data from Dittmann and Maug (2007), who estimate the CEO's non-firm wealth which results from past salary and bonus awards, and sales of stock and options previously granted by the firm.¹⁹ We then add the CEO's current wealth invested in the firm, from stock and options, to give a total wealth measure. Unfortunately, it is not possible to obtain data on the wealth of U.S. CEOs that does not stem from past or current executive compensation (e.g. real estate ownership or holdings of other securities), but this is a reasonable benchmark. CEOs in 2006 have been in their current position for a median of 5.6 years. This is an estimate of $L/2$ (since most CEOs will continue in office after 2006) and so it corresponds to $L = 11.2$. Taking the benchmark case of log utility ($\Gamma = 1$) and using the exact expression (17) yields $L_{RA} = 0.34W = \$1.7$ billion. Note that this method is not suitable for estimation of L_{Alloc} : since χ'_n enters with a positive exponent, L_{Alloc} would be extremely sensitive to outliers in χ'_n . Indeed, χ'_n exhibits significant outliers, due to outliers in volatility, wages, or stock and option holdings. The estimation of L_{RA} is much less affected since χ'_n enters with a negative exponent.²⁰ In sum, the preliminary calibration suggests total efficiency losses from imperfect monitoring (while retaining optimal contracting) of approximately \$9 billion, twice the aggregate CEO salary.

We now turn to the losses from random allocation, given in Proposition 3. We focus on the case of $M = 1$ since this does not require estimation of the χ_n 's. Since this means that each firm is guaranteed to end up with a top-500 CEO, our results will represent a lower bound. To estimate $\frac{1}{1-\alpha\gamma}$, we use the identity²¹:

$$\frac{1}{1-\alpha\gamma} = E \left[\left(\frac{S}{S_*} \right)^\gamma \mid S \geq S_* \right]. \quad (32)$$

Indeed, for an arbitrary size cutoff S_* in the support of S , $P(S \geq xS_* \mid S \geq S_*) = x^{-1/\alpha}$, so

$$E \left[\left(\frac{S}{S_*} \right)^\gamma \mid S \geq S_* \right] = \int_1^\infty x^\gamma x^{-1/\alpha-1} \frac{1}{\alpha} dx = \frac{1}{\alpha} \left[\frac{-x^{-1/\alpha+\gamma}}{\frac{1}{\alpha} - \gamma} \right]_1^\infty = \frac{1}{1-\alpha\gamma}.$$

¹⁹This dataset is available at <http://people.few.eur.nl/dittmann/data.htm>. We thank Ingolf Dittmann for generously making this data available.

²⁰In addition, as with any calibration, we assume real-world data is optimal and thus can be used to estimate Λ_n . If, however, governance failures manifest in suboptimal contracting, our measure of Λ_n is inaccurate.

²¹Our calibration uses Zipf's law and constant returns to scale ($\alpha = \gamma = 1$). L_{Rand} in (22) is thus on the cusp of being divergent (with a weak, logarithmic divergence), and so it is easier to estimate with (32).

Taking $\gamma = 1$, the term $E \left[\left(\frac{S}{S_*} \right)^\gamma \mid S \geq S_* \right]$ can be estimated as the mean firm size above a cutoff S_* , divided by S_* . We define S_* as the size of the median firm in our top 500 (the 250th largest firm), which yields 5.3. From (22) we have $L_{Rand} \sim 5.3W / (5/3) = 3.2W \sim \16 billion. This is markedly greater than the combined losses from second-best assignment in Proposition 2. Moreover, since our estimate of L_{Rand} is a lower bound (it assumes $M = 1$) and the estimate of L_{Alloc} is an upper bound, the true difference is likely to be significantly greater. Naturally, the losses from both imperfect monitoring and random allocation will be significantly higher if we consider all top executives, rather than just the CEO.

4 Conclusion

This paper studies how CEO assignment, pay and incentives depend on talent, talent impact, firm size, risk and disutility in market equilibrium. The model's closed-form solutions allow the determinants of these three outcomes to be transparent, and clear empirical predictions. In talent assignment models without an effort conflict, the most talented managers are assigned to the largest firms. We show that this efficient allocation is distorted in the presence of moral hazard – a firm that is riskier or involves greater disutility hires a less talented CEO. The loss in efficiency is decreasing in the dispersion of firm size and size elasticity of talent, and increasing in the dispersion of managerial ability. If poor corporate governance instead manifests in a random assignment of CEOs to firms, the losses are significantly higher, and affected by the above parameters in the opposite direction.

Cross-sectional changes in risk and disutility increase the level of pay. Thus, risky firms not only hire less talented CEOs, but also pay their CEOs highly (relative to their skill level) as compensation. However, aggregate changes in these variables have no impact as they affect the current firm and outside options equally. The strength of incentives is increasing in the disutility of effort, but independent of risk and risk aversion if the CEO only affects mean returns. If value-enhancing actions by the CEO also increase firm risk, the contract slope generally rises and exhibits a positive relationship with both risk and risk aversion.

While a number of the model's predictions regarding pay and incentives are consistent with existing empirical findings, some predictions regarding talent assignment are yet to be tested (given the difficulties of measuring talent) and are potentially fruitful topics for future empirical research.²² In terms of future theoretical directions, it would be interesting to extend the analysis to a dynamic model where CEOs can be fired or voluntarily move between jobs. Axelson and Bond (2009) consider a dynamic market equilibrium under risk-neutrality, and Tsuyuhara (2009) assumes homogeneous agents and firms. Whether tractability can be preserved under the combination of dynamics, risk aversion and skill differences is an open question.

²²See Acharya, Gabarro and Volpin (2010), Chang, Dasgupta and Hilary (2010) and Nguyen and Nielsen (2010) for promising approaches to measuring talent.

A Proofs

Proof of Proposition 1

This is a special case of Theorem 1 of EG. EG have the utility function $u(v(c) - g(a))$; our utility function (4) is a particular case of this with $u(x) = \frac{e^{(1-\Gamma)x}}{1-\Gamma}$ and $v(c) = \ln c$. We refer the reader to EG for the full proof, which does not use first-order conditions and rules out contracts that are stochastic or depend on messages. Here we give a heuristic proof that conveys the intuition, so that the intuition is self-contained within this paper. Given $r = a + \eta + \mu$, the agent's expected utility is given by

$$E[U] = E \left[\frac{(c(a + \eta + \mu) e^{-g(a)})^{1-\Gamma}}{1-\Gamma} \right].$$

Since η is known when the agent takes his action, we can remove the expectations operator. The IC condition is thus:

$$\bar{a} \in \arg \max_{a \in [\underline{a}, \bar{a}]} c(a + \eta + \mu) e^{-g(a)} \quad (33)$$

Taking the first order condition yields:

$$c'(\bar{a} + \eta + \mu) e^{-g(\bar{a})} - g'(\bar{a}) e^{-g(\bar{a})} c(\bar{a} + \eta + \mu) = 0$$

i.e.

$$\frac{c'(r)}{c(r)} = g'(\bar{a}) = \Lambda.$$

Since this must hold state-by-state (i.e. for every possible η and r found on the equilibrium path), this integrates to

$$\ln c = \Lambda r + K.$$

Proof that maximum effort is optimal if s is sufficiently large

Consider a CEO with a reservation utility $\underline{u} = u(\ln v_n)$, where v_n is the “effective” dollar wage (defined later in equation (9)). Call $[\underline{\eta}, \bar{\eta}]$ the support of η , $f(\eta)$ its density, and $\bar{F}(x) = P(\eta > x)$ the complementary cumulative distribution function. Define

$$Q_n \equiv \frac{1}{e^{\bar{a}_n + \underline{\eta}_n}} \left(g'_n(\bar{a}_n) + g''_n(\bar{a}_n) \sup_{\eta} \frac{\bar{F}_n(\eta)}{f_n(\eta)} \right) e^{g'_n(\bar{a}_n) + (\bar{\eta}_n - \underline{\eta}_n) g'_n(\bar{a}_n)}$$

From condition (28) in EG (applied to $b(a, \eta) = e^{a+\eta}$ and $v(c) = \ln c$, and $e^{u^{-1}(\underline{u})} = v_n$), the firm wishes to implement \bar{a} for all η if:

$$s_n > Q_n v_n.$$

This condition requires firm value s_n to be sufficiently large compared to the CEO's effective

wage v_n .

Proof of Equation (8)

The CEO's wage is:

$$\ln c = \Lambda r + K = \Lambda(a + \eta) + K = \eta' + K'$$

with $\eta' = \Lambda\eta$, $K' = \Lambda a + K$. Thus his expected wage is:

$$w = E[c] = E[e^{\eta'+K'}] = E[e^{\eta'}] e^{K'} = e^{\ln E[e^{\eta'}]+K'}$$

His expected utility is:

$$\begin{aligned} U_n &= \frac{1}{1-\Gamma} E \left[(ce^{-g(a)})^{1-\Gamma} \right] = \frac{1}{1-\Gamma} E \left[e^{(\ln c - g(a))(1-\Gamma)} \right] \\ &= \frac{1}{1-\Gamma} E \left[e^{(\eta'+K'-g(a))(1-\Gamma)} \right] = \frac{1}{1-\Gamma} E \left[e^{(\eta')(1-\Gamma)} \right] e^{(K'-g(a))(1-\Gamma)} \\ &= \frac{1}{1-\Gamma} e^{(1-\Gamma) \left[\frac{1}{1-\Gamma} \ln E[e^{\eta'(1-\Gamma)}] + K' - g(a) \right]}. \end{aligned}$$

Hence, with:

$$\bar{\Gamma}(\Lambda^2\sigma^2)/2 = \ln E[e^{\eta'}] - \frac{1}{1-\Gamma} \ln E[e^{(1-\Gamma)\eta'}],$$

we have

$$\begin{aligned} U_n &= \frac{1}{1-\Gamma} e^{(1-\Gamma) \left[\ln E[e^{\eta'}] - \bar{\Gamma}(\Lambda^2\sigma^2)/2 + K' - g(a) \right]} \\ &= \frac{1}{1-\Gamma} e^{(1-\Gamma) \left[-(g(a) + \bar{\Gamma}(\Lambda^2\sigma^2)/2) + \ln E[e^{\eta'}] + K' \right]} \\ &= \frac{1}{1-\Gamma} e^{(1-\Gamma) [-\chi + \ln w]} \\ &= \frac{(we^{-\chi})^{1-\Gamma}}{1-\Gamma} \end{aligned}$$

The CEO receives the same utility as if he had to exert no effort, and received a fixed wage $we^{-\chi}$.

Proof of Theorem 1

This proof consists of four steps.

Step 1: Effective sizes. This is derived in the proof in the main paper.

Step 2: Distribution of effective sizes. We use the notations

$$\alpha' = \alpha\gamma, \quad \kappa_n = \chi_n/\alpha', \quad \bar{\kappa} = \bar{\chi}/\alpha'.$$

We use the interpretation of n as a quantile to simplify the algebra. Since $S(n) = An^{-\alpha}$, the

distribution of sizes follows $P(S \geq x) = (x/A)^{-1/\alpha}$. Averaging over all χ_n yields the following distribution function for effective sizes $\widehat{S}_n = S_n e^{-\chi_n/\gamma}$:

$$\begin{aligned}\widehat{F}(x) &= P\left(\widehat{S}_n \geq x\right) = P\left(S_n e^{-\chi_n/\gamma} \geq x\right) = P\left(S_n \geq x e^{\chi_n/\gamma}\right) \\ &= E\left[\left(x e^{\chi_n/\gamma}/A\right)^{-1/\alpha}\right] = \left(\frac{x}{A}\right)^{-1/\alpha} E\left[e^{-\chi_n/(\alpha\gamma)}\right] \\ &= \left(\frac{x}{A}\right)^{-1/\alpha} e^{-\bar{\kappa}}.\end{aligned}$$

We will use the r to denote the rank in effective size, and n for the rank in actual size. The effective size $\widehat{S}(r)$ of the firm of rank r satisfies $\widehat{F}\left(\widehat{S}(r)\right) = r$, i.e. the effective size of the firm of quantile rank r is:

$$\widehat{S}(r) = A e^{-\alpha\bar{\kappa}} r^{-\alpha}.$$

Step 3. Assignment in effective sizes. A firm with effective rank r optimizes over the talent rank q of the manager it wishes to hire:

$$\max_q C \widehat{S}^\gamma(r) T(q) - v(q),$$

which yields $C \widehat{S}^\gamma(r) T'(q) - v'(q) = 0$. In the competitive equilibrium, there is matching between talent and effective size, $q = r$. Hence:

$$C \widehat{S}^\gamma(r) T'(r) = v'(r). \quad (34)$$

Let \underline{v}_N denote the effective reservation wage of the least talented CEO ($n = N$). We obtain the classic assignment equation (Sattinger (1993), Terviö (2008)):

$$v(r) = - \int_r^N C \widehat{S}(u)^\gamma T'(u) du + \underline{v}_N.$$

Using the functional forms $\widehat{S}(u) = A e^{-\alpha\bar{\kappa}} u^{-\alpha}$ and $T'(u) = -B u^{\beta-1}$, we obtain:

$$\begin{aligned}v(r) &= A^\gamma e^{-\alpha'\bar{\kappa}} BC \int_r^N u^{-\alpha\gamma+\beta-1} du + \underline{v}_N = A^\gamma e^{-\alpha'\bar{\kappa}} BC \left[\frac{u^{-\alpha'+\beta}}{-\alpha'+\beta} \right]_r^N + \underline{v}_N \\ &= \frac{A^\gamma BC}{\alpha' - \beta} e^{-\alpha'\bar{\kappa}} \left(r^{-(\alpha'-\beta)} - N^{-(\alpha'-\beta)} \right) + \underline{v}_N\end{aligned}$$

In the limit $(r/N) \rightarrow 0$, the term $r^{-(\alpha'-\beta)}$ dominates the other two, and we have:

$$v(r) = \frac{A^\gamma BC}{\alpha' - \beta} e^{-\alpha'\bar{\kappa}} r^{-(\alpha'-\beta)}.$$

Step 4. Wages. The rank of a firm with effective size $S_n e^{-\chi_n/\gamma}$ is:

$$r = \widehat{F} \left(S_n e^{-\chi_n/\gamma} \right) = \left(\frac{S_n e^{-\chi_n/\gamma}}{A} \right)^{-1/\alpha} e^{-\bar{\kappa}} = \left(\frac{S_n}{A} \right)^{-1/\alpha} e^{\kappa_n - \bar{\kappa}}.$$

In other words, a firm with size rank n hires a manager with size talent rank $r = n e^{\kappa_n - \bar{\kappa}}$ (at least in the upper tail, i.e. in the domain of the power law specification). It pays an effective wage of $v(r)$, and thus a monetary wage of

$$\begin{aligned} w_n &= v(r) e^{\chi_n} \\ &= \frac{A^\gamma BC}{\alpha' - \beta} e^{-\alpha' \bar{\kappa}} \left(\left(\frac{S_n}{A} \right)^{-1/\alpha} e^{\kappa_n - \bar{\kappa}} \right)^{-(\alpha' - \beta)} e^{\alpha' \kappa_n} \\ &= \frac{A^{\beta/\alpha} BC}{\alpha' - \beta} S_n^{\gamma - \beta/\alpha} e^{\beta(\kappa_n - \bar{\kappa})}. \end{aligned}$$

Finally, substituting $S_{n_*} = A n_*^{-\alpha}$, we obtain:

$$\begin{aligned} w_n &= \frac{(S_{n_*} n_*^\alpha)^{\beta/\alpha} BC}{\alpha' - \beta} S_n^{\gamma - \beta/\alpha} e^{\beta(\kappa_n - \bar{\kappa})} \\ &= D(n_*) S_{n_*}^{\beta/\alpha} S_n^{\gamma - \beta/\alpha} e^{\beta(\kappa_n - \bar{\kappa})}, \end{aligned}$$

with $D(n_*) = \frac{n_*^\beta BC}{\alpha' - \beta}$.

A sufficient condition for the risk premium to increase with the wage

Consider a general utility function $U = \frac{e^{(1-\Gamma)(v(c)-g(a))}}{1-\Gamma}$ where v is concave; the core model corresponds to $v(c) = \ln c$. From EG, the optimal contract is $c = v^{-1}(\Lambda r + K)$. From the proof of equation (8) in Appendix A, we have

$$E[U] = \frac{e^{(1-\Gamma)(\Lambda a + K + H)}}{1-\Gamma}$$

where $H = -g(a) + \ln E[e^{\Lambda \eta}] - \bar{\Gamma}(\Lambda^2 \sigma^2)/2 < 0$. Hence, v (effective wage) $= \Lambda a + K + H$, and so the risk premium is given by

$$w - f(\Lambda a + K + H)$$

where $f = v^{-1}$.

Suppose now that w increases a small amount Δ and $\Lambda r + K$ increases δ such that $w + \Delta = E[f(\Lambda r + K + \delta)]$. The risk premium becomes $w + \Delta - f(\Lambda a + K + \delta + H)$. For the risk premium to be increasing in w , we require

$$w + \Delta - f(\Lambda a + K + \delta + H) > w - f(\Lambda a + K + H),$$

i.e.

$$\Delta > f(\Lambda a + K + \delta + H) - f(\Lambda a + K + H) = \delta f'(\Lambda a + K + H)$$

since δ is small. We have

$$\Delta = E[f(\Lambda r + K + \delta)] - E[f(\Lambda r + K)] = \delta E[f'(\Lambda r + K)]$$

and so we require

$$E[f'(\Lambda r + K)] > f'(\Lambda a + K + H).$$

Since v is concave, f is convex and so $f'' > 0$. Since also $H < 0$, it is sufficient to show that $E[f'(\Lambda r + K)] \geq f'(\Lambda a + K)$. Thus, it is sufficient for f' to be weakly convex (i.e. $f''' \geq 0$) for the risk premium to be increasing in the wage; multiplicative preferences are not necessary.

Proof that the market equilibrium is constrained-efficient

We prove that if a social planner faces the same informational constraints as the agents in the model (in particular, she cannot observe CEO effort), she cannot find a Pareto-dominant allocation. Using the same argument as in the “Proof that maximum effort is optimal if s is sufficiently large”, for a given CEO-firm pair, the social planner wishes the CEO to exert maximum effort (because the benefits of effort are sufficiently large) and seeks the cheapest contract that implements this effort level. We have shown that this is the contract in Proposition 1. Given this, the market assignment is Pareto optimal, as is well-known (see, e.g., Gretsky, Ostroy and Zame (1999)). For completeness, we provide a proof sketch.

Consider two firms a and b , who are matched in a decentralized equilibrium with two CEOs, a and b . We normalize $\gamma = 1$ and define $A_i = e^{x_i}$. Thus, if firm i hires CEO j , it must pay an effective wage v_j and a dollar wage $A_i v_j$. Since firm a appoints CEO a rather than CEO b :

$$S_a T_a - v_a A_a \geq S_a T_b - v_b A_a \tag{35}$$

and likewise because firm b appoints CEO b rather than CEO a :

$$S_b T_b - v_b A_b \geq S_b T_a - v_a A_b \tag{36}$$

We study whether the social planner can achieve a Pareto improvement, i.e. increase total production $\sum T_i S_j$ net of wages, subject to each CEO j receiving a utility at least v_j , the utility given by the market outcome. If the planner pairs firm a with CEO b , while paying CEO a at least $v_a A_b$ (and doing the symmetrical arrangement for firm b and CEO a), the surplus he achieves is $S_a T_b + S_b T_a - v_a A_b - v_b A_a$. By adding (35) and (36), this is weakly less than the initial surplus, $S_a T_a + S_b T_b - v_a A_a - v_b A_b$. This completes the proof.

Proof of Proposition 2

We start with L_{RA} . We have

$$\begin{aligned} L_{RA} &= \int \left[w(n) - w(n) e^{-\chi'_n} \right] dn \\ &= W - \int w(n) e^{-\chi'_n} dn \\ &= W - \left(\int w(n) dn \right) E \left[e^{-\chi'_n} \right], \end{aligned}$$

since χ'_n and $w(n)$ are independent. This yields $L_{RA} = WE [1 - e^{-\chi'_n}]$.

Turning to L_{Alloc} , as we have shown in the full proof of Theorem 1, a firm with size rank n hires a manager with talent rank $ne^{\kappa_n - \bar{\kappa}}$. Thus, $\hat{T}(n) = T(ne^{\kappa_n - \bar{\kappa}})$ and the value loss is

$$L_{Alloc} = \int CS(n)^\gamma (T(n) - T(ne^{\kappa_n - \bar{\kappa}})) dn.$$

Since $T(n) = T_{\max} - \frac{B}{\beta} n^\beta$, we have

$$L_{Alloc} = \int CS(n)^\gamma \frac{B}{\beta} n^\beta (e^{\beta(\kappa_n - \bar{\kappa})} - 1) dn.$$

As the κ_n are drawn independently of $S(n)$, we have:

$$L_{Alloc} = \left(\int CS(n)^\gamma B n^\beta dn \right) \frac{1}{\beta} E [e^{\beta(\kappa_n - \bar{\kappa})} - 1].$$

Next, we observe that $Bn^\beta = -T'(n)n$, and so

$$\begin{aligned} \int_0^N CS(n)^\gamma B n^\beta dn &= - \int_0^N CS(n)^\gamma T'(n) n dn \\ &= - \int_0^N w'(n) n dn \text{ by (34)} \\ &= [- (w(n) - w(N)) n]_0^N + \int_0^N (w(n) - w(N)) dn = W \\ \int_0^N CS(n)^\gamma B n^\beta dn &= \int_0^N (w(n) - w(N)) dn = W. \end{aligned} \tag{37}$$

This yields

$$L_{Alloc} = \frac{1}{\beta} E [e^{\beta(\kappa_n - \bar{\kappa})} - 1] W.$$

With small distortions, we can take Taylor expansions. As only $\kappa_n - \bar{\kappa}$ matters, it is sufficient to consider the case $E[\kappa_n] = 0$. The definition of $\bar{\kappa}$ gives:

$$e^{-\bar{\kappa}} = E [e^{-\kappa_n}] = E \left[1 - \kappa_n + \frac{\kappa_n^2}{2} + o(\kappa_n^2) \right] = 1 + \frac{\text{var}(\kappa_n)}{2} + o(\text{var}(\kappa_n)),$$

$$\bar{\kappa} = \frac{-\text{var}(\kappa_n)}{2} + o(\text{var}(\kappa_n)).$$

We next have

$$\begin{aligned} E[e^{\beta(\kappa_n - \bar{\kappa})} - 1] &= E\left[\beta(\kappa_n - \bar{\kappa}) + \frac{\beta^2}{2}(\kappa_n^2 - 2\kappa_n\bar{\kappa} + \bar{\kappa}^2)\right] \\ &= -\beta\bar{\kappa} + \frac{\beta^2}{2}E[\kappa_n^2] + o(\text{var}(\kappa_n)) \\ &= \frac{\text{var}(\kappa_n)}{2}(\beta + \beta^2) + o(\text{var}(\kappa_n)), \end{aligned}$$

and hence

$$L_{Alloc} \sim (1 + \beta) \frac{\text{var}(\kappa_n)}{2} W = \frac{(1 + \beta) \text{var}(\chi'_n)}{2\alpha^2\gamma^2} W.$$

Proof of Proposition 3

Using again $\alpha' = \alpha\gamma$, we observe that (37) yields:

$$\begin{aligned} W &= \int_0^N CS(n)^\gamma Bn^\beta dn = \int_0^N A^\gamma BCn^{-\alpha\gamma+\beta} dn \\ &= A^\gamma BC \frac{N^{1-\alpha'+\beta}}{1-\alpha'+\beta}. \end{aligned}$$

Given $T(n) = T_{\max} - \frac{B}{\beta}n^\beta$, we have

$$\bar{T} = \frac{1}{MN} \int_0^{MN} T(n) dn = T_{\max} - \frac{B}{\beta(1+\beta)} (MN)^\beta.$$

Thus:

$$\begin{aligned} L_{Rand} &= \int_0^N CS(n)^\gamma T(n) - \int_0^N CS(n)^\gamma \bar{T} dn \\ &= \int_0^N CS(n)^\gamma \left(\frac{B}{\beta(1+\beta)} M^\beta N^\beta - \frac{B}{\beta} n^\beta \right) dn \\ &= \int_0^N \frac{A^\gamma BC}{\beta} n^{-\alpha\gamma} \left(\frac{1}{1+\beta} M^\beta N^\beta - n^\beta \right) dn \\ &= \frac{A^\gamma BC}{\beta} \left[\frac{n^{1-\alpha'}}{1-\alpha'} \frac{M^\beta N^\beta}{1+\beta} - \frac{n^{1-\alpha'+\beta}}{1-\alpha'+\beta} \right]_0^N \\ &= \frac{A^\gamma BC}{\beta} \frac{N^{1-\alpha'+\beta}}{(1-\alpha'+\beta)} \left[\frac{1-\alpha'+\beta}{(1-\alpha')(1+\beta)} M^\beta - 1 \right] \\ &= \frac{W}{\beta} \left[\frac{1-\alpha'+\beta}{(1+\beta)(1-\alpha')} M^\beta - 1 \right]. \end{aligned}$$

Proof of Proposition 5

(i) In the corporate sector, χ_n increases by ξ , while it remains constant in the non-corporate sector. $\bar{\chi}$ thus changes to:

$$\begin{aligned}\bar{\chi}' &= -\alpha\gamma \ln E \left[e^{-\tilde{\chi}'_n/(\alpha\gamma)} \right] \\ &= -\alpha\gamma \ln E \left[e^{-\tilde{\chi}_n/(\alpha\gamma)} \left((1-\pi) e^{-\xi/(\alpha\gamma)} + \pi \right) \right] \\ &= -\alpha\gamma \ln E \left[e^{-\tilde{\chi}'_n/(\alpha\gamma)} \right] - \alpha\gamma \ln \left((1-\pi) e^{-\xi/(\alpha\gamma)} + \pi \right) \\ &= \bar{\chi} - \alpha\gamma \ln \left((1-\pi) e^{-\xi/(\alpha\gamma)} + \pi \right).\end{aligned}$$

Therefore, in the limit of small ξ , we have

$$\bar{\chi}' = \bar{\chi} + (1-\pi)\xi + O(\xi^2).$$

From equation (12), the wage in the corporate sector changes by:

$$\begin{aligned}\Delta \ln w_n &= \frac{\beta}{\alpha\gamma} (\Delta\chi_n - \Delta\bar{\chi}) \\ &= \frac{\beta}{\alpha\gamma} [\xi - (1-\pi)\xi] + O(\xi^2) \\ &= \frac{\beta}{\alpha\gamma} \pi\xi + O(\xi^2).\end{aligned}$$

(ii) Given the Pareto firm size distribution $S(n) = An^{-\alpha}$, the number of firms with a size greater than S is $KS^{-1/\alpha}$ for a constant $K = NA^{1/\alpha}$.²³ We normalize the initial χ to 0. For a non-corporate firm, the effective size equals its actual size. Given the increase in disutility, a corporate firm with effective size S has real size $Se^{\xi/\gamma}$. Thus, the probability that a firm has an effective size greater than S is:

$$(1-\pi)K(Se^{\xi/\gamma})^{-1/\alpha} + \pi KS^{-1/\alpha}.$$

Thus, the talent corresponding to a firm with effective size S is:

$$n' = n + \Delta n = KS^{-1/\alpha} \left(1 - \frac{(1-\pi)\xi}{\alpha\gamma} \right) + O(\xi^2).$$

Hence, a corporate firm of size S and thus effective size $Se^{-\xi/\gamma}$ hires a manager of talent (for

²³The proof is thus. $S = An^{-\alpha}$ implies $n = (S/A)^{-1/\alpha}$ and thus $n/N = (S/A)^{-1/\alpha}/N$. The left-hand side n/N is the number of firms larger than S , and the right-hand side can be rewritten $KS^{-1/\alpha}$.

small ξ):

$$\begin{aligned}
n' &= K (S e^{-\xi/\gamma})^{-1/\alpha} \left(1 - \frac{(1-\pi)\xi}{\alpha\gamma} \right) \\
&= K S^{-1/\alpha} \left(1 + \frac{\pi}{\alpha\gamma} \xi \right) \\
&= n \left(1 + \frac{\pi}{\alpha\gamma} \xi \right).
\end{aligned}$$

(iii) The value created by the corporate sector is given by $X = \int CS^\gamma(n) T(n) dn$, for n in the corporate sector. The loss of value creation in the corporate sector is:

$$\begin{aligned}
-\Delta X &= \int CS^\gamma(n) T(n) dn - \int CS^\gamma(n) T(n + \Delta n) dn \\
&= - \int CS^\gamma(n) T'(n) n \frac{\pi\xi}{\alpha\gamma} dn + O(\xi^2) \\
&= \frac{\pi\xi}{\alpha\gamma} W + O(\xi^2),
\end{aligned}$$

since (37) showed that $-\int_0^N CS(n)^\gamma T'(n) n dn = W$.

Proof of Theorem 2

We define $V(r_1, r_2) = v(c(r_1, r_2))$. At $t = 2$, the IC condition is:

$$\bar{a} \in \arg \max_{a_2} u(V(r_1, a_2 + \sigma(a_1)\eta_2 + \mu(a_1)) - g_1(a_1) - g_2(a_2)).$$

Note that there is no expectations operator here, since all noise has been realized when the CEO chooses a_2 – this highlights the role of our timing assumption in achieving tractability. We can thus remove $u(\cdot)$ to yield:

$$\bar{a} \in \arg \max_{a_2} V(r_1, a_2 + \sigma(a_1)\eta_2) + \mu(a_1) - g_1(a_1) - g_2(a_2)$$

The first order condition is:

$$\frac{\partial}{\partial r_2} V(r_1, r_2) - g_2'(\bar{a}) = 0,$$

which integrates to:

$$V(r_1, r_2) = K(r_1) + g_2'(\bar{a}) r_2. \quad (38)$$

for some function $K(r_1)$ to be determined.

We now consider the $t = 1$ IC constraint:

$$\bar{a} \in \arg \max_{a_2} E_1 [u(K(a_1 + \eta_1) + \Lambda_2(\bar{a} + \sigma(a_1)\eta_2 + \mu(a_1)) - g_1(a_1) - g_2(\bar{a}))] \quad (39)$$

where E_1 is the expectation conditional on η_1 . From $u(x) = e^{(1-\Gamma)x}/(1-\Gamma)$, we have the following certainty equivalent formula for any constant x and random variable \tilde{y} :

$$E[u(x + \tilde{y})] = u\left(x + \frac{1}{1-\Gamma} \ln E[e^{(1-\Gamma)\tilde{y}}]\right).$$

Applying this to (39) yields

$$\bar{a} \in \arg \max_{a_1} u\left(K(a_1 + \eta_1) + \Lambda_2 \bar{a} + \frac{1}{1-\Gamma} \ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2 + \mu(a_1))}]\right) - g_1(a_1) - g_2(\bar{a})$$

for any η_1 . As above, we can remove the u function to yield:

$$\bar{a} \in \arg \max_{a_1} K(a_1 + \eta_1) + \Lambda_2 \bar{a} + \frac{1}{1-\Gamma} \ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2 + \mu(a_1))}] - g_1(a_1) - g_2(\bar{a})$$

Hence we must have:

$$K'(r_1) + \frac{d}{da_1} \frac{1}{1-\Gamma} \ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2 + \mu(a_1))}] - g_1'(a_1) \Big|_{a_1=\bar{a}} = 0,$$

which implies:

$$K(r_1) = K_0 + \Lambda_1 r_1$$

for some constant K_0 , and

$$\begin{aligned} \Lambda_1 &= g_1'(\bar{a}) - \frac{1}{1-\Gamma} \frac{d}{da_1} \ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2 + \mu(a_1))}] \Big|_{a_1=\bar{a}} \\ &= g_1'(\bar{a}) - \Lambda_2 \mu'(a_1) - \frac{1}{1-\Gamma} \frac{d}{da_1} \ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2)}] \Big|_{a_1=\bar{a}}. \end{aligned}$$

Combining this with (38) yields

$$V(r_1, r_2) = K(r_1) + \Lambda_2 r_2 = K_0 + \Lambda_1 r_1 + \Lambda_2 r_2,$$

which generates the contract in Theorem 2.

Note that the above proof considers contracts that are message-free, deterministic and differentiable. The techniques in EG formally prove that the optimal contract satisfies all of these criteria.

In the limit of small noises, we have:

$$\ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2)}] \sim \frac{1}{2} (1-\Gamma)^2 (\Lambda_2)^2 \sigma(a_1)^2$$

and

$$\frac{1}{1-\Gamma} \frac{d}{da_1} \ln E[e^{(1-\Gamma)\Lambda_2(\sigma(a_1)\eta_2)}] \Big|_{a_1=a_1^*} \sim (1-\Gamma) (\Lambda_2)^2 \sigma(a_1) \sigma(a_1)'$$

References

- [1] Acharya, Viral, Marc Gabarro and Paolo Volpin (2010): “Competition for Managers, Corporate Governance and Incentive Compensation.” Working paper, New York University
- [2] Antràs, Pol, Luis Garicano and Esteban Rossi-Hansberg (2006): “Offshoring in a Knowledge Economy.” *Quarterly Journal of Economics* 121, 31-77
- [3] Axelson, Ulf and Philip Bond (2009): “Investment Banking (and Other High Profile) Careers.” Working paper, Stockholm School of Economics.
- [4] Baker, George P. (1992): “Incentive Contracts and Performance Measurement.” *Journal of Political Economy* 100, 598-614.
- [5] Bandiera, Oriana, Luigi Guiso, Andrea Prat and Raffaella Sadun (2010): “Matching Firms, Managers, and Incentives.” Working paper, London School of Economics
- [6] Baranchuk, Nina, Glenn Macdonald and Jun Yang (2009): “The Economics of Super Managers.” Working paper, University of Texas at Dallas.
- [7] Bebchuk, Lucian and Jesse Fried (2004): “Pay Without Performance: The Unfulfilled Promise of Executive Compensation.” Harvard University Press, Cambridge
- [8] Bennedsen, Morten, Francisco Perez-Gonzalez and Daniel Wolfenzon (2009): “Do CEOs Matter?” Working paper, Copenhagen Business School.
- [9] Biais, Bruno, Thomas Mariotti, Guillaume Plantin and Jean-Charles Rochet (2007): “Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications.” *Review of Economic Studies* 74, 345-390.
- [10] Chang, Candie, Sudipto Dasgupta and Gilles Hilary (2010): “CEO Ability, Pay, and Firm Performance.” Working paper, Hong Kong University of Science and Technology
- [11] Cooley, Thomas and Edward Prescott (2005): “Economic Growth and Business Cycles,” in Thomas Cooley, ed., *Frontiers in Business Cycle Research*. Princeton: Princeton University Press.
- [12] Core, John E. and Wayne R. Guay (1999): “The Use of Equity Grants to Manage Optimal Equity Incentive Levels.” *Journal of Accounting and Economics* 28, 151–84.
- [13] Core, John E., Wayne R. Guay and David F. Larcker (2003): “Executive Equity Compensation and Incentives: A Survey.” *Federal Reserve Bank of New York Economic Policy Review* 9, 27-50.

- [14] Core, John E., Wayne R. Guay, and Robert E. Verrecchia (2003): “Price Versus Non-Price Performance Measures in Optimal CEO Compensation Contracts.” *The Accounting Review* 78, 957-981.
- [15] DeMarzo, Peter M. and Michael J. Fishman (2007): “Optimal Long-Term Financial Contracting.” *Review of Financial Studies* 20, 2079-2127.
- [16] Demsetz, Harold and Kenneth Lehn (1985): “The Structure of Corporate Ownership: Causes and Consequences”. *Journal of Political Economy* 93, 1155-1177.
- [17] Dicks, David L. (2009): “Executive Compensation, Incentives, and the Role for Corporate Governance Regulation.” Working paper, University of North Carolina.
- [18] Dittmann, Ingolf, and Ernst Maug (2007): “Lower Salaries and No Options? On the Optimal Structure of Executive Pay.” *Journal of Finance* 62, 303-343.
- [19] Dittmann, Ingolf and Ko-Chia Yu (2009): “How Important are Risk-Taking Incentives in Executive Compensation?” Working paper, Erasmus University Rotterdam.
- [20] Edmans, Alex and Xavier Gabaix (2009): “Tractability in Incentive Contracting.” NBER Working paper No. 15545.
- [21] Edmans, Alex, Xavier Gabaix and Augustin Landier (2009): “A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium.” *Review of Financial Studies* 22, 4881-4917.
- [22] Eisfeldt, Andrea and Camelia M. Kuhnen (2009): “CEO Turnover in a Competitive Assignment Framework.” Working paper, Northwestern University.
- [23] Gabaix, Xavier and Augustin Landier (2008): “Why Has CEO Pay Increased So Much?” *Quarterly Journal of Economics* 123, 49-100.
- [24] Galichon, Alfred and Bernard Salanié (2009): “Matching With Trade-Offs: Revealed Preferences Over Competing Characteristics.” Working paper, Ecole Polytechnique
- [25] Garen, John E. (1994): “Executive Compensation and Principal-Agent Theory.” *Journal of Political Economy*, 102, 1175-1199.
- [26] Gayle, George-Levi and Robert Miller (2009): “Has Moral Hazard Become a More Important Factor in Managerial Compensation?” *American Economic Review* 99, 1740-69.
- [27] Gibbons, Robert and Kevin J. Murphy (1992): “Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence.” *Journal of Political Economy* 100, 468-505.

- [28] Gretsky, Neil E., Joseph M. Ostroy and William R. Zame (1999): “Perfect Competition in the Continuous Assignment Model.” *Journal of Economic Theory* 88, 60-118.
- [29] Hall, Brian J. and Jeffrey B. Liebman (1998): “Are CEOs Really Paid Like Bureaucrats?” *Quarterly Journal of Economics*, 113, 653-691.
- [30] Harris, Milton and Artur Raviv (1979): “Optimal Incentive Contracts With Imperfect Information.” *Journal of Economic Theory* 20, 231-259.
- [31] Hermalin, Benjamin E. (2005): “Trends in Corporate Governance.” *Journal of Finance* 60, 2351-2384.
- [32] Himmelberg, Charles P., and R. Glenn Hubbard (2000): “Incentive Pay and the Market for CEOs: An Analysis of Pay-for-Performance Sensitivity.” Working paper, Columbia University.
- [33] Holmstrom, Bengt and Paul Milgrom (1987): “Aggregation and Linearity in the Provision of Intertemporal Incentives.” *Econometrica* 55, 308-328.
- [34] Ittner, Christopher D., Richard A. Lambert and David F. Larcker (2003): “The Structure and Performance Consequences of Equity Grants to Employees of New Economy Firms.” *Journal of Accounting and Economics* 34, 89–127.
- [35] Jensen, Michael C. and Kevin J. Murphy (2004): “Remuneration: Where We’ve Been, How We Got to Here, What are the Problems, and How to Fix Them.” Working paper, Harvard University.
- [36] Kihlstrom, Richard E. and Jean-Jacques Laffont (1979): “A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion.” *Journal of Political Economy* 87, 719-748.
- [37] Lambert, Richard A. (1986): “Executive Effort and Selection of Risky Projects.” *Rand Journal of Economics* 17, 77-88.
- [38] Llense, Fabienne (2009): “French CEO Compensations: What is the Cost of a Mandatory Upper Limit?” Working paper, University of Paris.
- [39] Malmendier, Ulrike and Geoffrey Tate (2009): “Superstar CEOs.” *Quarterly Journal of Economics* 124, 1593-1638.
- [40] Murphy, Kevin J. (1999): “Executive Compensation” in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3b. New York and Oxford: Elsevier/North-Holland, 2485-2563.

- [41] Murphy, Kevin J. (2003): “Stock-Based Pay in New Economy Firms.” *Journal of Accounting and Economics* 34, 129–147.
- [42] Nguyen, Bang Dang and Kasper Meisner Nielsen (2010): “What Death Can Tell: Are Executives Paid for Their Contributions to Firm Value?” Working paper, Chinese University of Hong Kong
- [43] Ou-Yang, Hui (2003): “Optimal Contracts in a Continuous-Time Delegated Portfolio Management Problem.” *Review of Financial Studies* 16, 173-208.
- [44] Oyer, Paul and Scott Schaefer (2004): “Why Do Some Firms Give Stock Options To All Employees? An Empirical Examination of Alternative Theories.” *Journal of Financial Economics* 76, 99-133.
- [45] Peters, Florian S. (2009): “Risk Premia in Executive Compensation: A Life-Cycle Perspective.” Working paper, Swiss Finance Institute.
- [46] Peters, Florian S. and Alexander F. Wagner (2009): “The Executive Turnover Risk Premium.” Working paper, Swiss Finance Institute.
- [47] Plehn-Dujowich, Jose M. and Ajay Subrahmaniam (2009): “Monopolistic Competition, Managerial Incentives, and the Distribution of Firms in General Equilibrium.” Working paper, Temple University.
- [48] Sappington, David (1983): “Limited Liability Contracts Between Principal and Agent.” *Journal of Economic Theory* 29, 1-21.
- [49] Sattinger, Michael (1993): “Assignment Models of the Distribution of Earnings.” *Journal of Economic Literature* 31, 831-880.
- [50] Sung, Jaeyoung (1995): “Linearity with Project Selection and Controllable Diffusion Rate in Continuous-Time Principal-Agent Problems.” *RAND Journal of Economics* 26, 720-743.
- [51] Terviö, Marko (2008): “The Difference that CEOs Make: An Assignment Model Approach.” *American Economic Review* 98, 642-668.
- [52] Tsuyuhara, Kunio (2009): “Repeated Moral Hazard With Worker Mobility Via Directed On-The-Job Search.” Working paper, University of Toronto.