ecgi

How Important are Risk-Taking Incentives in Executive Compensation?

Ingolf Dittmann
Erasmus University Rotterdam and ECGI

Ko-Chia Yu
Shanghai University of Finance and Economics

Dan Zhang
Norwegian Business School

This paper can be downloaded without charge from:
http://ssrn.com/abstract_id=1176192

www.ecgi.org/wp

How Important are Risk-Taking Incentives in Executive Compensation?

Ingolf Dittmann
Ko-Chia Yu
Dan Zhang

Abstract

We consider a model in which shareholders provide a risk-averse CEO with risk-taking incentives in addition to effort incentives. We show that the optimal contract protects the CEO from losses for bad outcomes, is convex for medium outcomes, and concave for good outcomes. We calibrate the model to data on 1,707 CEOs and show that it explains observed contracts much better than the standard model without risk-taking incentives. An application to contracts that consist of base salary, stock, and options yields that options should be issued in the money. Our model also helps us rationalize the universal use of at-the-money options when the tax code is taken into account. Moreover, we propose a new measure of risk-taking incentives that trades off the expected value added to the firm and the additional risk a CEO has to take.

Ingolf Dittmann*
Professor of Finance
Erasmus University Rotterdam, School of Economics
P.O. Box 1738
3000 DR, Rotterdam, The Netherlands
phone: +31 10 4081283
e-mail: dittmann@ese.eur.nl

Ko-Chia Yu
Assistant Professor of Finance
Shanghai University of Finance and Economics, School of Finance
100 Wudong Rd.
Shanghai, China 200433
phone: +86 21 65908373
e-mail: KoChia.Yu@gmail.com

Dan Zhang
Assistant Professor of Finance
Norwegian Business School
Nydalsveien 37
0442 Oslo, Norway
phone: +47 46 410927
e-mail: danielle.zhang@bi.no

*Corresponding Author

How Important are Risk-Taking Incentives in Executive Compensation?*

Ingolf Dittmann[†]   Ko-Chia Yu[‡]   Dan Zhang[§]

July 22, 2016

Abstract

We consider a model in which shareholders provide a risk-averse CEO with risk-taking incentives in addition to effort incentives. We show that the optimal contract protects the CEO from losses for bad outcomes, is convex for medium outcomes, and concave for good outcomes. We calibrate the model to data on 1,707 CEOs and show that it explains observed contracts much better than the standard model without risk-taking incentives. An application to contracts that consist of base salary, stock, and options yields that options should be issued in the money. Our model also helps us rationalize the universal use of at-the-money options when the tax code is taken into account. Moreover, we propose a new measure of risk-taking incentives that trades off the expected value added to the firm and the additional risk a CEO has to take.

JEL Classifications: G30, M52

Keywords: Stock Options, Effort Aversion, Executive Compensation, Risk Aversion, Risk-Taking Incentives, Optimal Strike Price

[†]Corresponding author: Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands. Email: dittmann@ese.eur.nl. Tel: +31 10 408 1283.

[‡]School of Finance, Shanghai University of Finance and Economics, 100 Wudong Rd., Shanghai, China 200433. Email: KoChia.Yu@gmail.com. Tel: +86 21 65908373.

[§]BI Norwegian Business School, Nydalsveien 37, 0442 Oslo, Norway. Email: danielle.zhang@bi.no. Tel: +47 4641 0927

1 Introduction

Can the inclusion of risk-taking incentives in the standard model of executive compensation rationalize observed compensation practice? Hall and Murphy (2002) and Dittmann and Maug (2007) demonstrate that the standard Holmström (1979) model cannot explain the observed compensation contracts. In this paper, we show that including risk-taking incentives in the Holmström (1979) model better fits the observed contract empirically. Specifically, we assume that shareholders take into account both effort incentives and risk-taking incentives when designing the compensation contract. Our model predicts similar patterns as in the observed compensation contracts that emphasize "carrots" over "sticks": Firms pay a flat wage for large stock price decreases and provide incentives only for medium and high stock price ranges.

Risk-taking incentives are important in CEO compensation contracts, because equity compensation exposes CEOs to firm-specific risk. Risk-averse CEOs will want to reduce the firm risk even if this destroys value. Therefore, we need risk-taking incentives to induce the CEO to take risks that benefit well-diversified shareholders (Smith and Stulz (1985) and Haugen and Senbet (1981)). Indeed, empirical evidence suggests that risk-taking incentives matter for CEOs' actual risk-taking (see, for example, Low (2009), Knopf, Nam, and Thornton (2002), Coles, Daniel, and Naveen (2006), and Acharya, Amihud, and Litov (2011)).

In our model, the CEO does not only exert costly effort but also determines the firm's strategy. We capture these dimensions by assuming that the CEO affects both the mean and the volatility of future firm value. If the contract does not provide sufficient risk-taking incentives, the risk-averse CEO chooses a strategy that avoids risk and depresses the firm value. The best way for shareholders to mitigate this inefficiency is to provide both effort and risk-taking incentives by rewarding good outcomes and not punishing bad outcomes. While high stock price realizations are a clear good signal, low stock price realizations are ambiguous: they can be indicative of low effort (which is bad) or of extensive risk-taking (which is good, given that the CEO leans towards inefficiently low risk).

The optimal contract in our model differs markedly from the one in the standard model without risk-taking incentives. As shown in Dittmann and Maug (2007), the standard model predicts a concave optimal contract that emphasizes "sticks" over "carrots", featuring large penalties for stock price decreases and small gains for stock price increases. The result is driven by the decreasing marginal utility so that it is inefficient to make the CEO pay very sensitive to performance at high levels of wealth. By comparison, our model predicts similar patterns as in the observed compensation contracts that emphasize "carrots": Firms pay a flat wage for poor performance, a convex wage for

medium performance, i.e., increasing wealth for higher stock prices, and a concave wage for high performance. This result is driven by two forces. First, the risk-taking incentives are provided to a risk averse agent by making the contract more convex for medium outcomes (see Ross (2004)). Second, a decreasing marginal utility leads to the contract being concave for high outcomes.

We calibrate our model to a sample of 1,707 U.S. CEOs and generate the optimal compensation contract for each individual. Then, we compare optimal contracts to observed contracts and find that our model can explain observed contracts much better than the standard model without risk-taking incentives. In particular, the average distance, i.e., the expected absolute value between the observed contract and the optimal contract, is 5.4% for our model as compared to 16.1% for the model without risk-taking incentives.

We also apply our model to contracts that consist of base salary, stock, and options and establish that in-the-money options are preferable to the portfolio of stock and at-the-money options that we observe in practice. In our sample, the median strike price should be 55.4% of the firm's stock price when issued. Compared to the observed portfolio contract, this in-the-money option contract provides higher incentives at the center of the distribution and lower incentives in the tails of the distribution. If we take into account the tax penalties that apply to in-the-money options in the U.S., we achieve optimality of the observed portfolio contract for a majority of the CEOs in our sample. Therefore, the universal use of at-the-money options, which is often seen as evidence for managerial rent-extraction (see Bebchuk and Fried (2004)), is consistent with efficient contracting if the tax code is taken into consideration.[1]

This paper makes several contributions to the literature. First, while it has been known for some time that risk-taking incentives can explain convex contracts, we are the first to calibrate such a model.[2] We bridge the gap between theoretical and empirical research by testing the quantitative,

---

[1]We are not the first to show that at-the-money options can be part of the optimal contract. Specifically, Hall and Murphy (2000) already make this point for a partial principal-agent model. We generalize their argument. We solve a complete principal-agent model and calibrate it to the data.

[2]Lambert (1986) and Core and Qian (2002) consider discrete volatility choices, where the agent must exert effort to gather information about investment projects. Feltham and Wu (2001) and Lambert and Larcker (2004) assume that the agent's choice of effort simultaneously affects the mean and the variance of the firm value distribution, so they reduce the two-dimensional problem to a one-dimensional problem. Two other papers (and our model) work with continuous effort and volatility choice: Hirshleifer and Suh (1992) analyze a rather stylized principal-agent model and solve it for special cases. Flor, Frimor, and Munk (2014) consider a similar model to ours but they work with the assumption that stock prices are normally distributed while we work with the lognormal distribution. Hellwig (2009), Sung (1995), and Ou-Yang (2003) solve models with continuous effort and volatility choice, but Hellwig (2009) assumes that the agent is risk-neutral and Sung (1995) assumes that the principal can observe (and effectively set) volatility. Ou-Yang (2003) considers delegated portfolio management and assumes that the principal can infer what the portfolio value would have been if the optimal strategy had been implemented; in our model, the principal does not know this benchmark. Manso (2011) considers a class of Bayesian decision models which make the agent uncertain about the true distribution of payoffs of the available actions. He also establishes that optimal contracts must not punish bad outcomes when risk-taking (innovation) needs to be encouraged. None of these papers have calibrated their models.

and not just the qualitative, implications of different models. This calibration also contributes to the recent literature on the calibrations of contracting models.[3] Second, we propose a new risk-taking incentives measure that better describes the trade-off between the expected firm value and the additional risk a CEO has to take. Empirical studies usually measure risk-taking incentives as "vega", i.e., the change in the manager's wealth with respect to the change in the firm's stock return volatility. However, the effect of "vega" can be mitigated by high "delta", i.e., the change in the manager's wealth with respect to the change in the firm's stock price. Our measure, called *risk avoidance*, combines both "vega" and "delta". Third, we provide an alternative approach to the empirical literature that suffers from endogeneity, where firm risk and managerial incentives are simultaneously determined in the compensation design. We model the endogeneity directly and demonstrate that the provision of risk-taking incentives is consistent with efficient contracting. Fourth, our setting captures a multitasking problem where a CEO exerts costly effort and determines the firm's volatility. The principal takes into account how the incentives to undertake one task affect the incentives to undertake other tasks.[4]

We acknowledge that alternative explanations may account for the convexity in the observed contract.[5] The only alternative model that can be readily calibrated to the data is Dittmann, Maug, and Spalt (2010) where CEOs are assumed to be loss-averse. We calibrate this model to our data and show that our model is more robust than the loss-aversion model to changes in the preference parameters. As a further robustness check, we introduce the threat of dismissal into the CEO's wealth contract and show that omitting CEO dismissals biases our risk avoidance measure downwards. We discuss other limitations of our model and offer several conjectures on how the optimal contract can change when dynamic elements, such as gradual vesting, new grants, and contract renegotiation, are introduced.

---

[3]See Dittmann and Maug (2007), Gabaix and Landier (2008), Edmans, Gabaix, and Landier (2009), Dittmann, Maug, and Spalt (2010), and Dittmann, Maug, and Spalt (2013).

[4]Holmström and Milgrom (1991) model a multitasking problem where the agent needs to allocate his effort among different tasks. They show that raising effort on one task raises the marginal cost of effort on the other task. Our model allows the agent to exert costly effort to affect the mean and costless effort to affect the volatility of the stock price. Dewatripont and Tirole (1999) model a direct conflict between tasks where two agents are hired to search for information about the pros and cons of a decision. In our model, there is no direct conflict between the agent's influence on the mean and volatility of the stock price.

[5]Inderst and Müller (2005) explain options as instruments that provide outside shareholders with better liquidation incentives. Edmans and Gabaix (2011) and Edmans et al. (2012) show that convex contracts can arise in dynamic contracting models. Peng and Röell (2014) analyze stock price manipulations in a model with multiplicative CEO preferences and find convex contracts for some parameterizations. Dittmann, Maug and Spalt (2010) show that optimal contracts are convex if CEOs are loss-averse. Chaigneau and Sahuguet (2015) model indexed options as a device to retain CEOs. Innes (1990) shows that stock options can be optimal in a model with limited liability and risk neutrality of both the principal and the agent. Chaigneau (2013a) explains the structure of CEO incentive pay with decreasing relative risk aversion.

Our analysis proceeds as follows. In the next section, we present our model and derive the shape of the optimal contract. Section 3 contains the calibration method. In a nutshell, we numerically search for the cheapest contract that provides the manager with the same incentives and the same utility as the observed contract. Section 4 describes the construction of the data set. In Section 5, we present our main results. Section 6 analyzes the optimal strike price in a standard option contract. Section 7 provides robustness checks, Section 8 discusses the limitations of the model, and Section 9 concludes the paper. The appendixes collect some technical material.

2 Optimal contracting with risk-taking incentives

2.1 Model

Our model is in the spirit of Holmström (1979), i.e., there are two points in time and the principal cannot observe the agent's actions. At time $t = 0$, the contract between a risk-neutral principal (the shareholders) and a risk-averse agent (CEO) is signed, and at time $t = T$, the contract period ends. At some point during the contract period $(0, T)$, the agent simultaneously makes two choices. He chooses effort $e \in [0, \infty)$ which results in private costs $C(e)$ for the agent and which affects the firm's expected value $E(P_T)$. In addition to Holmström (1979), we explicitly allow the CEO to choose the firm's stock return volatility σ which also affects the firm's expected value $E(P_T)$. We refer to σ interchangeably as "firm risk". We follow Innes (1990) to assume that the agent can costlessly destroy output. Therefore, the wage scheme $w(.)$ must be non-decreasing.

2.1.1 Volatility

The choice of volatility can be attributed to a choice of strategy or investment.[6] We assume that volatility cannot be contracted upon as the CEO can arbitrarily inflate volatility. He could, for instance, make the firm riskier by investing free cash flows in speculative assets or by taking a short position in some risky trades without changing the firm's core strategy.[7] As a consequence, the

---

[6]We think of the strategy as a feasible combination of many different actions that affect issues including project choice, mergers and acquisitions, capital structure, and financial transactions. For instance, part of the strategy could be an R&D project that increases value and risk. Another part could be financial hedging of some input factors which reduces value and risk. Due to its richness, we do not model the agent's choice of strategy in detail. Instead, the undiversified and risk averse CEO himself is interested in a low volatility, as the disutility he suffers from taking an extra unit of risk dominates the utility he gains from the increase in wealth due to the extra unit of risk. We assume that the CEO chooses a strategy that minimizes the firm risk σ given the expected value $E(P_T)$ or, equivalently, that maximizes the expected value $E(P_T)$ given the risk σ.

[7]More strictly speaking, the observed volatility σ^{obs} is equal to the productive volatility σ which is depicted in Figure 1 and an unproductive volatility σ_1. If the volatility is omitted from the contract, $\sigma_1 = 0$ and the observed volatility σ^{obs} is equal to the productive volatility σ (which we consider in the paper). If the volatility entered into the

manager's wealth $W_T = w(P_T)$ only depends on the end-of-period stock price P_T.

We assume that there is a first-best firm volatility σ^{**} that maximizes the firm value (given effort e). If the agent wants to reduce the risk to some value below σ^{**}, he can do so in two ways. Either he drops some risky but profitable projects (e.g., an R&D project), or he takes an additional action that reduces the risk but also the profits (e.g., costly hedging). In both cases, a reduction in volatility σ leads to a reduction in firm value $E(P_T)$.

2.1.2 Production

After the contract details have been disclosed, we can write $E(P_T) = f(e, \sigma)$, where f is a production function. Therefore, we assume that $f(e, \sigma)$ is increasing and concave in σ as long as $\sigma < \sigma^{**}|e$. In the region above $\sigma^{**}|e$, the production function $f(e, \sigma)$ is weakly decreasing in σ; it is flat if the agent can take on additional risk at no costs (e.g., with financial transactions). Finally, we assume that the production function $f(e, \sigma)$ is increasing and concave in e (given volatility σ). One advantage of our model is that we do not have to assume a specific functional form for how the firm value changes with firm risk. We only need to assume that the production function is increasing and concave in risk for risk levels below first-best.

We assume that the end-of-period stock price P_T is lognormally distributed:

$$P_T(u|e, \sigma) = f(e, \sigma) \exp\left\{ -\frac{\sigma^2}{2}T + u\sqrt{T}\sigma \right\}, \quad u \sim N(0, 1). \tag{1}$$

The market value of the firm at time $t = 0$ is $P_0(e, \sigma) = E(P_T(u|e, \sigma)) \exp\{-r_f T\}$, where r_f is the risk-free rate.[8] Therefore, we can write $E(P_T) = P_0(e, \sigma) \exp\{r_f T\} = f(e, \sigma)$:

$$P_T(u|e, \sigma) = P_0(e, \sigma) \exp\left\{ (r_f - \frac{\sigma^2}{2})T + u\sqrt{T}\sigma \right\}, \quad u \sim N(0, 1). \tag{2}$$

---

contract, the agent would choose the unproductive volatility σ_1 larger than zero, more precisely, $\sigma_1 = \sigma^* - \sigma$, where σ^* is the target volatility from the shareholders' optimization problem, thereby circumventing the solution. There are more arguments for the assumption that it is not possible to contract upon volatility. For example, the stock price volatility over the period $[0; T)$ does not necessarily match the risk of the project; for instance, if the CEO is induced to be active in the M&A market which involves constantly surveying the situation and being ready for any forthcoming M&A activity. It may be that an appropriate M&A activity does not emerge until T and therefore the volatility in $[0; T)$ is low.

[8]We follow Dittmann and Maug (2007) and Dittmann, Maug, and Spalt (2010) and assume that either there is no premium for systematic risk or the firm has no exposure to systematic risk, so that the risk-free rate r_f is the appropriate stock return. This assumption allows us to abstract from the agent's portfolio problem, because in our model the only alternative to an investment in the own firm is an investment at the risk-free rate. We effectively reduce a two-dimensional problem where one innovation drives the systematic and another innovation the unsystematic risk to a one-dimensional problem. If we allowed the agent to earn a risk-premium on the shares of his firm, he could value these above their actual market price, because investing into his own firm is then the only way of earning the risk-premium. Our assumption effectively means that all risk in the model is firm-specific.

2.1.3 Utility function

The manager's utility is additively separable in wealth and effort and has a constant relative risk-aversion with parameter γ with respect to wealth W_T:

$$U(W_T, e) = V(W_T) - C(e) = \frac{W_T^{1-\gamma}}{1-\gamma} - C(e).$$ (3)

If $\gamma = 1$, we define $V(W_T) = \ln(W_T)$. The costs of effort are assumed to be increasing and convex in effort, i.e. $C'(e) > 0$ and $C''(e) > 0$. We normalize $C(0) = 0$. There is no direct cost associated with the manager's choice of volatility. Volatility σ affects the manager's utility indirectly via the stock price distribution and the utility function $V(.)$. Finally, we assume that the manager has outside employment opportunities that give him expected utility $\underline{U}$.

2.2 Optimal contract

As incentives for a risk-averse CEO are costly, shareholders implement a level of volatility $\sigma^* \leq \sigma^{**}$ as well as a given effort e^* and solve the following optimization problem:

$$\max_{W_T} E\left[P_T - W_T(P_T)|e^*, \sigma^*\right]$$ (4)

$$\text{subject to } \frac{dW_T(P_T)}{dP_T} \geq 0 \text{ for all } P_T$$ (5)

$$E\left[V(W_T(P_T))|e^*, \sigma^*\right] - C(e^*) \geq \underline{U}$$ (6)

$$\{e^*, \sigma^*\} \in \text{argmax } \{E\left[V(W_T(P_T))|e, \sigma\right] - C(e)\}$$ (7)

Hence, shareholders choose the wage schedule $W_T(P_T)$ that minimizes the contracting costs subject to three constraints: The monotonicity constraint (5), the participation constraint (6), and the incentive compatibility constraint (7). We replace (7) with its first-order conditions. Appendix A contains a discussion of the validity of the first-order approach

$$\frac{dE\left[V(W_T(P_T))\right]}{de} - \frac{dC}{de} = 0,$$ (8)

$$\frac{dE\left[V(W_T(P_T))\right]}{d\sigma} = 0.$$ (9)

We call condition (8) the effort incentive constraint and (9) the volatility incentive constraint.

Proposition 1. *(Optimal contract): The optimal contract that solves the shareholders' problem*

(4), (5), (6), (8), and (9) has the following functional form:

$$\left[\frac{dV(W_T^*)}{dW_T}\right]^{-1} = \begin{cases} c_0(\sigma) + c_1(\sigma)\ln P_T + c_2(\sigma)(\ln P_T)^2 & \textit{if } ln(P_T) > -\frac{c_1(\sigma)}{2c_2(\sigma)} \\ c_0(\sigma) - \frac{(c_1(\sigma))^2}{4c_2(\sigma)} & \textit{if } ln(P_T) \leq -\frac{c_1(\sigma)}{2c_2(\sigma)} \end{cases} \tag{10}$$

where $c_0(\sigma)$, $c_1(\sigma)$, and $c_2(\sigma)$ depend on the distribution of P_T and the Lagrange multipliers of the optimization problem, with $c_2(\sigma) > 0$. For constant relative risk aversion, we obtain

$$W_T^* = \begin{cases} \left[c_0(\sigma) + c_1(\sigma)\ln P_T + c_2(\sigma)(\ln P_T)^2\right]^{1/\gamma} & \textit{if } ln(P_T) > -\frac{c_1(\sigma)}{2c_2(\sigma)} \\ \left[c_0(\sigma) - \frac{(c_1(\sigma))^2}{4c_2(\sigma)}\right]^{1/\gamma} & \textit{if } ln(P_T) \leq -\frac{c_1(\sigma)}{2c_2(\sigma)} \end{cases} \tag{11}$$

The proof of Proposition 1 and full expressions for parameters $c_0(\sigma)$, $c_1(\sigma)$, and $c_2(\sigma)$ can be found in Appendix B. To develop an intuition for the optimal contract (11), it is instructive to first look at the optimal contract without any risk-taking incentives. This contract has the form $\widetilde{W}_T = (c_0 + c_1 \ln P_T)^{1/\gamma}$ and is globally concave as long as $\gamma \geq 1$ (see Dittmann and Maug (2007) for a problem with exogenous σ). The comparison with $W_T = \left(c_0(\sigma) + c_1(\sigma)\ln P_T + c_2(\sigma)(\ln P_T)^2\right)^{1/\gamma}$ shows that risk-taking incentives are provided by the additional quadratic term $c_2(\sigma)(\ln P_T)^2$. This term makes the contract more convex and limits its downside, two features that make risk-taking more attractive for a risk-averse agent. To satisfy the monotonicity constraint, the downward sloping part of the wage function due to the quadratic term is replaced by a flat wage. The resulting contract (11) is flat below some threshold $\widetilde{P} = \exp\{-\frac{c_1(\sigma)}{2c_2(\sigma)}\}$, increasing and convex for some region above this threshold, and finally concave, because the concavity of the logarithm dominates the convexity of the quadratic term asymptotically.

2.3 Risk-taking incentives in our model

In the empirical literature on executive compensation, risk-taking incentives are usually measured by the vega of the manager's equity portfolio, i.e., by the partial derivative of the manager's wealth with respect to his own firm's stock return volatility.[9] An exception is Lambert, Larcker, and Verrecchia (1991) who work with what we call the "utility adjusted vega", i.e., the partial derivative of the manager's expected *utility* with respect to stock return volatility. However, there is another effect of volatility on managerial utility that, to the best of our knowledge, has been ignored in the empirical literature on risk-taking incentives. If the CEO has too little incentive to take risks, a rise in volatility

---

[9]See, among others, Guay (1999), Rajgopal and Shevlin (2002), Knopf, Nam, and Thornton (2002), Habib and Ljungqvist (2005), and Coles, Daniel, and Naveen (2006).

increases the firm value and, due to the CEO's equity portfolio, also increases the managerial utility. Conversely, if he has too much risk-taking incentive, a further rise in volatility decreases the firm value and therefore also decreases the managerial utility. In this subsection, we derive this result formally from our model, and propose a new measure of risk-taking incentives that combines the two effects.

In our model, risk-taking incentives are described in the volatility incentive constraint (9). Substituting $\frac{dP_T}{d\sigma}$ from (2) and rearranging (9) yields

$$PPS^{ua}\frac{dP_0}{d\sigma} = -\nu^{ua}, \tag{12}$$

$$\text{where } PPS^{ua} := E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}\frac{P_T}{P_0}\right] \tag{13}$$

$$\text{and} \qquad \nu^{ua} := E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}P_T\left(-\sigma T + u\sqrt{T}\right)\right]. \tag{14}$$

Here, PPS^{ua} is the utility adjusted pay-for-performance sensitivity, or the utility adjusted delta, which measures how much the manager's expected utility rises for a marginal stock price increase. Likewise, ν^{ua} is the utility adjusted vega, i.e. the marginal increase in the manager's expected utility for a marginal increase in volatility - assuming that the firm value P_0 stays constant.

The first-order condition (12) equates the marginal benefits to the marginal costs of an increase in volatility from the agent's point of view. Figure 1 shows benefits and costs as a function of σ. The benefit, represented by the solid line, stems from the response of the firm value to an increase in volatility: $dP_0/d\sigma$ is increasing for low $\sigma < \sigma^{**}$ and weakly decreasing for high $\sigma > \sigma^{**}$. Consequently, $PPS^{ua}dP_0/d\sigma$ is increasing for low values of σ and decreasing for high values. The cost, represented by the dashed line, is due to the decrease in the manager's utility $-\nu^{ua}$ with volatility, as managers are assumed to be risk averse. When the volatility incentive constraint (12) is binding, the two lines cross and the two effects cancel out at σ^*. This is the point where the CEO optimally chooses the level of volatility. It lies to the left of the maximum of the solid line because the manager is risk averse as his stock and option compensation exposes the manager to firm-specific risk. If the manager were risk neutral, it would lie exactly at the maximum of the solid line, i.e., at the level that maximizes firm value.

The agent will take an action if and only if its benefits exceed its cost, i.e., if

$$PPS^{ua}\frac{dP_0}{d\sigma} > -\nu^{ua} \Leftrightarrow \frac{dP_0}{d\sigma}\frac{1}{P_0} > -\frac{\nu^{ua}}{PPS^{ua}}\frac{1}{P_0}. \tag{15}$$
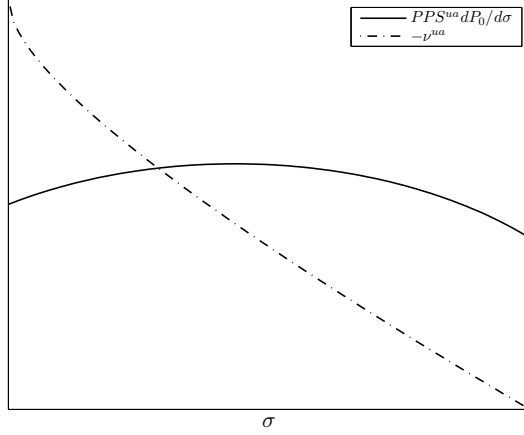
Figure 1: The figure depicts benefits ($PPS^{ua}dP_0/d\sigma$) and costs ($-\nu^{ua}$) as a function of σ (see equation (12)) for a stylized contract. PPS^{ua} is the utility adjusted pay-for-performance sensitivity, or the utility adjusted delta, which measures how much the manager's expected utility rises for a marginal stock price increase. ν^{ua} is the utility adjusted vega, which captures the marginal increase in the manager's expected utility for a marginal increase in volatility - assuming that firm value P_0 stays constant.

Therefore, we define the incentives to avoid risk as

$$\rho := -\frac{\nu^{ua}}{PPS^{ua}}\frac{1}{P_0}. \tag{16}$$

Equation (16) defines a hurdle rate: the CEO will take a new project only if it increases the firm value by ρ times the percentage increase in the firm risk. Consider a project that would increase the firm risk by one percentage point, e.g., from 30% to 31%, and let $\rho = 2$. Then, the agent will take this project only if it increases the firm value by at least 2%. All positive NPV projects that generate less than a 2% increase in firm value for each percent of additional risk will thus be passed up. On the other hand, if $\rho < 0$, the agent has incentives to take on risky projects with negative NPV. In the above example of a project that increases the firm risk by one percentage point, $\rho = -2$ means that the agent is willing to undertake this project as long as it does not destroy more than 2% of the firm value. If $\rho = 0$, the CEO is indifferent to firm risk and will therefore implement all profitable projects irrespective of their riskiness. We refer to ρ as risk avoidance, and to $-\rho$ as risk-taking incentives.

Our main conceptual result is that the utility adjusted vega alone is not the best measure of risk-taking incentives, but that it should be scaled by the utility adjusted delta. To understand why this scaling is necessary, first consider the case where vega is negative, and so the manager wishes to avoid risky, positive NPV projects. However, this effect is mitigated if the CEO has a high delta as

9

this means that he gains from taking positive NPV actions. Second, consider the case where vega is positive, and thus the manager has an incentive to take risky projects even if they are negative NPV. Once more, this effect is mitigated if the CEO has a high delta as it means that he is hurt by taking negative NPV actions. Regardless of the sign of vega, the incentives to take too little or too much risk are offset by a high delta, so the measure of risk-taking incentives depends on the ratio of vega to delta.

3   Calibration

In this section, we present formulae for the calibration of the optimal or model contract (11) to the data. We assume that shareholders want to implement a certain action $\{e^*, \sigma^*\}$ in the observed contract. We effectively suppose that the firm has already induced the optimal level of CEO effort and firm risk as these are orders of magnitude higher than the cost of incentivizing the CEO, which is then left for the calibration method to verify.[10] Under this assumption, we can reformulate the shareholder's optimization problem (4), (5), (6), (8), and (9) as follows:

$$\min_{c_0, c_1, c_2} E\left[W_T^*(P_T | c_0, c_1, c_2)\right] \tag{17}$$

$$\text{subject to } E\left[V(W_T^*(P_T | c_0, c_1, c_2))\right] \geq E\left[V(W_T^d(P_T))\right] \tag{18}$$

$$PPS^{ua}(W_T^*(P_T | c_0, c_1, c_2)) = PPS^{ua}(W_T^d(P_T)) \tag{19}$$

$$\rho(W_T^*(P_T | c_0, c_1, c_2)) = \rho(W_T^d(P_T)), \tag{20}$$

where $W_T^d(P_T) = (W_0 + \phi^d) exp(r_f T) + n_S^d P_T + n_O^d \max\{P_T - K^d, 0\}$ is the observed contract (d for "data") that we construct from the data as described in Section 4. Equations (17) to (20) can be calibrated to the data.

We derive equations (17) to (20) as follows. First, as the principal is risk-neutral, it does not matter if he maximizes (4) or minimizes (17). Second, we rewrite the effort incentive constraint (8) so that the left-hand side of the equation does not contain any quantities that we cannot compute and the right-hand side does not contain the wage function:

$$PPS^{ua}(W_T(P_T)) = E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_0}\right] = \frac{C'(e)}{\frac{dP_0}{de}} \tag{21}$$

---

[10]This calibration method was first been used by Dittmann and Maug (2007). It corresponds to the first stage of the two-stage procedure in Grossman and Hart (1983). We cannot repeat this task for alternative effort/volatility levels, because this would require knowledge of the production and cost functions. Therefore, we cannot analyze the optimal level of effort or volatility, which is the second stage in Grossman and Hart (1983).

Under the hypothesis that the model is descriptive of the data (i.e., the optimal contract fulfills all the incentive constraints and the participation constraints), the effort incentive constraint in our calibration problem can be written as (19). Third, for the volatility incentive constraint (9), equations (15) and (16) imply

$$\rho(W_T(P_T)) = \frac{dP_0}{d\sigma^{obs}} \frac{1}{P_0}. \tag{22}$$

Note that this equation once more separates quantities that we cannot compute $(dP_0/d\sigma^{obs})$ from quantities that depend on the shape of the optimal contract (ρ). Therefore, we likewise obtain (20). Fourth, for the participation constraint (6), we first note that it is restricted by the condition $\phi \geq -W_0$. Therefore, we can shift the wage function downward until it binds or $\phi = -W_0$ holds. The participation constraint can likewise be written as (18).

Intuitively, we search for the contract $W_T(P_T|c_0, c_1, c_2)$ with shape (11) that achieves three objectives. First, it provides the same effort and risk-taking incentives for the agent as the observed contract (conditions (19) and (20)). Second, it provides the agent with the same utility as the observed contract (condition (18)), and, third, it is as cheap as possible for the firm (objective (17)). If our model is correct and descriptive of the data, the cheapest contract found in this optimization will be identical to the observed contract. If the new contract differs substantially, we can reject the hypothesis that contract shape (11) is optimal, because it is possible to find a cheaper contract that implements the same effort and the same volatility as the observed contract. In this case, either the compensation practice is inefficient or the model is incorrect. In both cases, the model is not descriptive of the data.[11]

4 Data set

To construct approximate CEO contracts, we start with the most recent compensation contract of all CEOs in ExecuComp during the fiscal years 2007-2012. We include all CEOs from the fiscal year 2012 plus those from 2007-2011 who are not covered in any later years. We start from the year 2007 because the new reporting standards on option grants allow us to obtain all necessary information for each option grant and to calculate the accurate option portfolios for each CEO. We stop at the fiscal year 2012 because this is the most recent year available on ExecuComp at the time of our analysis.

---

[11]Edmans, Gabaix, Sadzik, and Sannikov (2012) consider a risk averse CEO in continuous time with multiplicative utility, not additive utility as in equation (3). In the empirical implementation, we have the advantage that the additive utility disposes of the cost function. However, the multiplicative utility keeps the cost function. Other models (DeMarzo and Sannikov (2006) and Zhu (2013)) consider risk neutral CEOs, an assumption which makes the pay *structure* (base salary, stock, and options) irrelevant.

Our selection process ensures that no CEO is counted twice and that there are as many CEOs as possible. As a robustness check, we also perform our main analysis for each individual year between 1997 and 2012 and the findings are qualitatively the same. Let us denote the year we selected as t. We first identify all persons in the database who were CEOs during the full year t and executive of the same company in $t-1$. This leaves us with 2,623 CEOs. We calculate the base salary ϕ (which is the sum of salary, bonus, other compensation, non-equity incentive plan compensation, and the change in pension value and nonqualified deferred compensation earnings from ExecuComp) from year t, and take information on stock and option holdings from the end of the fiscal year $t-1$. We subsume bonus payments under the base salary, because previous research has shown that bonus payments are only weakly related to firm performance (see Hall and Liebman, 1998).

We take the firm's market capitalization P_0 from the end of the fiscal year $t-1$. While our formulae above abstract from dividend payments for the sake of simplicity, we take dividends into account in our empirical work and use the dividend rate d from $t-1$. We estimate the firm's stock return volatility σ from daily CRSP stock returns over the fiscal year t and drop all firms with fewer than 220 daily stock returns on CRSP. We use the CRSP/Compustat Merged Database to link ExecuComp with CRSP data. The risk-free rate is set to the U.S. government bond yield with five-year maturity from January of year t.

Many CEOs in our sample have more than one option grant in their option portfolio. In this case, we aggregate this portfolio into one representative option. This aggregation is necessary to arrive at a parsimonious wage function that can be calibrated to the data. Our model is static and therefore cannot accommodate option grants with different maturities. The representative option is determined so that it has a similar effect as the actual option portfolio on the agent's utility, his effort incentives, and his risk-taking incentives. More precisely, we numerically calculate the number of options n_O, the strike price K, and the maturity T so that the representative option has the same Black-Scholes value, the same option delta, and the same option vega as the estimated option portfolio. Hence, we solve the following system of three equations in three variables:

$$n_O \cdot BS(P_0, K, T, \sigma, r_f) = \sum_i n_O^i \cdot BS(P_0, K^i, 0.7T^i, \sigma, r_f)$$
$$n_O \cdot delta(P_0, K, T, \sigma, r_f) = \sum_i n_O^i \cdot delta(P_0, K^i, 0.7T^i, \sigma, r_f)$$
$$n_O \cdot vega(P_0, K, T, \sigma, r_f) = \sum_i n_O^i \cdot vega(P_0, K^i, 0.7T^i, \sigma, r_f),$$

where n_O^i, K^i, and T^i are the number, the strike price, and the maturity of the ith option in the

CEO's option portfolio. We take into account the fact that most CEOs exercise their stock options before maturity by multiplying T^i by 0.7 before calculating the representative option (see Huddart and Lang, 1996, and Carpenter, 1998).[12]

We need a wealth estimate for the utility functions: We approximate the non-firm wealth W_0 of each CEO from the ExecuComp database by assuming that all historic cash inflows from salary and the sale of shares minus the costs of exercising options have been accumulated and invested year after year at the one-year risk-free rate. We assume that the CEO had zero wealth when he entered the database (which biases our estimate downward) and that he did not consume since then (which biases our estimate upward).[13] To arrive at meaningful wealth estimates, we discard all CEOs who do not have a history of at least five years (i.e., from $t-5$ to $t-1$) on ExecuComp. During this period, they need not be a CEO. This procedure results in a data set with 1,707 CEOs. In Section 7.1, we will show that the potential survivorship bias has a limited effect on our results.

[Insert Table 1 here]

Table 1 Panel A provides an overview of our data set. The median CEO owns 0.35% of the stock of his company and has options on an additional 0.50%. The median base salary is \$2.02 million, and the median non-firm wealth is \$19.1 million. The representative option has a median maturity of 4.4 years and is in the money with a moneyness (K/P_0) of 84.3%. Most stock options are granted at the money in the United States (see Murphy, 1999), but after a few years they are likely to be in the money. This is the reason why the representative option grants are in the money for two thirds of the CEOs in our sample.

We report descriptive statistics for the risk avoidance measure ρ in our sample for six values of risk aversion γ in Table 1, Panel B. Appendix C contains all the necessary formulae to calculate ρ.[14] For all six values of γ ranging from 0.5 to 6, risk avoidance ρ is positive for the majority of CEOs; for $\gamma \geq 3$ it is positive for 94.1% of all CEOs. This suggests that the majority of CEO will not adopt a project that increases firm risk if it leads to a drop in firm value. Therefore, the risk avoidance measure is consistent with our result that the compensation contracts chosen by the firm do not

---

[12] In these calculations, we use the stock return volatility for the lagged fiscal year (with at least 220 daily stock returns on CRSP) and, for the risk-free rate, the U.S. government bond yield with 5-year maturity from January of year t. Data on risk-free rates have been obtained from the Federal Reserve Board's website. For CEOs who do not have any options, we set $K = P_0$ and $T = 10$ (multiplied by 0.7) as these are typical values for newly granted options.

[13] These wealth estimates can be downloaded for all years and all executives in ExecuComp from http://people.few.eur.nl/dittmann/data.htm. They have also been used by Dittmann and Maug (2007) and Dittmann, Maug, and Spalt (2010).

[14] These measures of risk avoidance can be downloaded for all years and all executives in ExecuComp from http://people.few.eur.nl/dittmann/data.htm.

make CEOs risk-seeking. For $\gamma = 3$, the average ρ is 1.36 and the median is 1.11. This implies that the average CEO in our sample passes up risky positive NPV projects if they increase the firm value by less than 1.36% per percentage points of additional volatility.

While risk-avoidance ρ is zero in the first-best optimum, it is positive in the second-best optimum as risk-taking incentives are costly (cf. Figure 1). It is difficult to judge, however, what a plausible optimal level for ρ is, because this depends on the availability of profitable risky projects: a firm that only has few such projects will not benefit much from an increase in the risk-taking incentives. Nevertheless, a median ρ of 1.11 for $\gamma = 3$ appears large when taking into account that CEO pay typically constitutes only about 1.0% of the firm value (see the median of "value of contract" and "firm value" in Table 1 Panel A). We agree that these values are high, but also note that they do necessarily follow from our assumptions that CEOs have CRRA preferences with $\gamma = 3$, which is high.[15] We still use $\gamma = 3$ as the base case in this paper because it is a standard choice and provide robustness checks for $\gamma = 0.5$ and $\gamma = 6$. This range includes the risk-aversion parameters used in previous research.[16]

We require that all CEOs in our data set are included in the ExecuComp database for the years $t - 5$ to t, and this requirement is likely to bias our data set towards surviving CEOs, namely those who are richer and work in bigger and more successful firms. Table 1 Panel C compares the full ExecuComp universe of 1,526 CEOs in 2012 and 1,196 ExecuComp CEOs in 2012 that are included in our sample. The two-sample t-test and the Wilcoxon test show that as compared to the larger sample, our CEOs hold a smaller portion of options relative to the total outstanding shares (0.35% less), receive higher salaries ($0.21m more), and work in bigger firms ($780m more firm value). However, there is no statistical significance in CEO stock holdings, CEO age, and the past five-year stock returns, indicating that our sample does not have a bias towards older CEOs and more successful firms. In a robustness check below, we show that the effect of the selection bias is negligible.

Table 1 Panel D displays the corporate governance variables which will be discussed in the next section. We construct four corporate governance variables using two data sources, namely Institutional Shareholder Services (formerly RiskMetrics) and Thomson Reuters Form 13F institutional

---

[15]Graham, Harvey, and Puri (2013) show that CEOs are less risk averse than the population average, so that the CRRA-parameter γ might be considerably below 3. Faccio, Marchica, and Mura (2011) show that major shareholders might not be well diversified and therefore want to take less risk than what is optimal in a model with risk-neutral shareholders. Their findings suggest that shareholders do not intend to reduce risk avoidance ρ to zero, but to some other positive value.

[16]Lambert, Larcker, and Verrecchia (1991) use values between 0.5 and 4. Carpenter (1998) and Hall and Murphy (2000) use $\gamma = 2$. Hall and Murphy (2002) use $\gamma = 2$ and 3.

holdings. *E-index* is a measure of CEO entrenchment, following the definition of Bebchuk, Cohen, and Ferrell (2009). *CC-Ownership* measures the total percentage of ownership of all independent compensation committee members. *Institutional ownership* captures the percentage of shares held by institutional owners. *Blockholder* measures whether there is an institutional owner who holds 5% shares or more. The data coverage of the corporate governance variables for our sample ranges from 76% to 90%. A median firm has an E-index of 2, 0.03% ownership for all compensation committee members, 81% institutional ownership, and at least one blockholder.

5 Empirical Results

5.1 Calibration Results

Figure 2 shows our calibration results for a representative CEO.[17] The solid line represents the model contract W_T^* which solves the optimization problem (17) to (20), and the dotted line is the observed contract W_T^d. The figure shows the CEO's end-of-period wealth W_T as a function of the end-of-period stock price P_T which we express as a multiple of the beginning-of-period stock price P_0. The model contract with risk-taking incentives protects the CEO from losses. It provides the CEO with a flat wealth of $29.7m if the stock price falls below 56% of the initial stock price. Intuitively compared to the observed contract, limiting the downside for bad outcomes provides better (i.e., cheaper) risk-taking incentives than rewarding good outcomes. In the region between 56% and 93%, the contract is increasing and convex. For larger stock prices, the contract is concave. The reason for the concavity is the CEO's decreasing marginal utility: the richer is the CEO, the less interested he is in additional wealth.

As a benchmark, we also calibrate the model contract without risk-taking incentives from Dittmann and Maug (2007); this is shown by the dashed line in Figure 2. For this purpose, we solve the optimization problem (17) to (19) without the volatility incentive constraint (20) and use the contract shape $W_T^*(P_T|c_0, c_1) = (c_0 + c_1 \ln P_T)^{1/\gamma}$. We call this contract the *benchmark contract* or the *CRRA contract* while we refer to the contract from the full model as the *RTI contract* or, more precisely, the *CRRA-RTI contract*. Figure 2 shows that the benchmark contract is globally concave and puts the agent's entire wealth at risk.

---

[17]For each parameter (observed salary ϕ^d, observed stock holdings n_S^d, observed option holdings n_O^d, wealth W_0, firm size P_0, stock return volatility σ, time to maturity T, and moneyness K/P_0) and each CEO, we calculate the absolute percentage difference between individual and median value. Then, we calculate the maximum difference for each CEO and select the CEO for whom this maximum difference is the smallest.

Figure 2: The figure shows end of period wealth W_T for the observed contract (dotted line), the optimal CRRA contract with risk-taking incentives (solid line), and the optimal CRRA contract without risk-taking incentives (dashed line) for a representative CEO whose parameters are close to the median of the sample. The parameters are $\phi = \$1.51\text{m}$, $n_S = 0.31\%$, $n_O = 0.69\%$ for the observed contract. Initial non-firm wealth is $W_0 = \$24.9\text{m}$. $P_0 = \$1.5\text{bn}$, $\sigma = 24.1\%$, and $K/P_0 = 81\%$, $T = 4.5$ years, $r_f = 0.8\%$, $d = 2.8\%$. All calculations are for $\gamma = 3$.

Since the results may be sensitive to γ, we repeat our analysis for $\gamma = 0.5$ and $\gamma = 6$ in Figure 3. Both plots show that the model contract with risk-taking incentives generates much better fits of the observed contract than the model without risk-taking incentives, especially when $\gamma = 0.5$. In addition, we find the same pattern as in Figure 2, i.e., that the optimal contract protects the CEO for bad outcomes when the stock price falls below 45% of the scaled stock price for $\gamma = 0.5$ and 56% for $\gamma = 6$, respectively. When the stock price is above 45%, the contract for $\gamma = 0.5$ is convex until 186% when it turns concave. When $\gamma = 6$, the contract is convex in the region between 56% and 78% and concave for a higher stock price.

Both Figure 2 and Figure 3 suggest that the model with risk-taking incentives (solid line) fits the observed contract (dotted line) much better than the model without risk-taking incentives (dashed line). To quantify this visual impression, we calculate for both models the average distance between the contract W_T^* predicted by the model and the observed contract W_T^d:

$$D_1 = E\left(\frac{\left|W_T^*(P_T) - W_T^d(P_T)\right|}{W_T^d(P_T)}\right). \tag{23}$$

We recognize that the observed contract we construct in Section 4 is a stark simplification of the
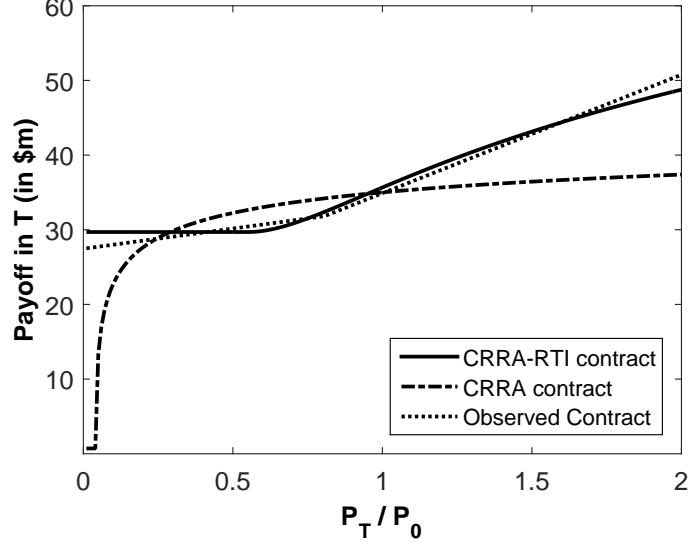
16
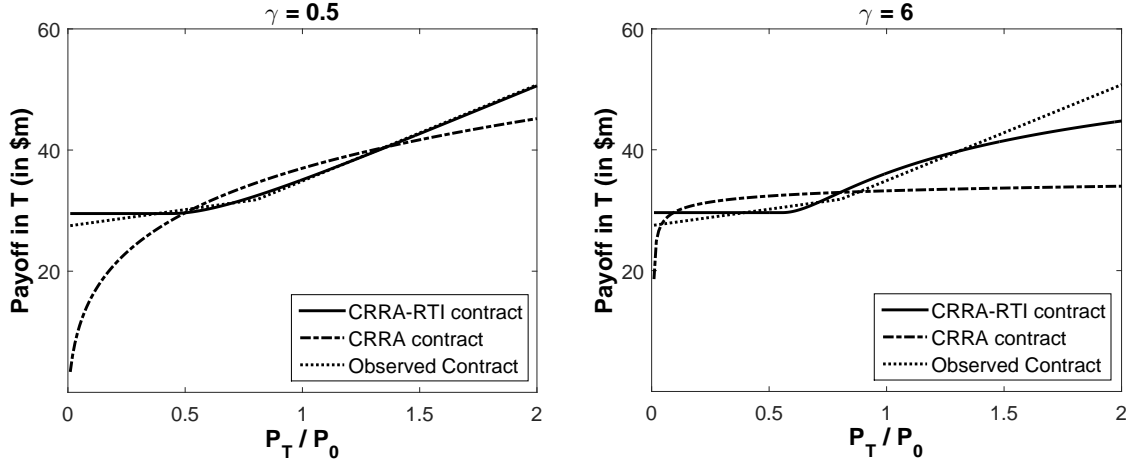
Figure 3: Both plots show end-of-period wealth W_T for the observed contract (dotted line), the optimal CRRA contract with risk-taking incentives (solid line), and the optimal CRRA contract without risk-taking incentives (dashed line) for the same representative CEO as Figure 2. The plots only differ in the value of parameter γ, with $\gamma = 0.5$ for the left plot and $\gamma = 6$ for the right plot.

contracts used in practice, especially because typical contracts contain several grants of stock options with different maturities and different strike prices. Therefore, contracts are in general not piecewise linear with just one kink but have a more complicated shape. To address this caveat, we consider a second distance metric

$$D_2 = E\left(\frac{\left|W_T^*(P_T) - W_T^{smth}(P_T)\right|}{W_T^{smth}(P_T)}\right), \tag{24}$$

where $W_T^{smth}(P_T)$ sums up the expected value of the sum of the base salary and all stock and option grants held by the CEO. For an option grant that has a maturity larger than T, this is just the Black-Scholes value for the remaining maturity, given P_T. For a grant that has a maturity smaller than T, we calculate the expected value of the option at maturity given P_0 and P_T and assume that this amount is invested at the risk-free rate for the remaining time between maturity and T. In this way, we obtain a smooth contract for all CEOs who have at least two different option grants. For CEOs with only one option grant, $W_T^{smth}(P_T) = W_T^d(P_T)$. We explain the construction and calculation of W_T^{smth} in more detail in Appendix D. For the representative CEO shown in Figure 2, the distance D_1 is 1.5% ($D_2 = 1.8\%$) for the contract with risk-taking incentives and 6.7% ($D_2 = 5.7\%$) for the contract without risk-taking incentives.

[Insert Table 2 here]

Table 2, Panel A shows the results for all CEOs in our sample. The left part of the table describes the model contract with risk-taking incentives (CRRA-RTI model) for three values of

constant relative risk-aversion γ. We do not tabulate the parameters c_0, c_1, and c_2, as they cannot be interpreted independently from each other. Instead, the table shows the mean and the median of some key variables that describe the contract. These variables include the two distance measures D_1 and D_2 from (23) and (24) and the manager's minimum wealth ($\min W_T^*(P_T)$) scaled by non-firm wealth W_0. In addition, the table shows two probabilities. First, the kink quantile is the probability that the contract pays out the minimum wage in the flat region of the contract; formally, this is $\Pr(ln(P_T) \leq -\frac{c_1}{2c_2})$ from equation (11). Second, the inflection quantile is the probability mass below the point where the contract curvature changes from convex to concave. Finally, the table also shows risk avoidance ρ from (16).

Table 2 demonstrates that the model contract provides the agent with comprehensive downside protection. For $\gamma = 3$, the median minimum wealth is 1.3 times the initial wealth W_0. None of the CEOs in our sample have a minimum wealth lower than their observed non-firm wealth W_0. The variable *Kink quantile* shows that the contract pays out the minimum wage for the worst 21.6% of all outcomes in the median. The median inflection quantile is 47.5%, so that the contracts are convex for mediocre performance between the 21.6% and the 47.5% quantile and concave for good performance above the 47.5% quantile.

Table 2, Panel A also shows the savings firms could realize when they switch from the observed contract to the model contract. These savings are defined as

$$savings = \left[E\left(W_T^d(P_T)\right) - E\left(W_T^*(P_T)\right)\right] / E\left(W_T^d(P_T)\right).$$

For $\gamma = 3$, the mean (median) savings are 10.3% (4.4%). The mean distance D_1 between the observed contract and the model contract is 5.4%, and the mean distance D_2 is 6.3%. For lower values of risk-aversion γ, we obtain a better fit: for $\gamma = 0.5$, the average distance D_1 is only 2.3%. Contracts are then convex over a larger range of stock prices from the 10.3% quantile to the 74.6% quantile for the median CEO. Conversely, we find a worse fit for higher values of risk-aversion γ. The region of convexity shrinks relative to our benchmark case $\gamma = 3$ and the distance to the observed contract increases according to all measures. By construction, the savings from recontracting are smaller for lower γ.[18]

---

[18]For $\gamma = 0$, the risk premium disappears and the problem becomes degenerate. Then, the compensation mix cannot be determined as it does not matter. Numerically, these problems already occur for $\gamma = 0.5$, when we have convergence for only (1151/1707) 67% of our observations (see the last line in Table 2 Panel A). We also experience numerical problems for $\gamma = 6$ for (1124/1707) 66% of our observations. The reason is that $V'(W_T) = W_T^{1-\gamma}$, so that for high values of γ we obtain very low values of marginal utility, even though we scale all dollar values by the firm's stock price. To learn more from the cases when the algorithm fails to converge, we provide more discussions in Appendix E

The right part of Table 2 displays the results for the benchmark model without any risk-taking incentives (CRRA model). It shows that the average risk avoidance in this model is 4.91 (for $\gamma = 3$) and therefore much higher than in the model with risk-taking incentives where it is 1.37. The benchmark contract does not contain any downside protection, so the CEO can potentially lose all his wealth. Moreover, it is globally concave for all CEOs if $\gamma \geq 1$, so that the kink quantile and the inflection quantile are both zero. Due to convergence problems, the sample size in Table 2, Panel A is not the same for the two sets of results. Therefore, we once more report the numbers in Panel B for the subsample of CEOs for whom we obtain convergence for both models. This panel shows that the model with risk-taking incentives approximates observed contracts much better than the benchmark model. For $\gamma = 3$, the average distance D_1 is 16.1% for the benchmark model as compared to 5.4% for the RTI model. The savings from recontracting are also much higher for the benchmark model than for the RTI model. The benchmark model suggests that shareholders leave 17.7% of the contracting costs on the table while the RTI model puts this number at 10.3% only. These numbers suggest that risk-taking incentives play an important role in observed compensation contracts. Observed contracts appear to be markedly more efficient when risk-taking incentives are taken into account.

A natural question to ask is how firm value would increase if the CEO counterfactually chose higher risk. Indeed, firm value P_0 and risk σ are related. Ceteris paribus, in the region where $\sigma < \sigma^{**}$ (see Figure 1), firm value is increasing in risk. If the CEO counterfactually chose a marginally higher risk, the firm value would increase. If we had a functional form of $P_0(e, \sigma)$, we could make predictions about the firm value. However, we merely assume that $P_0(e, \sigma)$ is increasing and concave in σ as long as $\sigma < \sigma^{**}$ and therefore we can make no precise prediction.

5.2 Risk avoidance and deviations from the optimal contract in sample splits

Table 3 displays median risk avoidance together with median distance D_1 for several subsamples. When we consider the sample split for banks and non-banks, the median of risk avoidance for non-bank firms is larger than for banks. This is in line with John, Saunders, and Senbet (2000) and Chaigneau (2013b) who show that it can be optimal for bank shareholders to design a CEO contract with excess risk taking incentives when they are partially protected by deposit insurance and too-big-to-fail implicit guarantees. However, banks still have a sizable positive median risk avoidance of 0.84. This conclusion is also true when we go back in time and consider the data for 2006 (not

---

on whether there are some differences between CEO/firms/contract where we obtain convergence and where we do not.

shown in the table). Our model suggests that risk-taking incentives in banks were not excessive from the perspective of bank shareholders, but they might still be excessive from a social perspective. Moreover, the result for median distances suggests that contracting is more efficient in banks than in non-bank firms.

[Insert Table 3 here.]

Table 3 also shows the split according to the book and market leverage within non-bank firms.[19] Our model does not include leverage and, accordingly, bankruptcy is impossible. Extreme leverage is therefore not covered by our model. We assume that the incentives are set before the CEOs make both the leverage and the project decisions. An increase in leverage constitutes a redistribution of wealth from bondholders to shareholders and increases the equity risk.[20] We indeed find that risk avoidance in the subsample with low leverage is higher than in the subsample with high leverage. The result for the distances shows that contracting is more efficient in the sample with high leverage. This is consistent with Jensen's (1986) argument that debt markets help to discipline managers, e.g., by removing free cash-flows.

Moreover, we compare risk avoidance measures and distances with the E-index. A low E-index can be interpreted as a measure of good governance. We find that risk avoidance ρ is 22% higher for more entrenched CEOs (E-index), which is consistent with the hypothesis that entrenchment can have adverse effects on management behavior and incentives. This could also be due to other factors, for instance, risk-taking is less important in some firms, such as more mature firms, which take fewer risks and which tend to have more entrenched CEOs. For the distances, we have, as expected, more efficient contract arrangements for firms whose E-index is low.

Finally, the risk avoidance is 39% higher for firms with higher ownership among the compensation committee members (CC-Ownership). Similarly, risk avoidance is 23% higher for firms with higher institutional ownership and it is 52% higher when there is an institutional owner who holds 5% shares or more. For the distances, we find that contracting is more efficient for firms with low compensation committee ownership, low institutional ownership, and no 5% blockholders. This finding seems counterintuitive, but it can be explained by Faccio, Marchica, and Mura (2011) that

---

[19]Shue and Townsend (2014) find causal evidence that a positive change in the CEO option grant increases the leverage. We calculate book leverage as (total long-term debt + total debt in current liabilities) / total assets and market leverage as (total long-term debt + total debt in current liabilities) / (total assets + market equity - book equity) where book equity is the sum of stockholders' equity, deferred taxes, and investment tax credit minus preferred stock.

[20]Note that we keep on referring to *firm risk* in other parts of the paper but, in this context, it would be more precise to talk about *equity risk*.

large shareholders might not be well diversified and therefore take less risk than what is optimal in a model with risk-neutral shareholders. Since our model does not allow for risk-averse principals, it overstates the savings for firms with undiversified large shareholders. It is also possible that our corporate governance proxies are correlated with some other variables that are driving the contract efficiency. For example, shareholders may appoint a blockholder to the compensation committee only when there is a more entrenched and powerful CEO. If the effect of CEO influence outweighs the effect of the blockholder, then we would expect firms with blockholders on the compensation committee (higher CC-Ownership) to have less efficient contracts.

6   Application: Optimal strike prices

In this section, we analyze the implications of the RTI model for optimal strike prices in a standard option contract. Therefore, we consider contracts that have the same structure as the stylized contract in Section 4, consisting of fixed salary ϕ, the number of stock n_S, and the number of options n_O with the strike price K:

$$W_T^{lin} = \phi + n_S P_T + n_O \max\{P_T - K, 0\}$$

For each CEO, we solve the optimization problem (17) to (20) with W_T^{lin} instead of W_T^*, where the principal's choice variables are ϕ, n_S, n_O, and K.[21]

[Insert Table 4 here]

Table 4 describes our results for five values of γ: 0.5, 2, 3, 4, and 6. In all cases, the RTI model predicts that the median CEO does not hold any stock. Instead, the median CEO would have more options (+70% for $\gamma = 3$; compare Table 4 with Table 1) and more base salary (+68%). For 97% of the CEOs in our sample, the strike price in the model contract is lower than that in the observed

---

[21]We need a few additional restrictions, so that the problem is well-defined. First, we assume that the number of shares n_S is non-negative. We allow for negative option holdings n_O and negative salaries ϕ, but we require that $n_O > -n_S \exp\{dT\}$ and $\phi > -W_0$ to prevent negative payouts. Negative option holdings or negative salaries are rarely seen in practice, but they are certainly possible. A negative salary would imply that the firm requires the CEO to invest this amount of his private wealth in firm equity. We do not allow for negative stockholdings, because compensation could then become non-monotonic in stock price, which violates one of our model assumptions.

We also need to restrict the strike price K, because options and shares become indistinguishable if K approaches zero, and the problem is poorly identified if K is small. We work with two lower bounds for K. We first solve the numerical problem with the restriction $K/P_0 \geq 20\%$. If we find a corner solution with $K/P_0 = 20\%$, we repeat the calibration with a lower bound $K/P_0 \geq 10\%$. If the second calibration does not converge, we use the (corner) solution from the first step.

In many cases, the objective function in our problem is rather flat around the optimal solution. In order to check whether an interior solution with $n_S^* > 0$ is indeed the optimal solution, we repeat our calibration with the additional restriction $n_S = 0$ whenever we obtain a solution with $n_S^* > 0$ in the original problem. For our empirical analysis, we always use the solution with the lowest costs.
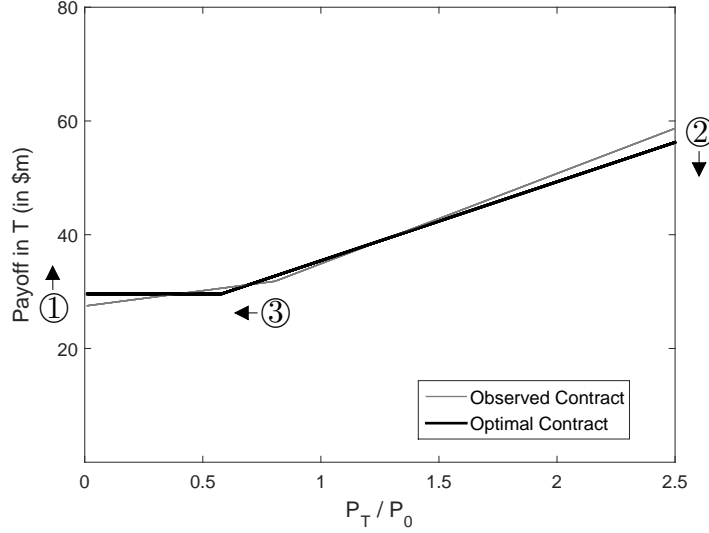
Figure 4: The figure shows end-of-period wealth W_T as a function of the end-of-period stock price P_T for the observed contract (thin grey line) and the optimal piecewise linear contract (thick black line) for one CEO in our sample. The arrows indicate the three main features of the optimal contract relative to the observed contract: (1) it punishes very bad outcomes less, (2) it rewards very good outcomes less, and (3) the strike price of the option grant is lower. The parameters for this CEO are $\phi = \$1.5m$, $n_S = 0.3\%$, $n_O = 0.69\%$ for the observed contract. Initial non-firm wealth is $W_0 = \$25m$. $P_0 = \$1526m$, $\sigma = 24.1\%$, and $K/P_0 = 80.6\%$, $T = 4.54$ years, $r_f = 0.8\%$, $d = 2.84\%$, and $\gamma = 3$.

contract. While the moneyness of the observed contract is 84.3% in the median in Table 1, it is 46.7% for the model contract. If we assume that observed option grants have been issued at the money and have moved into the money only because of the general stock price increase in the years before year t, our results imply that options should have been issued 55.4% (= 46.7%/84.3%) in the money.

The general picture is that the stock and option holdings in the observed contract are replaced by option holdings that are considerably deeper into the money. As options are less valuable than shares, this exchange is accompanied by an increase in the base salary, so that the new contract provides the same expected utility to the agent as the observed contract. The savings generated by switching to the model contract are limited, however. The median firm would just save 2.1% of the contracting costs for $\gamma = 3$ and the average is 5.7%. This is hardly a savings potential that would trigger shareholder activism or takeovers. The comparatively small savings imply that a portfolio of stock and at-the-money options constitutes a good substitute for in-the-money options. The numerical flip side of low savings is that the objective function (after taking into account the constraints) is rather flat. While this is certainly a complication when it comes to solving the model numerically, it is not a *problem* of our model but rather a *result*.

Figure 4 illustrates our main results. It shows the wealth function $W_T^{lin}(P_T)$ of the observed contract and the model contract for one CEO in our sample. This CEO is not representative for our sample; for the representative CEO, the two contracts are more difficult to be distinguish visually. The three arrows in Figure 4 illustrate the main features of the model contract as compared to the observed contract that consists of a portfolio of stock and at-the-money options. The first feature of the model contract is that it provides for less punishment in the bad states of the world than the observed contract, which improves the risk-taking incentives. On the other hand, the model contract also gives fewer rewards in the best states of the world (feature 2), which reduces the risk-taking incentives. Effort incentives, on the other hand, are reduced by both features (1) and (2). Moving the strike price more into the money (feature 3), however, increases the effort incentives and offsets the effect of features (1) and (2). Therefore, the model contract moves some of the effort incentives from the tails of the distribution to its center. Finally observe that features (1) and (2) make the model contract less risky than the observed contract. Therefore, the agent demands a lower risk premium and the model contract is cheaper for shareholders. The same effects can be found for the general optimal contract depicted in Figure 2.

In-the-money options are rare in U.S. compensation practice. A potential reason is that the U.S. tax system strongly discriminates against in-the-money options (see Walker, 2009). In the remainder of this section, we therefore describe the optimal option contract if realistic taxes are taken into account. According to IRC Section 409A, income from in-the-money options is subject to a 20% penalty tax that has to be paid by the executive at the time of vesting. Shares, at-the-money options, or out-of-the-money options are not subject to this additional tax. Moreover, in-the-money options (like restricted stock) do not automatically qualify as performance-based pay under IRC Section 162(m) and therefore count towards the $1 million per executive that are tax deductible at the firm level. However, this rule can easily be circumvented by subjecting in-the-money options to specific performance criteria. Therefore, we concentrate on the 20% penalty tax from Section 409A and neglect the potential effects of Section 162(m) in the following analysis.[22]

We repeat our numerical analysis for $\gamma = 3$ with a 20% tax penalty on in-the-money options. We assume that this tax must be paid if and only if the actual strike price is lower than the observed strike price, so we effectively assume that all options in the observed contract have been issued at-the-

---

[22]In addition, Section 409A requires that the difference between the stock price and the strike price be recognized as income at the time of vesting, rather than on exercise. Thus this rule accelerates income recognition from the exercise date to the vesting date (see Alexander, Hirschey, and Scholz (2007)). Our model does not distinguish between exercise date and vesting date, so we cannot model this effect.

money. If the 20% tax were not taken into account, the mean tax revenues from issuing in-the-money stock options would be $3,602,000. However, if the tax is taken into account, the mean tax revenues will be $1,258,000. The mean deadweight loss is 5.7% - 2.1% = 3.6% (i.e., the savings if no taxes are taken into account minus the savings if the 20% are taken into account). In this setting, we find that the 74.6% of the 1,686 CEOs for whom our algorithm converges have exactly the same optimal contracts (including salary, number of stock and options) as the observed contract. All numbers listed in this paragraph are not shown in the tables.

Many other countries (including the U.K., Canada, Germany, and France) discourage the use of in-the-money options, so the United States is not an exception (see Walker, 2009).[23] A potential reason is that in-the-money options induce some costs that are not included in our model and that justify government intervention. Our results in Table 4 show that the use of in-the-money options is associated with large increases in the base salary. These might be difficult to explain to shareholders and the general public, and might cause social unrest and higher wage demands. A commitment to using only at-the-money options would reduce the CEO's base salary, and our analysis shows that the costs of such a commitment are rather low (compare the savings in Table 4).

7 Robustness checks

7.1 Sample selection bias

Our data set is subject to a moderate survivorship bias, as we require that CEOs are covered by the ExecuComp database for at least five years. Table 1 Panel C indicates that smaller firms are underrepresented in our data set and our CEOs have lower option holdings (due to the larger number of outstanding shares) and higher salaries which are usually associated with larger firm size. To see how the bias towards bigger firms affects our results, we divide our sample into quintiles according to four variables: CEOs' option holdings, fixed salary, CEOs' non-firm wealth W_0, and firm value P_0. Table 5 displays for these subsamples the average distance D_1, and, in the last line, the p-value of the Wilcoxon test that the average distance is identical in the first and the fifth quintile. This analysis is done for $\gamma = 3$.

[Insert Table 5 here.]

---

[23]Australia is the only country for which we could find evidence that in-the-money options are commonly used. See Rosser and Canil (2004).

The table shows that the model fit is worse for CEOs with larger option holdings, lower salaries, and lower non-firm wealth. For the 20% largest CEO option holdings, the 20% lowest salaries, and the 20% least wealthy CEOs, we find an average distance of 7.2%, 6.7%, and 7.1%, respectively, compared to 5.4% for the full sample (see Table 2). Given that our sample is biased towards wealthier CEOs with smaller option holdings and higher salaries, the average distance in the unbiased sample would be somewhat higher than shown in Table 2. We find no significant difference in the average distance along the firm value dimension, thereby suggesting that the effect of the selection bias towards big firms is negligible. Altogether, the effect of the sample selection bias on our results is therefore small.

7.2 Robustness check for 1997-2012

So far, we use the augmented year 2012 data set for our analyses. As a robustness check, we repeat our main analysis for each individual year between 1997 and 2012. The sample for this robustness check starts from the fiscal year 1997, because the wealth estimate for a CEO needed for the utility function requires at least five years of history and ExecuComp starts in 1992. Before 2006, the proxy statement does not disclose any complete data on previously granted options. Therefore, we estimate each CEO's option portfolio with the method proposed by Core and Guay (2002). After 2006, we are able to obtain all necessary information for each option grant and calculate the accurate option portfolios for each CEO. For the years before 2006, we do not take into account pension benefits, because they are not available in ExecuComp and are difficult to compile. Pensions can be regarded as negative risk-taking incentives (see Sundaram and Yermack (2007), and Edmans and Liu (2011)), so that we overestimate the risk-taking incentives in observed contracts.

[Insert Table 6 here]

We take $\gamma = 3$ and apply the same calibration procedure for the CRRA-RTI model as in Table 2 Panel A. The results are summarized in Table 6. First of all, the means and medians for all variables listed in the table are of similar magnitudes as those in Table 2 Panel A, indicating that our findings are independent of the sample period. Second, two distance measures D_1 and D_2 and savings are relatively high for the periods 1998-2002 and 2007-2009, suggesting that our model better fits to the CEO contract in regular times than in the crisis period (IT bubbles and the recent financial crisis).

7.3 CEO dismissals

Our model does not incorporate any CEO dismissals. However, when the past stock price performance is bad, the CEO is likely to lose his position and suffers a negative shock on his wealth and human capital (e.g., Coughlan and Schmidt (1985), Kaplan (1994), Fee and Hadlock (2004), Jenter and Kanaan (2015), Kaplan and Minton (2012), and Peters and Wagner (2014)). The absence of CEO dismissal is of particular concern for our analysis where an important layer of risk perceived by the CEO is missing in our model. Therefore, we address these shortcomings here by introducing the threat of dismissal into the CEO's wealth function.

[Insert Table 7 here]

Specifically, we follow Dittmann and Maug (2007) and estimate a logit regression for CEO dismissals. The dummy variable for CEO dismissal is equal to one if a CEO who is in the data set in 2008 leaves the company within five years and ExecuComp records "resigned" as the reason for leaving. We regress this dummy variable on the 5-year stock return from 2008 to 2012.[24] Then, we use these parameters of the logistic function to estimate the probability of CEO dismissal as a function of terminal stock returns $p(P_T/P_0)$. We assume that the CEO loses all her compensation in the event of dismissal, which is most likely an overstatement as severance pay is ignored here. That is, we redefine the end-of-period wealth (compare W_T^d on page 10) as

$$W_T^d = W_0 exp(r_f T) + (1 - p(P_T/P_0)) \left(\phi^d exp(r_f T) + n_S^d P_T + n_O^d \max\{P_T - K^d, 0\} \right) \qquad (25)$$

We rerun the analysis in Table 1 Panel B and compute the risk avoidance using the new definition of W_T^d. The results are summarized in Table 7. Comparing the numbers in this table with those in Table 1, we can see that risk avoidance with the threat of dismissal is slightly higher than that without the threat of dismissal for all levels of risk aversion. For example, for $\gamma=3$, the median risk avoidance is 1.17 which is 5% higher than 1.11 when CEO dismissal is absent. Therefore, omitting CEO dismissals biases our risk avoidance measure downwards, though the difference is not particularly large. However, we assume that the CEO loses everything when dismissed while Yermack (2006) argues that managers are partially compensated for their dismissal. If we take severance pay into account, the bias in our risk avoidance measure will be even smaller.

---

[24]If we regress the dummy variable on the 5-year *abnormal* return, which is the difference between the 5-year gross stock return and the 5-year market return as in Dittmann and Maug (2007), we get exactly the same result. We keep the gross return as the regressor for its simplicity.

7.4 Optimal contracts when CEOs are loss averse

Dittmann, Maug, and Spalt (2010) propose an alternative model without risk-taking incentives where the manager is loss averse. They also calibrate the model to the data and show that it fits the data well. In this section, we therefore compare the CRRA-RTI model and the loss-aversion model (henceforth: LA model) and investigate whether the LA model can be further improved by taking into account risk-taking incentives.

The standard loss-aversion model

Loss-aversion preferences are given by (see Tversky and Kahneman, 1992)

$$V^{LA}(W_T) = \begin{cases} (W_T - W^R)^\alpha & if \quad W_T \geq W^R \\ -\lambda(W^R - W_T)^\beta & if \quad W_T < W^R \end{cases} \quad , \text{ where } 0 < \alpha, \beta < 1 \text{ and } \lambda \geq 1. \quad (26)$$

Here, W^R is the agent's reference wealth level. Payouts above this level are coded as gains, while payouts below this level are coded as losses. The agent is risk-averse over gains and risk-seeking over losses, and losses receive a higher weight than gains ($\lambda > 1$). The utility $U^{LA}(W_T, e) = V^{LA}(W_T) - C(e)$ then replaces equation (3). Following Dittmann, Maug, and Spalt (2010), we use $\alpha = \beta = 0.88$ and $\lambda = 2.25$ and parameterize reference wealth W^R by $W_t^R = W_0 + \phi_{t-1} + \theta \cdot MV(n_{t-1}^S, n_{t-1}^O, P_t)$, where $MV(.)$ denotes the market value of last year's stock and option portfolio evaluated at this year's market price. Reference wealth therefore equals the sum of non-firm wealth W_0, last year's fixed salary ϕ, and a portion θ of today's market value of the stock and options held last period. Dittmann, Maug, and Spalt (2010) show that the model fits the data best for $\theta = 0.1$ and we therefore consider three values of θ: 0.1, 0.5, and 0.9.

[Insert Table 8 here]

Table 8 Panel A displays our results for the LA model for three different values of reference wealth as parameterized by θ. In addition to the mean and the median of the two distance metrics D_1 and D_2, and the savings, the table shows the average probability that the terminal payout is zero (the "jump quantile"), the inflection quantile where the contract changes from convex to concave, and risk avoidance ρ. We find that the LA model with $\theta = 0.1$ approximates the observed contract better than the CRRA-RTI model with $\gamma = 3$. The median distance D_1 is 2.3% for the LA model with $\theta = 0.1$ as compared to 4.0% for the CRRA-RTI model (see Table 2).[25] For higher reference

---

[25] Across all models and all specifications, the CRRA-RTI model with $\gamma = 0.5$ has the best fit. However, we do not

wealth, however, the LA model is considerably worse than the RTI model for any of the risk-aversion parameters considered ($\gamma = 0.5$, 3, and 6). The reason is that the probability that the CEO ends up with zero wealth is low only for very low reference points: for $\theta = 0.5$, the average jump quantile is 3% and for $\theta = 0.9$ it is 11.8% as compared to 0.6% for $\theta = 0.1$. Therefore, we conclude that the LA model is superior only for a rather specific choice of parameterization. In contrast, the CRRA-RTI model offers a reasonable approximation of the observed contract that is more robust to changes in the preference parameter.

Risk-taking incentives in the loss-aversion model

We follow similar procedures as in Dittmann, Maug, and Spalt (2010) to derive the shape of the optimal loss-aversion contract that takes risk-taking incentives into account in the Internet Appendix. We refer to this contract by the acronym LA-RTI.

The results are shown in Table 8 Panel B which is similar to Table 2 Panel B. The table shows that the probability that the CEO ends up with zero wealth is much lower for the LA-RTI model compared to the LA model. For $\theta = 0.5$, this probability decreases from 5.2% to 2% on average. Removing the punishment for poor outcomes increases the risk-taking incentives, and the LA-RTI model has a slightly better fit than the LA model if $\theta \leq 0.5$. For $\theta = 0.9$, however, the average distance metrics are higher for the LA-RTI model as compared to the LA model. The number of observations displayed in this table is quite small due to the difficulty in solving the optimization problem numerically. Altogether, we therefore conclude that the LA-RTI model does not yield any significant improvement over the LA model because the risk avoidance is low in Table 8 Panel A. For $\theta = 0.1$ in particular, the average risk avoidance is 0.16 and the median risk avoidance is 0.03.

7.5 Constant absolute risk aversion

The CEO's attitude to risk is central to our model. So far, we have assumed that the CEO's preferences exhibit constant relative risk aversion (CRRA). To see whether our results are robust to alternative assumptions on CEO risk-aversion, we repeat our analysis from Table 2 with constant absolute risk aversion (CARA), so that $V^{CARA}(W_T) = -\exp(-\eta W_T)$ replaces $V(W_T)$ in equation (3).

All our results continue to hold with CARA utility (see Table IA.1 in the Internet Appendix). In particular, the CARA-RTI model generates a much better fit than the CARA model as it guarantees

---

regard the CRRA model with $\gamma = 0.5$ as reasonable, because the model then implies unrealistic portfolio decisions. A CEO with $\gamma = 0.5$ would borrow heavily and invest much more than his entire wealth into the market portfolio.

a minimum payout that is always higher than the CEO's non-firm wealth, and it is convex for intermediate payouts and concave for good payouts. According to the two distance measures, CRRA-RTI dominates CARA-RTI.

8 Limitations of the model

8.1 The convexity of contracts

Our model predicts that contract payoffs are concave in firm value when the firm value is high. The observed contract is different from this prediction, since CEOs' payout is never concave when it is paid in long positions in stock and options. To evaluate how much our model prediction W_T^* deviates from the observed contract $W_T^{smth}(P_T)$ (which is the expected value of the sum of the base salary and all stock and option grants held by the CEO; see Appendix D), we take the following approach to capture the convexity and the deviation. First, we summarize the convexity of the model contracts by tabulating the proportion of the model contracts that are convex at the xth percentile of P_T (denoted by $P_{x\%}$ hereafter). Second, we compare the convexity of the model contracts with that of the observed contracts by taking the difference between the second derivative of the model contracts $W_T^*(P_T)$ at $P_T = P_{x\%}$ and the second derivative of the observed contracts $W_T^{smth}(P_T)$ at $P_T = P_{x\%}$. Third, we calculate the average distance between the model contract W_T^* and the observed contract W_T^{smth} for the right tail of the distribution:

$$D_2^{x\%} = E_{P_T \geq P_{x\%}} \left(\frac{\left| W_T^*(P_T) - W_T^{smth}(P_T) \right|}{W_T^{smth}(P_T)} \right). \tag{27}$$

[Insert Table 9 here]

In Table 9, we tabulate three measures described above for the 80th, 90th, and 95th percentile of the firm value. First, for $\gamma = 3$, the proportion of model contracts that are convex decreases from 11.2% at $P_{80\%}$ to 0.8% at $P_{95\%}$. Second, 0.7% of the contracts are more convex than the observed contract for $\gamma = 3$ at $P_{80\%}$ and the number goes down slightly to 0.1% $P_{95\%}$. Third, the median deviation for the top 20% of P_T when $\gamma = 3$ counts for 35.2% of the total deviation for the whole distribution of P_T as measured by the distance D_2. This number is 19.6% for the top 5% of P_T. Finally, the table also shows that the deviation at the right tail of the distribution is of less concern when the risk aversion is low. However, the deviation can be quite sizable when the risk aversion increases. In the next subsection, we offer some speculations on dynamic contracting issues, which

could potentially explain part of the deviation.

8.2 The dynamics of contracts

One limitation of our analysis is that our model is static and only considers two points in time: the time of contract negotiation and the time when the final stock price is realized. We are aware that other events that are not specified in our model might occur between these two time points. Specifically, our model does not include any new grants, gradual vesting, and contract renegotiation.

Our model faces the challenge that the initial optimal contract may lose its incentive effect over time in a dynamic world. One way of restoring CEO incentives is to award new grants and another way is to renegotiate the contract. First, when a firm performs poorly after the initial incentives are provided, the option will be deep out of the money and provide little incentives. A potential solution is to issue new option grants at the money. An alternative way is to renegotiate the contract and lower the exercise price of the out-of-the-money options (Acharya, John, and Sundaram (2000) and Brenner, Sundaram, and Yermack (2000)), but it is controversial as it seems to reward the CEO for failure. Second, when a firm performs extremely well, the options will be well into the money and they resemble stock that have little or no risk-taking incentives. In this scenario, firms may give additional option grants to introduce convexity into the contract. In both cases, the inclusion of new grants and contract renegotiation can introduce more convexity into the contract and better align our model to the data. This leads to a lower risk avoidance measure and a higher pay for performance sensitivity.

Core and Guay (1999) provide empirical evidence that firms use new equity grants to move CEOs towards their optimal incentive levels. They estimate a cross-sectional model of CEO incentives and take the residual of this model to predict any new grants to executives in the following year. They do not consider risk taking incentives and risk aversion. An alternative modeling approach would be to estimate an appropriate cross-sectional model similar to that of Core and Guay (1999). It would likely increase the precision of our risk avoidance measure, increase the convexity of our model and, consequently, increase the degree to which the theory lines up with the data.

In a dynamic setting on the theoretical side, single-period contracts can encourage the CEO to engage in short-termism by inflating the current stock price at the expense of long-term firm value (see, for example, Peng and Röell (2008, 2014) and Goldman and Slezak (2006)). One remedy is to introduce gradual vesting of equity grants. For example, Edmans et al. (2012) proposed a "Dynamic Incentive Account" of which a fraction is paid to the CEO every year and the remainder remains

escrowed to deter myopia. Furthermore, the CEO regularly sells stock and exercises options to keep up his consumption. Zhu (2016) shows that "bonus banks" that pay out a fraction of bonuses to the manager each period also help to deter myopia. Chaigneau (2015) models progressive learning about firm value due to exogenous shocks to explain the multiple vesting horizons that are commonly seen in practice. In our model, we constrain short-termism and myopic behaviors by imposing longer vesting periods. However, lengthening the vesting periods can be costly. Peng and Röell (2014) argue that long-term compensation potentially exposes the manager to risk outside his control and thus firms need to compensate the manager for bearing the additional risk. Therefore, allowing gradual vesting would make the optimal contract cheaper and the risk avoidance measure lower.

9   Conclusions

We argue that shareholders take into account risk-taking incentives when designing CEO contracts, because CEOs are often heavily exposed to firm-specific risk through their large stock and option holdings and bear the employment risk. If CEOs are risk averse, then they will want to reduce the firms' risk even if doing so destroys the value. We contribute to the literature by introducing risk-taking incentives into the standard contracting model. Specifically, CEOs in our model do not only exert costly effort but also determine the firm's strategy so that they affect both the mean and the volatility of future firm value. The contracts are designed to efficiently induce a given level of effort and choice of volatility from the CEO. Our calibration analysis demonstrates that the extended model explains the observed executive compensation contracts significantly better than the standard model without risk-taking incentives. We also propose a new measure of risk-taking incentives that captures the tradeoff between the expected value added to the firm and the additional risk a CEO must take. This measure essentially combines both the utility-adjusted vega and the utility-adjusted delta.

In this paper, we do not argue how strong those risk-taking incentives provided by firms should be. Instead, we take the observed strength of risk-taking incentives as given and search for the cheapest contract that provides those incentives. Besides, Lambert, Larcker, and Verrecchia (1991), Carpenter (2000), Ross (2004), and Lewellen (2006) argue and show that stock options can make managers more averse to increases in firm risk, so that stock options might be counter-productive if risk-taking incentives need to be provided. Our paper shows that options are indeed part of an optimal contract. They can be detrimental to risk-taking incentives, but wreak less havoc than stock.

Having neither stock nor options is not an alternative, because such a contract would not provide any effort incentives. In addition, we do not allow CEOs to hedge their exposure to firm risk. Gao (2010) shows that the CEO pay-for-performance sensitivity decreases with his hedging cost, so the firm reacts to the possibility of hedging by awarding even more options or stock to the CEO. This argument suggests that, for many firms, our risk avoidance measure would be lower in both models with and without risk-taking incentives. Finally, this paper takes an important step forward by modeling both risk and effort incentives, but future research is needed to determine how the results would change when we allow for gradual vesting, new grants, or contract renegotiation.

Appendix A: Validity of the first-order approach

Like most of the theoretical literature on executive compensation, we work with the first-order approach: we replace the incentive compatibility constraint (7) by the two first-order conditions (8) and (9). This approach is only valid if the utility that the agent maximizes has exactly one optimum, and it is a sufficient condition is that this utility is globally concave. In our model, this sufficient condition does not hold, and it is possible that the first-order approach is violated.

A violation of the first-order approach has two potential consequences. First, the agent might choose a different combination of effort e and volatility σ than under the observed contract. The reason is that our optimization routine only ensures that the pair $\{e^d, \sigma^d\}$ (which is implemented by the observed contract) remains a *local* optimum under the new contract, but we do not require it to be the global optimum (see Lambert and Larcker (2004)). Second, a violation of the first-order approach implies that there might be more than one solution to the optimization problem. We tackle the second problem by repeating our numerical optimizations with different starting values, but we do not find any indication that there are multiple solutions for any CEO in our sample. In this appendix, we therefore concentrate on the first problem. In particular, we analyze whether the agent has an incentive to shirk under the optimal contract $W^*(P_T)$, i.e., to choose effort $e \neq e^d$ or volatility $\sigma \neq \sigma^d$ such that $P_0(e, \sigma) < P_0^d = P_0(e^d, \sigma^d)$. We ignore deviations that lead to an increase of firm value as shareholders will not worry about this case. For expositional convenience, we say that the first-order approach is violated if the agent shirks under the optimal contract $W^*(P_T)$. In the remaining part of this appendix, we derive two conditions under which the first-order approach is not violated. To simplify the argument, we normalize $P_0(e = 0, \sigma) = P_0(e, \sigma = 0) = 0$ and $C(e = 0) = 0$.

Condition 1. *The agent has no incentives to choose $e = 0$ or $\sigma = 0$, i.e., $E(V(W_T^*)|P_0 = 0) < E(V(W_T^*)|P_0 = P_0^d) - C(e^d) = \underline{U}$.*

The optimal contract W_T^* from (11) features a lower bound on the payout to the agent. If this lower bound is higher than the agent's outside option $\underline{U}$, the agent might not exert any effort and might choose the lowest feasible volatility. Consequently, the first-order approach might be violated. Our first condition therefore states that this is not the case. This assumption appears reasonable, because for the median CEO, the minimum payout ($1.3m, from Table 2, Panel A for $\gamma = 3$) is only 8.1% of the expected payout ($16.0m, from Table 1). The strong rise in executive compensation during the past three decades has been attributed to a higher outside option or higher rents, but not to an increase in the costs of effort. Therefore, Condition 1 is plausible: No CEO will stop working

when he gets a minimum payment of 8.1% of what he can expect with normal effort.

Next, we consider more general (and less extreme) deviations from the target values of effort e^d and volatility σ^d. We show that these deviations are not profitable for the agent when Condition 1 and the following condition hold:

Condition 2. *The production function $P_0(e, \sigma)$ is concave enough, i.e., it is steep enough in e and σ for $e < e^d$ and $\sigma < \sigma^d$ and it is not too steep in e and σ for $e > e^d$ and $\sigma > \sigma^d$.*

We distinguish three cases. First, consider a choice $e \leq e^d$ and $\sigma \leq \sigma^d$, where $e < e^d$ or $\sigma < \sigma^d$. The agent will not deviate in this way if

$$E(V(W_T^*)|e, \sigma) - C(e) < E(V(W_T^*)|e^d, \sigma^d) - C(e^d).$$

This inequality holds if the firm value $P_0(e, \sigma)$ associated with the deviation to (e, σ) is low enough to render this choice unattractive. This is the case if Condition 1 holds and if $P_0(e, \sigma)$ is steep enough in e and σ.

The second case is obtained if $e < e^d$ and $\sigma > \sigma^d$. To rule out such a deviation, the punishment for the downward deviation in e must not be fully compensated by the reward for the upward deviation in σ. This is achieved if $P_0(e, \sigma)$ is steep enough in e for $e < e^d$ and not too steep in σ for $\sigma > \sigma^d$. A similar argument applies to the third case if $e > e^d$, $\sigma < \sigma^d$.

Appendix B: Proof of Proposition 1

Note that the monotonicity constraint (5) must hold for every P_T, so that it is actually a continuum of an infinitely number of restrictions. We first rewrite the restriction as a function of W_T. Let $h(.)$ be the function that maps P_T into W_T: $W_T = h(P_T)$. Then, $P_T = h^{-1}(W_T)$, and $\frac{dW_T}{dP_T}(P_T) = h'(h^{-1}(W_T))$. Hence, (5) can be rewritten as

$$h'(h^{-1}(W_T)) \geq 0. \tag{28}$$

For every W_T, (5) provides one restriction, so the Lagrangian for the differentiation at W_T is:

$$L_{W_T} = \int_0^\infty [P_T - W_T] g(P_T|e, \sigma) dP_T + \lambda_{PC} \left(\int_0^\infty V(W_T, e) g(P_T|e, \sigma) dP_T - C(e) - \underline{U} \right)$$

$$+ \lambda_e \left(\int_0^\infty V(W_T) g_e(P_T|e, \sigma) dP_T - \frac{dC}{de} \right) + \lambda_\sigma \int_0^\infty V(W_T) g_\sigma(P_T|e, \sigma) dP_T$$

$$+ \lambda_{W_T} h'(h^{-1}(W_T)),$$

34

where $g(P_T|e,\sigma)$ is the (lognormal) density function of the end-of-period stock price P_T:

$$g(P_T|e,\sigma) = \frac{1}{P_T\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{(\ln P_T - \mu(e,\sigma))^2}{2\sigma^2 T}\right] \tag{29}$$

with

$$\mu(e,\sigma) = \ln P_0(e,\sigma) + (r_f - \sigma^2/2)T. \tag{30}$$

g_e and g_σ are the derivatives of $g(.)$ with respect to e and σ. The first-order condition is then

$$g(P_T|e,\sigma) = \lambda_{PC}V_{W_T}g(P_T|e,\sigma) + \lambda_e V_{W_T}g_e(P_T|e,\sigma) + \lambda_\sigma V_{W_T}g_\sigma(P_T|e,\sigma) \tag{31}$$
$$+ \lambda_{W_T}\frac{h''(h^{-1}(W_T))}{h'(h^{-1}(W_T))}.$$

While there is one multiplier λ_{W_T} for each value of W_T, the other three multipliers λ_{PC}, λ_e, and λ_σ are the same across all values of W_T. If the constraint (28) is binding, equation (31) defines the Lagrange multiplier λ_{W_T}, and the solution is determined by the binding monotonicity constraint. If (28) is not binding, λ_{W_T} is zero and the first-order condition (31) simplifies with some rearranging to

$$\frac{1}{V_{W_T}(W_T)} = \lambda_{PC} + \lambda_e\frac{g_e}{g} + \lambda_\sigma\frac{g_\sigma}{g}. \tag{32}$$

Consequently, the solution is given by (32) as long as it is monotonically increasing, and flat otherwise.

For the log-normal distribution (29), we get:

$$g_e = g \cdot \frac{\ln P_T - \mu(e,\sigma)}{\sigma^2 T} \cdot \mu_e(e,\sigma) \tag{33}$$

$$g_\sigma = g \cdot \frac{[\ln P_T - \mu(e,\sigma)] \cdot \mu_\sigma(e,\sigma) \cdot \sigma^2 T + [\ln P_T - \mu(e,\sigma)]^2 \sigma T}{(\sigma^2 T)^2} - \frac{g}{\sigma}$$

$$= g \cdot \frac{[\ln P_T - \mu] \cdot \mu_\sigma \cdot \sigma + [\ln P_T - \mu]^2}{\sigma^3 T} - \frac{g}{\sigma}. \tag{34}$$

Substituting this into the first-order condition (32) yields:

$$\frac{1}{V_{W_T}(W_T)} = \lambda_{PC} + \lambda_e\frac{[\ln P_T - \mu] \cdot \mu_e}{\sigma^2 T} + \lambda_\sigma\left(\frac{[\ln P_T - \mu] \cdot \mu_\sigma \cdot \sigma + [\ln P_T - \mu]^2}{\sigma^3 T} - \frac{1}{\sigma}\right).$$

The optimal wage contract can be written as (10) with parameters $c_0(\sigma)$, $c_1(\sigma)$, and $c_2(\sigma)$:

$$c_0(\sigma) = \lambda_{PC} - \lambda_e\frac{\mu_e \cdot \mu}{\sigma^2 T} - \lambda_\sigma\left(\frac{\mu \cdot \mu_\sigma}{\sigma^2 T} - \frac{\mu^2}{\sigma^3 T} + \frac{1}{\sigma}\right),$$

$$c_1(\sigma) = \lambda_e\frac{\mu_e}{\sigma^2 T} + \lambda_\sigma\left(\frac{\mu_\sigma}{\sigma^2 T} - \frac{2\mu}{\sigma^3 T}\right),$$

$$c_2(\sigma) = \lambda_\sigma\frac{1}{\sigma^3 T} \geq 0.$$

Equation (11) then immediately follows with $V(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}$ for $\gamma \neq 1$ and $V(W_T) = ln(W_T)$ for $\gamma = 1.$∎

Appendix C: User's guide on how to calculate risk avoidance ρ

This appendix contains formulae for our measure of risk avoidance ρ from (16) that can readily be implemented in a computer program. We start with a few definitions:

$$PC = P_0 \exp\left\{\left(r_f - d - \frac{\sigma^2}{2}\right)T\right\}, \quad CV = \sigma\sqrt{T},$$

$$TW = (\phi + W_0)\exp\left\{r_f T\right\}, \qquad MD2 = \frac{\ln(K) - \ln(PC)}{CV},$$

$MD2$ is the point where options are just at the money. With these definitions, we can calculate PPS^{ua} and ν^{ua} as follows:

$$
\begin{aligned}
PPS^{ua} = & \frac{PC}{P_0} \left[\int_{-\infty}^{MD2} (TW + n_S \exp\{dT\} PC \exp\{CVu\})^{-\gamma} n_S \exp\{dT + CVu\} f(u)du \right. \\
& + \int_{MD2}^{\infty} (TW + (n_S \exp\{dT\} + n_O) PC \exp\{CVu\} - n_O K)^{-\gamma} \\
& \left. (n_S \exp\{dT\} + n_O) \exp\{CVu\} f(u)du \right] \\
\nu^{ua} = & \int_{-\infty}^{MD2} (TW + n_S \exp\{dT\} PC \exp\{CVu\})^{-\gamma} n_S \exp\{dT + CVu\} \\
& PC \left(-\sigma T + u\sqrt{T}\right) f(u)du \\
& + \int_{MD2}^{\infty} (TW + (n_S \exp\{dT\} + n_O) PC \exp\{CVu\} - n_O K)^{-\gamma} (n_S \exp\{dT\} + n_O) \\
& PC \exp\{CVu\} \left(-\sigma T + u\sqrt{T}\right) f(u)du,
\end{aligned}
$$

where $f(u)$ is the standard normal density function. Our measure of risk avoidance then follows from (16).

Appendix D: Representing the observed contract

Let N be the number of option grants. Each grant i is characterized by the strike price K^i, the maturity T^i, and the number of options n_O^i. We define

$$W_T^{smth}(P_T) := \phi e^{r_f T} + n_S P_T + \sum_{i=1}^{N} n_O^i E\left(\max\left\{P_{T^i} - K^i, 0\right\} | P_{min\{T^i, T\}}\right) e^{r_f max\{T - T^i, 0\}}. \quad (35)$$

If $T^i > T$, this is simply the Black-Scholes value of the option i over the remaining maturity $T^i - T$. If $T^i < T$, we assume that the option is exercised at time T^i if it is in the money and that the

proceeds are invested at the risk-free rate until time T.

Note that, for each option grant i with $T^i < T$, $W_T^{smth}(P_T)$ contains a separate integral with respect to the stock price at T^i conditional on P_T . Therefore, D_2 is an $(m+1)$-dimensional integral, where m is the number of option grants with $T^i < T$. As we cannot solve this numerically, we approximate D_2 by a sum over 1,001 equally spaced stock prices P_T over the range of stock prices that covers 99.9% of the probability mass.

Appendix E: Model convergence

In Table 2 and footnote 18, we experience numerical problems in the calibration and the convergence rates for $\gamma = 0.5$ and $\gamma = 6$ are 67% and 66%, respectively. To better understand the economics behind the numerical problem, we repeat our calibration process for $\gamma = 2, 4, 5$ in addition to 0.5, 3, 6 reported in Table 2. The convergence rates for the CRRA-RTI model are 97% for $\gamma = 2$, 93% for $\gamma = 4$, and 78% for $\gamma = 5$. It seems that our model and calibration work well for the more common range of the risk parameter γ ranging from 2 to 4, but not for the more extreme values such as $\gamma = 0.5$ and $\gamma = 6$.

Among 1,707 CEOs, we have 14 cases (less than 1%) where the calibration fails for all three values ($\gamma = 0.5, 3, 6$). For these 14 cases, we argue that our model is a bad description of that given firm/CEO pair. For the rest of the sample (more than 99%), we can calibrate our model for at least one of the assumed values of γ, suggesting that the model works for at least one parameter value of γ.

[Insert Table 10 here]

In addition, we split the sample into two based on whether the calibration in Table 2 Panel A converges and compare the two sub-samples for all variables listed in Table 1. The medians of the variables in each sub-sample as well as the p-value of the two-sample Wilcoxon test are displayed in Table 10. For $\gamma = 0.5$, we find no significant difference between the convergence and the non-convergence subsamples. For $\gamma = 3$, the only difference between the two sub-samples is that the percentage of CEO stock holdings is lower for the non-convergence group. For $\gamma = 6$, the non-convergence group has a higher base salary and non-firm wealth.

References

Acharya, Viral V., Kose John, and Rangarajan K. Sundaram, 2000, On the optimality of resetting executive stock options, *Journal of Financial Economics* 57, 65-101.

Acharya, Viral V., Yakov Amihud, and Lubomir Litov, 2011, Creditor rights and corporate risk-taking, *Journal of Financial Economics* 102, 150-166.

Alexander, Raquel Meyer, Mark Hirschey, and Susan Scholz, 2007, Backdating employee stock options: Tax implications, *CPA Journal* 77(10), 24-28.

Bebchuk, Lucien, Alma Cohen, and Allen Ferrell, 2009, What matters in corporate governance, *Review of Financial Studies* 22, 783-827.

Bebchuk, Lucien, and Jesse M. Fried, 2004, *Pay without performance - the unfulfilled promise of executive compensation* (Harvard University Press, Cambridge (MA) and London).

Carpenter, Jennifer N., 1998, The exercise and valuation of executive stock options, *Journal of Financial Economics* 48, 127-158.

Carpenter, Jennifer N., 2000, Does option compensation increase managerial risk appetite?, *Journal of Finance* 55, 2311-2331.

Chaigneau, Pierre, 2013a, Explaining the structure of CEO incentive pay with decreasing relative risk aversion, *Journal of Economics and Business* 67, 4-23.

Chaigneau, Pierre, 2013b, Risk-shifting and the regulation of bank CEOs' compensation, *Journal of Financial Stability* 9, 778-789.

Chaigneau, Pierre, 2015, The optimal timing of CEO Compensation, Working paper, HEC Montreal.

Chaigneau, Pierre, and Nicolas Sahuguet, 2015, The effect of monitoring on CEO pay practices in a matching equilibrium, Working paper, HEC Montreal.

Coles, Jeffrey L., Naveen D. Daniel, and Lalitha Naveen, 2006, Managerial incentives and risk-taking, *Journal of Financial Economics* 79, 431-468.

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, 151-184.

Core, John E., and Wayne R. Guay, 2002, Estimating the value of stock option portfolios and their sensitivities to price and volatility, *Journal of Accounting Research* 40, 613-640.

Core, John E., and Qian, Jun, 2002, Project selection, production, uncertainty, and incentives, Working paper, University of Pennsylvania.

Coughlan, Anne T., and Ronald M. Schmidt, 1985, Executive compensation, management turnover, and firm performance: An empirical investigation, *Journal of Accounting and Economics* 17, 43-66.

DeMarzo, Peter M., and Yuliy Sannikov, 2006, Optimal security design and dynamic capital structure in a continuous-time agency model, *Journal of Finance* 61, 2681-2724.

Dewatripont, Mathias, and Jean Tirole, 1999, Advocates, *Journal of Political Economy* 107, 1-39.

Dittmann, Ingolf, and Ernst Maug, 2007, Lower salaries and no options? On the optimal structure of executive pay, *Journal of Finance* 62, 303-343.

Dittmann, Ingolf, Ernst Maug, and Oliver Spalt, 2010, Sticks or carrots: Optimal CEO compensation when managers are loss-averse, *Journal of Finance* 65, 2015-2050.

Dittmann, Ingolf, Ernst Maug, and Oliver Spalt, 2013, Indexing executive compensation contracts, *Review of Financial Studies* 26, 3182-3224.

Edmans, Alex, and Qi Liu, 2011, Inside debt, *Review of Finance* 15, 75-102.

Edmans, Alex, and Xavier Gabaix, 2011, Tractability in incentive contracting, *Review of Financial Studies* 24, 2865-2894.

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.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov, 2012, Dynamic CEO compensation, *Journal of Finance*, 67, 1603-1647.

Faccio, Mara, Maria-Teresa Marchica, and Roberto Mura, 2011, Large shareholder diversification and corporate risk-taking, *Review of Financial Studies* 24, 3601-3641.

Fee, Edward C., and Charles J. Hadlock, 2004, Management turnover across the corporate hierarchy, *Journal of Accounting & Economics* 37, 3-38.

Feltham, Gerald A., and Martin G. H. Wu, 2001, Incentive efficiency of stock versus options, *Review of Accounting Studies* 6, 7-28.

Flor, Christian Riis, Hans Frimor, and Claus Munk, 2014, Options in compensation: Promises and pitfalls, *Journal of Accounting Research* 52, 703-732.

Gabaix, Xavier, and Augustin Landier, 2008, Why has CEO pay increased so much?, *Quarterly Journal of Economics* 123, 49-100.

Gao, Huasheng, 2010, Optimal compensation contracts when managers can hedge, *Journal of Financial Economics* 97, 218-238.

Goldman, E. and S.L. Slezak, 2006, An equilibrium model of incentive contracts in the presence of information manipulation, *Journal of Financial Economics* 80, 603-626.

Graham, John R., Campbell R. Harvey, and Manju Puri, 2013, Managerial attitudes and corporate actions, *Journal of Financial Economics* 109, 103-121.

Grossman, Sanford J., and Oliver D. Hart, 1983, An analysis of the principal-agent problem, *Econometrica* 51, 7-45.

Guay, Wayne R., 1999, The sensitivity of CEO wealth to equity risk: An analysis of the magnitude and determinants, *Journal of Financial Economics* 53, 43-71.

Habib, Michel A., and Alexander Ljungqvist, 2005, Firm value and managerial incentives: A stochastic frontier approach, *Journal of Business* 78, 2053-2093.

Hall, Brian J., and Jeffrey B. Liebman, 1998, Are CEOs really paid like bureaucrats?, *Quarterly Journal of Economics* 113, 653-691.

Hall, Brian J., and Kevin J. Murphy, 2000, Optimal exercise prices for executive stock options, *American Economic Review* 90, 209-214.

Hall, Brian J., and Kevin J. Murphy, 2002, Stock options for undiversified executives, *Journal of Accounting and Economics* 33, 3-42.

Haugen, Robert A., and Lemma W. Senbet, 1981, Resolving the agency problems of external capital through options, *Journal of Finance* 36, 629-647.

Hellwig, Martin F., 2009, A reconsideration of the Jensen-Meckling model of outside finance, *Journal of Financial Intermediation* 18, 495-525.

Hirshleifer, David and Yoon Suh, 1992, Risk, managerial effort and project choice, *Journal of Financial Intermediation* 2, 308-345.

Holmström, Bengt, 1979, Moral hazard and observability, *Bell Journal of Economics* 10, 74-91.

Holmström, Bengt, and Paul Milgrom, 1991, Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design, *Journal of Law, Economics, & Organization,* 24-52.

Huddart, Steven J., and Mark Lang, 1996, Employee stock option exercises - an empirical analysis, *Journal of Accounting and Economics* 21, 5-43.

Inderst, Roman, and Holger M. Müller, 2005, Benefits of Broad-Based Option Pay, CEPR Discussion Paper no. 4878.

Innes, Robert D, 1990, Limited liability and incentive contracting with ex-ante action choices, *Journal of Economic Theory* 52, 45-67.

Jensen, Michael C., 1986, Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, *American Economic Review* 76, 323-329.

Jenter, Dirk and Faid Kanaan, 2015, CEO turnover and relative performance evaluation, *Journal of Finance* 70, 2155-2184.

John, Kose, Anthony Saunders, and Lemma W. Senbet, 2000, A theory of bank regulation and management compensation, *Review of Financial Studies* 13, 95-125.

Kaplan, Steven N., 1994, Top executive rewards and firm performance: A comparison of Japan and the United States, *Journal of Political Economy* 102, 510-546.

Kaplan, Steven N., and Bernadette A. Minton, 2012, How has CEO turnover changed, *International Review of Finance* 12, 57-87.

Knopf, John D., Jouahn Nam, and John H. Thornton, Jr., 2002, The volatility and price sensitivities of managerial stock option portfolios and corporate hedging, *Journal of Finance* 57, 801-813.

Lambert, Richard A., 1986, Executive effort and selection of risky projects, *Rand Journal of Economics* 17, 77-88.

Lambert, Richard A., and David F. Larcker, 2004, Stock options, restricted stock, and incentives, discussion paper, Wharton School.

Lambert, Richard A., David F. Larcker, and Robert Verrecchia, 1991, Portfolio considerations in valuing executive compensation, *Journal of Accounting Research* 29, 129-149.

Lewellen, Katharina, 2006, Financing decisions when managers are risk averse, *Journal of Financial Economics* 82, 551-589.

Low, Angie, 2009, Managerial risk-taking behavior and equity-based compensation, *Journal of Financial Economics* 92, 470-490.

Manso, Gustavo, 2011, Motivating innovation, *Journal of Finance* 66, 1823-1860.

Murphy, Kevin J., 1999, Executive compensation, in O. Ashenfelter and D. Cards, eds.: *Handbook of Labor Economics* 3 (North Holland, Amsterdam).

Ou-Yang, Hui, 2003, Optimal contracts in a continuous-time delegated portfolio management problem, *Review of Financial Studies* 16, 173-208.

Peters, Florians S., and Alexander F. Wagner, 2014, The executive turnover risk premium, *Journal of Finance* 69, 1529-1563.

Peng, Lin, and Ailsa Röell, 2008, Manipulation and equity-based compensation, *American Economic Review: Papers & Proceedings* 98, 285-290.

Peng, Lin, and Ailsa Röell, 2014, Managerial incentives and stock price manipulation, *Journal of Finance* 69, 487-526.

Rajgopal, Shivaram, and Terry Shevlin, 2002, Empirical evidence on the relation between stock option compensation and risk taking, *Journal of Accounting and Economics* 33, 145-171.

Ross, Stephen A., 2004, Compensation, incentives, and the duality of risk aversion and riskiness, *Journal of Finance* 60, 207-225.

Rosser, Bruce A., and Jean M. Canil, 2004, Executive stock options: Evidence that premium and discount awards do matter, discussion paper, University of Adelaide.

Shue, Kelly, and Richard Townsend , 2014, Swinging for the Fences: Executive Reactions to Quasi-Random Option Grants, discussion paper, University of Chicago.

Smith, Clifford W., and René M. Stulz, 1985, The determinants of firms' hedging policies, *Journal of Financial and Quantitative Analysis* 20, 391-405.

Sundaram, Rangarajan K., and David Yermack, 2007, Pay me later: Inside debt and its role in managerial compensation, *Journal of Finance* 62, 1551-1588.

Sung, Jaeyoung, 1995, Linearity with project selection and controllable diffusion rate in continuous-time principal-agent problems, *RAND Journal of Economics* 26, 720-743.

Tversky, Amos, and Daniel Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297-323.

Walker, David I., 2009, The Non-Option: Understanding the Dearth of Discounted Employee Stock Options, *Boston University Law Review* 89, 1505-1563.

Yermack, David, 2006, Golden handshakes: Separation pay for retired and dismissed CEOs, Journal of Accounting and Economics 41, 237-256.

Zhu, John Y., 2013, Optimal contracts with shirking, *Review of Economic Studies* 80, 812-839.

Zhu, John Y, 2016, Myopic agency, Working paper, The Wharton School.

Table 1: Description of the dataset

This table displays mean, median, standard deviation, and the 10% and 90% quantile of the variables in our dataset. Stock holdings n_S and option holdings n_O are expressed as a percentage of all outstanding shares. Panel A describes our sample of 1707 CEOs in the augmented year 2012. Panel B displays descriptive statistics for risk avoidance ρ from equation (16) for six different values of the CRRA-parameter γ. Panel C compares 1,526 executives in the ExecuComp universe who are CEOs in 2012 and 1,196 ExecuComp CEOs in 2012 who are included in our sample. The last two columns of Panel C display the p-values of the two-sample t-test and the two-sample Wilcoxon test. Panel D summarizes the data coverage and statistics for corporate governance variables for the augmented year 2012, including the entrenchment index (E-index), the total ownership stake of all independent compensation committee members (CC-ownership), institutional ownership, and the presence of a 5% institutional blockholder.

Panel A: Data set with 1,707 U.S. CEOs

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Stock (%)	n_S	1.53%	4.43%	0.04%	0.35%	3.18%
Options (%)	n_O	0.86%	1.10%	0.00%	0.50%	2.17%
Base Salary ($m)	ϕ	3.04	3.43	0.71	2.02	6.52
Value of Contract ($m)	π_0	78.8	852.5	3.9	16.0	84.1
Non-firm Wealth ($m)	W_0	59.8	349.0	4.7	19.1	88.6
Firm Value ($m)	P_0	7,749	23,562	287	1,778	15,990
Strike Price ($m)	K	10,422	149,733	263	1,475	12,814
Moneyness (%)	K/P_0	109.7%	218.9%	44.1%	84.3%	151.3%
Maturity (years)	T	4.7	2.5	2.0	4.4	7.0
Stock Volatility (%)	σ	40.1%	38.0%	17.9%	33.5%	62.6%
Dividend Rate (%)	d	1.42%	2.17%	0.00%	0.62%	3.80%
CEO Age (years)		56.9	6.9	48	57	65
Past 5-Year Stock Return (%)		1.3%	16.8%	-16.9%	1.6%	19.3%

Panel B: Risk avoidance in the full sample

γ	Obs.	Mean	Standard Deviation	10% Quantile	Median	90% Quantile	Proportion with $\rho > 0$
0.5	1707	0.09	0.53	-0.53	0.06	0.67	59.1%
1	1707	0.44	0.73	-0.33	0.31	1.31	75.3%
2	1707	0.97	1.00	-0.03	0.75	2.27	88.7%
3	1707	1.36	1.18	0.11	1.11	2.95	94.1%
4	1707	1.67	1.32	0.25	1.40	3.49	95.6%
6	1707	2.14	1.51	0.47	1.87	4.18	97.6%

Panel C: Comparison of the ExecuComp universe and our sample in 2012

Variable		ExecuComp Universe 1,526 CEOs in 2012		Our Sample 1,196 CEOs in 2012		Difference (p-value)	
		Mean	Median	Mean	Median	Mean	Median
Stock (%)	n_S	1.82%	0.35%	1.53%	0.35%	0.10	0.93
Options (%)	n_O	1.12%	0.48%	0.77%	0.41%	0.00	0.04
Fixed Salary ($m)	ϕ	3.01	2.01	3.22	2.15	0.11	0.02
Firm Value ($m)	P_0	8,001	1,761	8,789	1,984	0.41	0.01
Age		56.5	56	56.7	56.0	0.47	0.36
Return 2007-2011 (%)		0.2%	0.6%	0.7%	1.1%	0.50	0.15

Panel D: Corporate governance variables for the augmented year 2012

	Coverage		Statistics				
	#	%	Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Data from Institutional Shareholder Services (formerly RiskMetrics)							
E-index	1523	89%	2.2	1.2	1	2	4
CC-Ownership	1289	76%	.32%	1.90%	.00%	.03%	.49%
Data from Thomson Reuters Form 13F institutional holdings							
Institutional ownership	1544	90%	79%	18%	55%	81%	99%
Presence of a 5% blockholder	1544	0.90	0.92	0.27	1	1	1

Table 2: Optimal CRRA contracts
with and without risk-taking incentives

This table describes the optimal contracts according to the CRRA-RTI model from equation (11) and the CRRA model from Dittmann and Maug (2007) for three different values of the CRRA parameter γ. The table displays the mean and the median of seven measures that describe the optimal contract. The two distance metrics D_1 and D_2 are defined in equations (23) and (24). *Savings* are the difference in compensation costs between the observed contract and the optimal contract expressed as a percentage of the costs of the observed contract. *Minimum wealth* is the lowest possible payout of the contract expressed as a multiple of the CEO's nonfirm wealth W_0. The kink quantile is the point where the contract shape starts to increase, and the inflection quantile is the point where it turns from convex to concave. Both the kink quantile and the inflection quantile are expressed as probabilities. *Risk avoidance ρ* is from equation (16). Panel A displays these statistics for all CEOs in our sample. Panel B shows results for those CEO-γ-combinations where we obtain the convergence for both models.

Panel A: All results

		CRRA-RTI Model			CRRA Model		
		$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$	$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$
Distance D_1	mean	2.3%	5.4%	7.2%	13.0%	16.1%	21.9%
	median	1.7%	4.0%	5.5%	9.6%	13.4%	17.0%
Distance D_2	mean	4.8%	6.3%	7.7%	11.7%	13.9%	20.9%
	median	3.0%	5.0%	5.5%	9.1%	10.9%	16.3%
Savings	mean	0.7%	10.3%	20.3%	3.8%	17.7%	33.0%
	median	0.1%	4.4%	13.9%	1.2%	10.7%	30.5%
Minimum wealth	mean	1.5	1.5	1.3	0.0	0.0	0.0
	median	1.3	1.3	1.2	0.0	0.0	0.0
	Prop < 1	48.9%	0.0%	27.8%	100.0%	100.0%	100.0%
Kink quantile	mean	18.7%	25.9%	24.2%	0.0%	0.0%	0.0%
	median	10.3%	21.6%	20.5%	0.0%	0.0%	0.0%
Inflection quantile	mean	64.8%	48.6%	39.5%	0.0%	0.0%	0.0%
	median	74.6%	47.5%	37.0%	0.0%	0.0%	0.0%
Risk avoidance ρ	mean	0.14	1.37	1.94	1.68	4.91	11.40
	median	0.10	1.12	1.70	1.36	4.98	13.05
Observations		1151	1658	1124	1695	1441	1051

Panel B: Results where numerical routine converges for both models

		CRRA-RTI Model			CRRA Model		
		$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$	$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$
Distance D_1	mean	2.3%	5.4%	6.6%	14.0%	15.9%	15.9%
	median	1.7%	4.2%	4.9%	10.1%	13.4%	13.9%
Distance D_2	mean	4.9%	6.3%	6.8%	12.7%	13.6%	15.5%
	median	3.1%	5.2%	4.9%	9.4%	10.9%	13.2%
Savings	mean	0.7%	9.3%	15.2%	4.3%	16.3%	30.7%
	median	0.1%	4.2%	8.8%	1.3%	10.7%	26.4%
Observations		1149	1402	694	1149	1402	694

Table 3: Risk avoidance and model fit for some sample splits on leverage and governance

This table considers sample splits on whether the company in question is a bank (SIC codes 6000-6099 and 6200-6299) or a non-bank (any SIC codes except for 6000-6999), and a median split of market leverage, book leverage, the entrenchment index (E-index), the total ownership stake of all independent compensation committee members (CC-ownership), institutional ownership, and the presence of a 5% blockholder. This table shows median *risk avoidance* ρ from equation (16) and median distance D_1 from equation (23). The table also displays the *p*-values of the two-sample Wilcoxon test. All calculations are for $\gamma = 3$.

Variable	Subsamples		Median ρ in		Wilco-xon	Median D_1 in		Wilco-xon	Obs.	
	S1	S2	S1	S2		S1	S2		S1	S2
Banks	yes	no	0.84	1.16	0.00	3.0%	4.4%	0.00	126	1318
Market leverage | non-bank	high	low	1.10	1.22	0.04	4.0%	4.8%	0.00	660	652
Book leverage | non-bank	high	low	1.07	1.24	0.00	4.1%	4.6%	0.04	660	653
E-Index	high	low	1.22	1.00	0.01	4.4%	3.8%	0.00	517	918
CC-Ownership	high	low	1.25	0.90	0.00	4.7%	3.3%	0.00	608	609
Institutional ownership	high	low	1.18	0.96	0.02	4.3%	3.7%	0.00	742	715
Presence of a 5% blockholder	yes	no	1.11	0.73	0.01	4.0%	3.1%	0.01	1347	110

Table 4: Optimal contracts that consist of salary, stock, and options

This table describes the optimal piecewise linear contract for five different values of the CRRA parameter γ. The table displays the mean and the median of the four contract parameters: base salary ϕ^*, stock holdings n_S^*, option holdings n_O^*, and the moneyness, i.e. the option strike price K^* scaled by the stock price P_0. Savings are the difference in compensation costs between observed contracts and optimal contracts as a percentage of total (observed) pay. The number of observations varies across different values of γ due to numerical problems and because we exclude all CEO-γ-combinations for which the observed contract implies negative risk-avoidance ρ from equation (16).

		$\gamma = 0.5$	$\gamma = 2$	$\gamma = 3$	$\gamma = 4$	$\gamma = 6$
Salary ϕ^*	mean	3.57	5.93	6.01	6.27	5.89
	median	2.11	3.41	3.39	3.36	3.18
Stock n_S^*	mean	0.75%	0.69%	0.59%	0.67%	0.54%
	median	0.00%	0.00%	0.00%	0.00%	0.00%
Options n_O^*	mean	1.43%	1.44%	1.25%	1.13%	0.88%
	median	1.00%	0.97%	0.85%	0.73%	0.57%
Moneyness K^*/P_0	mean	71.5%	56.6%	54.6%	55.7%	48.7%
	median	59.9%	48.8%	46.7%	44.1%	38.8%
	Prop.<K/P0	95.7%	96.7%	97.0%	96.6%	97.3%
Savings	mean	0.3%	3.0%	5.7%	8.2%	13.2%
	median	0.1%	0.9%	2.1%	3.7%	7.8%
Observations		620	1151	1208	1173	991

Table 5: Model fit for subsamples

This table shows mean distance D_1 from equation (23) for quintiles formed according to four variables: CEO option holdings, fixed salary, initial non-firm wealth W_0, and firm value P_0. The risk-aversion parameter γ is set equal to 3. The last row shows the p-value of the two-sample Wilcoxon signed rank test that the average D_1 is identical in Quintile 1 and Quintile 5.

Quin-tile	Options (%)		Fixed Salary (in $m)		Wealth W_0 (in $m)		Firm Value P_0 (in $m)	
	Mean	D_1	Mean	D_1	Mean	D_1	Mean	D_1
1	0.01%	4.1%	0.68	6.7%	4.5	7.1%	295	5.1%
2	0.19%	4.0%	1.30	6.0%	10.8	6.2%	863	5.7%
3	0.50%	5.3%	2.03	5.1%	19.2	5.0%	1,848	5.4%
4	1.02%	6.1%	3.25	5.1%	35.7	4.3%	4,407	5.1%
5	2.58%	7.2%	7.93	3.9%	229.0	4.1%	31,360	5.4%
P-Value Q1-Q5	0.00		0.00		0.00		0.82	

Table 6: Optimal CRRA-RTI contracts for years 1997-2012

This table describes the optimal contracts according to the CRRA-RTI model from equation (11) for $\gamma = 3$. The table displays the mean and the median of seven measures that describe the optimal contract. The two distance metrics D_1 and D_2 are defined in equations (23) and (24). *Savings* are the difference in compensation costs between the observed contract and the optimal contract expressed as a percentage of the costs of the observed contract. *Minimum wealth* is the lowest possible payout of the contract expressed as a multiple of the CEO's nonfirm wealth W_0. The kink quantile is the point where the contract shape starts to increase, and the inflection quantile is the point where it turns from convex to concave. Both the kink quantile and the inflection quantile are expressed as probabilities. *Risk avoidance ρ* is from equation (16).

		1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Distance D_1	mean	8.0%	9.1%	9.6%	9.3%	9.0%	8.9%	6.6%	5.7%	5.6%	5.5%	8.6%	9.9%	6.6%	5.5%	6.5%	4.9%
	median	6.9%	8.0%	8.2%	8.3%	7.9%	7.7%	5.4%	4.5%	4.2%	3.8%	6.8%	8.4%	5.5%	4.5%	5.5%	3.7%
Distance D_2	mean	8.0%	9.0%	9.5%	9.2%	9.2%	8.9%	6.7%	5.8%	5.7%	5.5%	9.2%	11.6%	7.4%	6.3%	7.3%	5.8%
	median	6.9%	8.0%	8.2%	8.2%	8.0%	7.7%	5.5%	4.6%	4.3%	3.8%	8.0%	9.4%	5.8%	5.2%	6.3%	4.5%
Savings	mean	13.0%	21.0%	21.6%	27.7%	25.5%	23.6%	13.0%	9.1%	7.7%	7.1%	12.5%	33.2%	22.1%	8.7%	11.8%	7.2%
	median	7.1%	14.2%	16.5%	21.1%	19.9%	17.5%	7.6%	4.4%	3.7%	3.3%	8.0%	29.4%	16.5%	5.3%	7.6%	3.1%
Minimum wealth	mean	1.6	2.0	1.5	1.4	1.3	1.4	1.4	1.5	1.5	1.4	1.5	1.5	1.3	1.4	1.4	1.5
	median	1.4	1.3	1.3	1.3	1.2	1.2	1.3	1.3	1.3	1.2	1.3	1.2	1.2	1.3	1.3	1.3
Kink quantile	mean	22.0%	28.6%	30.9%	37.5%	35.4%	35.2%	29.9%	23.6%	21.7%	20.2%	24.1%	41.8%	42.3%	26.7%	27.8%	21.6%
	median	17.9%	23.1%	26.0%	32.2%	30.4%	30.8%	24.3%	18.7%	17.6%	16.7%	20.2%	39.2%	40.0%	24.1%	26.0%	18.7%
Inflection quantile	mean	40.1%	43.7%	45.9%	51.2%	49.2%	49.6%	48.8%	44.0%	43.5%	43.0%	39.9%	53.9%	59.1%	49.0%	47.7%	46.4%
	median	37.8%	40.5%	43.4%	49.4%	47.7%	48.2%	47.0%	42.2%	42.3%	41.9%	37.9%	52.8%	58.6%	47.9%	47.7%	45.6%
Risk avoidance	mean	2.32	2.71	2.71	2.89	2.95	2.76	2.10	1.96	1.82	1.77	1.96	2.42	1.71	1.30	1.49	1.25
	median	2.22	2.61	2.65	2.83	2.87	2.68	1.94	1.80	1.66	1.64	1.79	2.05	1.44	1.16	1.31	1.01
Observations		518	579	578	556	582	719	742	810	806	705	701	927	922	1091	1141	1132

Table 7: Risk avoidance with CEO dismissals

This table displays descriptive statistics for risk avoidance ρ from equation (16) for six different values of the CRRA-parameter γ. In order to specify the probability of dismissal, we estimate a logit regression in which the dependent variable is equal to one if a CEO who is in the data set in 2008 leaves the company within the next 5 years and if ExecuComp records "resigned" as the reason for leaving. We regress this dummy variable on the 5-year stock return from 2008-2012. The parameter estimates (standard errors) are -3.743 (0.163) for the intercept and -0.358 (0.144) for the slope. The risk avoidance is then calculated using equation (25).

γ	Obs.	Mean	Standard Deviation	10% Quantile	Median	90% Quantile	Proportion with $\rho > 0$
0.5	1707	0.10	0.53	-0.51	0.08	0.70	60.4%
1	1707	0.47	0.74	-0.30	0.33	1.35	77.7%
2	1707	1.02	1.02	0.01	0.81	2.32	90.2%
3	1707	1.42	1.20	0.15	1.17	3.06	94.7%
4	1707	1.74	1.33	0.28	1.47	3.57	95.8%
6	1707	2.23	1.53	0.52	1.97	4.35	97.9%

Table 8: Optimal LA contracts with and without risk-taking incentives

Panel A describes the optimal contract according to the LA model from Dittmann, Maug, and Spalt (2010) for three different levels of reference wealth W^R parameterized by θ. Panel B describes the optimal contracts according to the LA-RTI model. The table displays the mean and the median of six measures that describe the optimal contract. The two distance metrics D_1 and D_2 are defined in equations (23) and (24). *Savings* are the difference in compensation costs between the observed contract and the optimal contract expressed as a percentage of the costs of the observed contract. *Jump quantile* is the point where the contract features a jump from the lowest possible payout to some payout above the reference wealth. *Inflection quantile* is the point where the contract turns from convex to concave. Both the jump quantile and the inflection quantile are expressed as probabilities. *Risk avoidance ρ* is from equation (16). The number of observations varies across different values of θ due to numerical problems.

Panel A: Optimal LA contracts without risk-taking incentives

		$\theta = 0.1$	$\theta = 0.5$	$\theta = 0.9$
Distance D_1	mean	4.5%	12.4%	29.6%
	median	2.3%	8.5%	26.5%
Distance D_2	mean	5.1%	11.6%	27.1%
	median	2.4%	8.5%	24.5%
Savings	mean	1.0%	4.9%	14.6%
	median	0.1%	3.6%	13.8%
Jump quantile	mean	0.6%	3.0%	11.8%
	median	0.0%	1.4%	9.8%
Inflection quantile	mean	100%	100%	100%
	median	100%	100%	100%
Risk avoidance ρ	mean	0.16	0.45	2.34
	median	0.03	0.19	2.17
Observations		1472	1259	971

Panel B: Optimal LA contracts with and without risk-taking incentives

		LA-RTI Model			LA Model		
		$\theta = 0.1$	$\theta = 0.5$	$\theta = 0.9$	$\theta = 0.1$	$\theta = 0.5$	$\theta = 0.9$
Distance D_1	mean	13.4%	16.5%	57.7%	14.4%	16.9%	40.5%
	median	9.1%	15.2%	50.4%	10.2%	15.9%	42.8%
Distance D_2	mean	13.8%	15.1%	50.0%	14.0%	15.5%	37.2%
	median	11.9%	13.4%	41.1%	11.1%	14.5%	37.4%
Savings	mean	4.9%	6.4%	8.1%	3.0%	6.7%	8.8%
	median	1.6%	6.5%	8.3%	2.1%	6.8%	8.8%
Jump quantile	mean	1.5%	2.0%	2.4%	5.2%	5.2%	18.1%
	median	0.6%	0.0%	0.0%	3.3%	4.6%	18.1%
Inflection quantile	mean	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	median	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Observations		43	152	59	43	152	59

Table 9: The convexity of CRRA contracts with risk-taking incentives

This table describes the convexity of the optimal contracts according to the CRRA-RTI model from equation (11) for three different values of the CRRA parameter γ. $P_{x\%}$ denotes the xth percentile of P_T. The distance metrics $D_2^{x\%}$ is defined in equation (28).

		$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$
Proportion of the model contracts that are convex at $P_{x\%}$	$x = 80$	54.1%	11.2%	1.1%
	$x = 90$	44.2%	3.1%	0.3%
	$x = 95$	28.1%	0.8%	0.2%
Proportion of the model contracts that are more convex than the observed contracts at $P_{x\%}$	$x = 80$	49.1%	0.7%	0.0%
	$x = 90$	32.0%	0.2%	0.0%
	$x = 95$	19.6%	0.1%	0.0%
Median ($D_2^{x\%}$ / D_2)	$x = 80$	27.6%	35.2%	52.0%
	$x = 90$	18.2%	25.7%	41.5%
	$x = 95$	10.2%	19.6%	27.4%
Observations		1151	1658	1124

Table 10: Convergence and non-convergence cases in Table 2

This table shows the median of the variables in our dataset for the convergence and non-convergence cases in Table 2. The *p*-values of the two-sample Wilcoxon test is also displayed.

Variable		Convergence			Non-Convergence			Difference (p-value)		
		$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$	$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$	$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$
Stock (%)	n_S	0.36%	0.35%	0.33%	0.33%	0.17%	0.29%	0.57	0.01	0.67
Options (%)	n_O	0.48%	0.50%	0.49%	0.53%	0.48%	0.40%	0.84	0.39	0.13
Base Salary ($m)	ϕ	2.00	2.02	2.03	2.07	2.23	2.65	0.77	0.78	0.00
Value of Contract ($m)	π_0	16.5	16.0	16.1	14.8	12.5	15.9	0.29	0.14	0.37
Non-firm Wealth ($m)	W_0	19.9	19.0	18.5	17.5	24.6	23.6	0.25	0.61	0.02
Firm Value ($m)	P_0	1,874	1,781	1,855	1,662	1,753	2,252	0.38	0.84	0.07
Strike Price ($m)	K	1,518	1,477	1,549	1,453	1,564	1,841	0.67	0.30	0.08
Moneyness (%)	K/P_0	83.9%	84.2%	85.4%	86.4%	86.9%	84.6%	0.53	0.12	0.82
Maturity (years)	T	4.4	4.4	4.3	4.5	4.2	4.5	0.64	0.54	0.93
Stock Volatility (%)	σ	33.0%	33.6%	32.8%	33.9%	31.3%	32.0%	0.67	0.33	0.73
Dividend Rate (%)	d	0.70%	0.61%	0.92%	0.38%	1.09%	0.92%	0.23	0.25	0.54
CEO Age (years)		57	57	57	57	57	57	0.97	0.69	0.38
Stock Return 2007-2011 (%)		1.9%	1.7%	1.6%	1.2%	0.1%	1.3%	0.83	0.10	0.81
Observations		1,151	1,658	1,124	556	49	583			

Internet Appendix

(Temporarily attached here)

Optimal loss-aversion contract

Proposition 2. (Optimal LA contract): *Under the assumptions that (i) the agent is loss-averse as described in (3) and (26) and (ii) the stock price P_T is lognormally distributed as described in (2), the optimal contract $W^*(P_T)$ that solves the shareholders' problem (4), (5), (6), (8), and (9) is:*

$$W_T^{*,LA} = \begin{cases} W^R + [\widetilde{w}(P_T)]^{\frac{1}{1-\alpha}} & \text{if } P_T > \widehat{P} \\ 0 & \text{if } P_T \leq \widehat{P} \end{cases}, \tag{36}$$

where $\widetilde{w}(P_T) := c_0 + c_1 \ln P_T + c_2 (\ln P_T)^2$ and $\widehat{P}$ is the largest solution to

$$\alpha W^R = \widetilde{w}(P_T) \lambda \left(W^R\right)^\beta + (1-\alpha) \left(\widetilde{w}(P_T)\right)^{\frac{1}{1-\alpha}}. \tag{37}$$

If no solution for $\widehat{P}$ exists to (37), the optimal contract is

$$W_T^{*,LA} = \begin{cases} W^R + [\widetilde{w}(P_T)]^{\frac{1}{1-\alpha}} & \text{if } ln(P_T) > -\frac{c_1}{2c_2} \\ W^R + \left(c_0 - \frac{c_1^2}{4c_2}\right)^{\frac{1}{1-\alpha}} & \text{if } ln(P_T) \leq -\frac{c_1}{2c_2} \end{cases}. \tag{38}$$

The parameters c_0, c_1, and c_2 depend on the distribution of P_T and the Lagrange multipliers of the optimization problem, with $c_2 > 0$.

Lemma 1 in Appendix A in Dittmann, Maug and Spalt (2010) continues to hold. This lemma states that the optimal contract never pays off in the interior of the loss space. Together with the assumption that the optimal contract is monotonically increasing, this immediately implies that either the contract pays out in the gain space only or there exists a cut-off value $\widehat{P}$ such that the optimal contract pays out in the gain space for all $P_T > \widehat{P}$ and 0 for all $P_T < \widehat{P}$. Therefore, we can rewrite the optimization problem as:

$$\min_{\widehat{P}, W_T \geq W^R} \int_{\widehat{P}}^{\infty} W_T g(P_T|e, \sigma) dP_T \tag{39}$$

$$s.t. \int_{\widehat{P}}^{\infty} V(W_T) g(P_T|e, \sigma) dP_T + V(0) G(\widehat{P}|e, \sigma) \geq \underline{U} + C(e), \tag{40}$$

$$\int_{\widehat{P}}^{\infty} V(W_T) g_e(P_T|e, \sigma) dP_T + V(0) G_e(\widehat{P}|e, \sigma) \geq C'(e), \tag{41}$$

$$\int_{\widehat{P}}^{\infty} V(W_T) g_\sigma(P_T|e, \sigma) dP_T + V(0) G_\sigma(\widehat{P}|e, \sigma) \geq 0. \tag{42}$$

Here, $G(P_T)$ is the cumulative distribution function of the lognormal stock price distribution. To keep the proof simple, we do not add the monotonicity constraint to the program at this point. Further below, we check whether the solution to this program satisfies the monotonicity constraint.

The derivative of the Lagrangian with respect to W_T at each point $P_T \geq \widehat{P}$ is:

$$\frac{\partial \mathcal{L}}{\partial W_T} = g(P_T|e, \sigma) - \lambda_{PC} V'(W_T) g(P_T|e, \sigma) - \lambda_e V'(W_T) g_e(P_T|e, \sigma)$$

$$- \lambda_\sigma V'(W_T) g_\sigma(P_T|e, \sigma) \tag{43}$$

Setting (43) to zero and solving gives the optimal contract in the gain space as:

$$V'(W_T) = \left[\lambda_{PC} + \lambda_e \frac{g_e(P_T|e, \sigma)}{g(P_T|e, \sigma)} + \lambda_\sigma \frac{g_\sigma(P_T|e, \sigma)}{g(P_T|e, \sigma)} \right]^{-1}. \tag{44}$$

For the Tversky and Kahneman (1992) preferences (26), we can rewrite (44) as:

$$W_T = W^R + \left[\alpha \left(\lambda_{PC} + \lambda_e \frac{g_e(P_T|e, \sigma)}{g(P_T|e, \sigma)} + \lambda_\sigma \frac{g_\sigma(P_T|e, \sigma)}{g(P_T|e, \sigma)} \right) \right]^{\frac{1}{1-\alpha}}. \tag{45}$$

Substituting the relevant expressions for the lognormal distribution from (33) and (34) and rearranging yields

$$W_T = W^R + \left[c_0 + c_1 \ln P_T + c_2 (\ln P_T)^2 \right]^{\frac{1}{1-\alpha}}, \tag{46}$$

where

$$c_0 = \alpha\lambda_{PC} - \alpha\lambda_e \frac{\mu_e \cdot \mu}{\sigma^2 T} - \alpha\lambda_\sigma \left(\frac{\mu \cdot \mu_\sigma}{\sigma^2 T} - \frac{\mu^2}{\sigma^3 T} + \frac{1}{\sigma} \right), \tag{47}$$

$$c_1 = \alpha\lambda_e \frac{\mu_e}{\sigma^2 T} + \alpha\lambda_\sigma \left(\frac{\mu_\sigma}{\sigma^2 T} - \frac{2\mu}{\sigma^3 T} \right), \tag{48}$$

$$c_2 = \frac{\alpha\lambda_\sigma}{\sigma^3 T} \geq 0. \tag{49}$$

Equation (46) provides the shape of the optimal contract for $P \geq \widehat{P}$ - provided that it is monotonic.

To find $\widehat{P}$ we take the derivative of the Lagrangian with respect to $\widehat{P}$:

$$\frac{\partial\mathcal{L}}{\partial\widehat{P}} = \left(-W(\widehat{P}) \right) g(\widehat{P}|e,\sigma) + \lambda_{PC} \left(V(W(\widehat{P})) - V(0) \right) g(\widehat{P}|e,\sigma)$$

$$+ \lambda_e \left(V(W(\widehat{P})) - V(0) \right) g_e(\widehat{P}|e,\sigma) + \lambda_\sigma \left(V(W(\widehat{P})) - V(0) \right) g_\sigma(\widehat{P}|e,\sigma) \tag{50}$$

$$= - \left(V(W(\widehat{P})) - V(0) \right) g(\widehat{P}|e,\sigma) \left[\frac{W(\widehat{P})}{V(W(\widehat{P})) - V(0)} - \lambda_{PC} - \lambda_e \frac{g_e\left(\widehat{P}|e,\sigma\right)}{g\left(\widehat{P}|e,\sigma\right)} - \lambda_\sigma \frac{g_\sigma\left(\widehat{P}|e,\sigma\right)}{g\left(\widehat{P}|e,\sigma\right)} \right].$$

$$\tag{51}$$

This derivative of the Lagrangian is zero if the term in squared brackets in (51) is zero. Substituting equation (44) and rearranging yields:

$$\frac{\partial\mathcal{L}}{\partial\widehat{P}} = 0 \Leftrightarrow V(W(\widehat{P})) - V(0) - V'\left(W\left(\widehat{P}\right) \right) W(\widehat{P}) = 0. \tag{52}$$

With Tversky and Kahneman (1992) preferences (26) we obtain:

$$\alpha W(\widehat{P}) - \lambda \left(W^R \right)^\beta \left(W(\widehat{P}) - W^R \right)^{1-\alpha} - \left(W(\widehat{P}) - W^R \right) = 0. \tag{53}$$

With (46) equation (53) becomes:

$$\alpha W^R = \left(c_0 + c_1 \ln\widehat{P} + c_2(\ln\widehat{P})^2 \right) \lambda \left(W^R \right)^\beta + (1-\alpha) \left(c_0 + c_1 \ln\widehat{P} + c_2(\ln\widehat{P})^2 \right)^{\frac{1}{1-\alpha}}. \tag{54}$$

This equation defines the threshold $\widehat{P}$.

As the wage function W_T from (46) is quadratic, the solution to condition (54) is not unique and might even not exist at all. If no solution exists, the contract always pays off in the gain space,

because paying off only in the loss space (i.e. always the minimum wealth 0) violates the participation constraint. With the same argument as the one put forth in the proof of Proposition 1, the optimal contract is then given by (46) as long as this function is monotone increasing. Otherwise, the optimal contract is constant. This proves (38).

Condition (54) might have exactly one solution, but this is a non-generic case. Generically, if there is one solution, there is also a second solution. Then, the general LA contract pays out in the gain space for very low and very high stock prices, while it pays the minimum wage for an intermediate range. Due to the monotonicity constraint, however, the contract is forced to pay out the minimum wage for all stock prices below the bigger of the two solutions to (54), and this proves (36). ∎

Constant absolute risk aversion contract

Corollary 1. *(Optimal CARA contract): If the agent exhibits constant absolute risk aversion with parameter η, the optimal contract has the following functional form:*

$$
W_T^* = \begin{cases} \frac{1}{\eta} \log \left\{ \eta \left[c_0 + c_1 \ln P_T + c_2 (\ln P_T)^2 \right] \right\} & \text{if } ln(P_T) > -\frac{c_1}{2c_2} \\ \frac{1}{\eta} \log \left\{ \eta \left[c_0 - \frac{c_1^2}{4c_2} \right] \right\} & \text{if } ln(P_T) \leq -\frac{c_1}{2c_2} \end{cases} \tag{55}
$$

To maintain comparability with our previous results, we calculate the coefficient of absolute risk aversion η from γ so that both utility functions exhibit the same risk-aversion at the expected end-of-period wealth. More precisely, we set $\eta = \gamma/(W_0 \exp(r_f T) + \pi_0)$, where π_0 is the market value of the manager's contract. The results are shown in Table IA.1.

Table IA.1 demonstrates that all our results continue to hold with CARA utility. In particular, the CARA-RTI model generates a much better fit than the CARA model as it guarantees a minimum payout that is always higher than the CEO's non-firm wealth, and it is convex for intermediate payouts and concave for good payouts. According to the two distance measures, CRRA-RTI (see Table 2) dominates CARA-RTI.

Table IA.1: Optimal contracts for CARA utility

This table contains the results from repeating our analysis from Table 2 under the assumption that the CEO has CARA utility. For three different values of γ, we calculate the CEO's coefficient of absolute risk aversion ρ as $\rho = \gamma / (W_0 \exp(r_f T) + \pi_0)$, where π_0 is the market value of his observed compensation package and W_0 is his initial non-firm wealth. The table displays the mean and the median of seven measures that describe the optimal contract. The two distance metrics D_1 and D_2 are defined in equations (23) and (24). *Savings* are the difference in compensation costs between the observed contract and the optimal contract expressed as a percentage of the costs of the observed contract. *Minimum wealth* is the lowest possible payout of the contract expressed as a multiple of the CEO's nonfirm wealth W_0. The kink quantile is the point where the contract shape starts to increase, and the inflection quantile is the point where it turns from convex to concave. Both the kink quantile and the inflection quantile are expressed as probabilities. *Risk avoidance ρ* is from equation (16). The results are shown for those CEO-γ-combinations only where we obtain convergence for both models.

		CARA-RTI Model			CARA Model		
		$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$	$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$
Distance D_1	mean	4.8%	6.1%	7.7%	17.6%	14.8%	15.0%
	median	3.3%	5.0%	6.6%	13.0%	12.7%	13.3%
Distance D_2	mean	6.7%	7.0%	8.1%	16.2%	13.3%	14.3%
	median	4.8%	5.7%	6.6%	11.6%	11.3%	12.6%
Savings	mean	6.0%	14.6%	19.7%	10.4%	23.0%	29.7%
	median	1.0%	8.7%	15.3%	3.9%	18.3%	27.7%
Minimum wealth	mean	1.6	1.6	1.5	0.0	0.0	0.0
	median	1.3	1.3	1.3	0.0	0.0	0.0
	Prop < 1	0.0%	0.0%	0.0%	100.0%	100.0%	100.0%
Kink quantile	mean	34.6%	29.1%	23.2%	0.0%	0.0%	0.0%
	median	30.6%	26.1%	19.8%	0.0%	0.0%	0.0%
Inflection quantile	mean	68.4%	49.0%	37.2%	0.0%	0.0%	0.0%
	median	68.5%	47.7%	35.4%	0.0%	0.0%	0.0%
Risk avoidance ρ	mean	0.50	1.42	1.91	2.51	3.92	5.36
	median	0.30	1.24	1.76	2.11	3.82	5.30
Observations		812	1161	1487	812	1161	1487

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI produces and disseminates high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It draws on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.org

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.org/wp) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Finance.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html

www.ecgi.org\wp