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

Ingolf Dittmann[†]   Ko-Chia Yu[‡]

December 11, 2010

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 727 CEOs and show that it can explain 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. Moreover, we propose a new measure of risk-taking (dis)incentives that measures the required profitability an additional risky project must exceed in order to be adopted by the CEO.

JEL Classifications: G30, M52

Keywords: Stock Options, Effort Aversion, Loss 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.

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

1 Introduction

This paper addresses the question to what extent the inclusion of risk-taking incentives in the standard model of executive compensation helps to rationalize observed compensation practice qualitatively and quantitatively. Our point of departure is the Holmström (1979) model, where shareholders wish to provide incentives to a risk-averse and effort-averse CEO to induce him to work hard. This model fails to rationalize observed compensation practice as Hall and Murphy (2002) and Dittmann and Maug (2007) demonstrate, because it cannot explain convex contracts. In this paper, we augment the standard model by assuming that shareholders take into account not only effort incentives but also risk-taking incentives when designing the compensation contract. We show that the augmented model predicts a contract that is flat for poor performance, convex for medium performance, and concave for high performance. We calibrate the optimal contract shape to the data and find that the augmented model approximates observed contracts much better than the model without risk-taking incentives.

The notion that shareholders might want to provide risk-taking incentives in addition to effort incentives goes back at least to Smith and Stulz (1985) and Haugen and Senbet (1981). CEOs not only exert effort and thereby shift the stock price distribution to the right, but they also make decisions that affect firm value and firm risk (i.e. location and dispersion of the stock price distribution). Accordingly, there is ample empirical evidence that risk-taking incentives matter for CEOs' actual risk-taking. Low (2009), for instance, investigates an exogenous increase in takeover protection. In a differences-in-differences analysis, she finds that those firms with little CEO risk-taking incentives experienced a sharp decline in firm risk and firm value.[1] Hirshleifer and Suh (1992) and Feltham and Wu (2001) show that optimal contracts are convex if they are designed to also provide risk-taking incentives. Therefore, the obvious way to fix the standard model and to introduce convexity into the optimal contract is the inclusion of risk-taking considerations.[2]

---

[1] Tufano (1996) and Knopf, Nam and Thornton (2002) show that CEOs respond to risk-taking incentives for hedging decisions, Rajgopal and Shevlin (2002) for investment decisions, Coles, Daniel and Naveen (2006) and Tchistyi, Yermack and Yun (2007) for capital structure decisions, and May (1995), Smith and Swan (2007), and Acharya, Amihud and Litov (2008) for corporate acquisitions. DeFusco, Johnson and Zorn (1990) and Billett, Mauer and Zhang (2010) investigate the reaction of stock and bond prices to first time equity grants and find that investors expect that these grants affect firm risk.

[2] There are several alternative explanations for the convexity in CEO contracts. Oyer (2004) models options as a device to retain employees when recontracting is expensive. Inderst and Müller (2005) explain options as instruments that provide outside shareholders with better liquidation incentives. Edmans and Gabaix (2009) and Edmans et al. (2009) show that convex contracts can arise in dynamic contracting models. Peng and Röell (2009) 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 the CEO is loss-averse. Hemmer, Kim, and Verrecchia (1999) assume gamma distributed stock prices and find convex contracts, but Dittmann and Maug (2007) show that these results are not robust.

The CEO in our model not only exerts costly effort but also determines the firm's strategy, i.e. he makes decisions on issues like project choice, mergers and acquisitions, capital structure, or financial transactions. The CEO is risk-averse and holds firm equity that provides him with effort incentives. If the contract does not provide sufficient risk-taking incentives, the CEO therefore chooses a strategy that avoids risk and depresses firm value. He might, for instance, pass up a profitable but very risky project, or might hedge his firm's risk at some cost. Shareholders can mitigate this inefficiency by providing risk-taking incentives such as rewarding the manager for extreme outcomes, but they must be careful not to impair effort incentives at the same time. While high stock price realizations are an unmistakably 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 best way to provide effort and risk-taking incentives therefore is to reward good outcomes and not to punish bad outcomes, i.e. the optimal contract features a limited downside.

The optimal contract in our model differs markedly from the one in the standard model without risk-taking incentives. As marginal utility rapidly declines with CEO wealth, the standard model predicts that the CEO is punished severely for bad outcomes while he effectively receives a fixed wage for medium and good outcomes. In our model, however, firms pay a flat wage for bad outcomes and provide incentives only for medium and high outcomes. Due to decreasing marginal utility, the payout function is convex for medium and concave for high outcomes.

We calibrate both models to the data on 727 U.S. CEOs and for each generate predictions about the optimal payout function. We then compare the optimal with the observed payout function and find that our model can explain observed contracts much better than the standard model without risk-taking incentives. In particular, the average distance between observed contract and optimal contract is 8.0% for our model compared to 28.8% for the model without risk-taking incentives.

We apply our model to contracts that consist of base salary, stock, and options and establish that, according to the model, in-the-money options are preferable to the portfolio of stock and at-the-money options that we observe in practice. On average in our sample, the strike price should be 74% of the firm's stock price at issue time. Compared to the observed portfolio contract, this in-the-money option contract moves incentives from the tails to the center of the stock price distribution. Incentives in the tails have either little effect (for high payouts) or induce the CEO to avoid risk (for low payouts). In contrast, steep payouts in the center of the distribution provide both effort and risk-taking incentives. When we take into account the tax penalties that apply to in-the-money

2

options in the U.S., we obtain optimality of the observed portfolio contract for a majority of the CEOs in our sample. Therefore, the universal use of at-the-money options, that is often seen as evidence for managerial rent-extraction (see Bebchuk and Fried, 2004), is perfectly consistent with efficient contracting.[3]

Our calibration approach bridges the gap between theoretical and empirical research on executive compensation and allows us to test the quantitative (and not just the qualitative) implications of different models. Moreover, this approach contributes to the empirical literature on CEO compensation as it circumvents the endogeneity problem that shareholders simultaneously determine firm risk and managerial incentives when they design the compensation contract. We model this endogeneity and test the predictions of the model. Under the assumptions that contracting is efficient and that CEOs are effort-averse and risk-averse, our results imply that the provision of risk-taking incentives is a major objective in executive compensation practice. We can reject the hypothesis that risk-taking incentives in observed contracts are a mere by-product of effort incentives.

Another contribution to the empirical literature is a new measure of risk-taking (dis-)incentives that combines the manager's risk preferences with the shape of his compensation contract and that we call *risk-avoidance*. It measures the required profitability an additional risky project must exceed in order to be adopted by the CEO. The median risk avoidance in our sample is 1.25 for a risk-aversion parameter of 2. Hence, the median CEO will adopt a project that increases firm risk by one percentage point if and only if it increases firm value by at least 1.25%. The standard measure for risk-taking incentives in the empirical literature is the vega of the CEO's option portfolio (see, e.g., Guay, 1999) or the utility-adjusted vega (see Lambert, Larcker, and Verrecchia (1991)). We argue that risk-taking incentives not only depend on the (utility-adjusted) vega but also on the (utility-adjusted) delta. While a negative utility-adjusted vega suggests that the CEO will pass up risky, positive-NPV projects, this effect is mitigated if the CEO has high (utility-adjusted) delta as this means that he gains from taking positive-NPV actions. Consequently, our proposed measure of risk-taking incentives is related to the ratio of utility-adjusted vega over utility-adjusted delta.

There are a few theory papers that also consider both effort-aversion and risk-taking incentives in models of executive compensation.[4] To our knowledge, this paper is the first, however, to calibrate

---

[3]There is an ongoing debate in the literature on whether executive stock options do provide risk-taking incentives. Intuitively, this seems obvious as the value of an option increases with the volatility of the underlying asset (see, e.g., Haugen and Senbet (1981) or Smith and Stulz (1985)). However, Carpenter (2000), Ross (2004), and Lewellen (2006) argue 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.

[4]Lambert (1986) and Core and Qian (2002) consider discrete volatility choices, where the agent must exert effort to

such a model and to test its quantitative implications. In this way, we also contribute to recent literature on calibrations of contracting models.[5]

We attribute the convexity in observed contracts to the provision of risk-taking incentives in this paper, and we acknowledge that there are alternative explanations for the use of options in executive compensation (see Footnote 2 above). The only alternative model that can be readily calibrated to the data is Dittmann, Maug, and Spalt (2010) where the CEO is assumed to be loss averse. We also calibrate this model to our data and find that its fit is comparable to the fit of our model. In addition, we show that the loss-aversion model does not improve much when shareholders take risk-taking incentives into account. The reason is that the standard loss-aversion model already predicts convex contracts with similar risk-taking incentives as the observed contract.

Our analysis proceeds as follows. In the next section, we present our model and derive the shape of the optimal contract. Section 3 describes the construction of the dataset, and Section 4 derives and empirically analyzes our proposed measure of CEO risk-taking incentives. In Section 5, we present our calibration method and our main results. In a nutshell, we numerically search for the cheapest contract with a given shape that provides the manager with the same incentives and the same utility as the observed contract. Section 6 provides robustness checks, and Section 7 analyzes the optimal strike price in a standard option contract. Section 8 contains our analysis for the loss-aversion model, and Section 9 concludes. The appendix collects some technical material.

2   Optimal contracting with risk-taking incentives

2.1   Model

We consider two points in time. 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. The market value of the firm at time $t = 0$ (after the contract details have been disclosed) is $P_0 = E(P_T) \exp\{-r_f T\}$, where r_f is the appropriate rate of return. At some point during the

---

gather information about investment projects. Feltham and Wu (2001) and Lambert and Larcker (2004) assume that the agent's choice of effort simultaneously affects mean and 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 (2006) 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) and Sung (1995) solve models with continuous effort and volatility choice, but Hellwig (2009) assumes that the agent is risk-neutral and Sung (1995) that the principal can observe (and effectively set) volatility. In a different type of model, Manso (2007) also establishes that optimal contracts must not punish bad outcomes when risk-taking (innovation) needs to be encouraged.

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

contract period $(0, T)$, the agent makes two choices. First, he chooses effort $e \in [0, \infty)$ that results in private costs $C(e)$ to the agent and that affects the firm's expected value $E(P_T)$. Second, he chooses a strategy s that affects the firm's expected value $E(P_T)$ and the firm's stock return volatility σ. We will refer to σ interchangeably as 'firm risk'. We can therefore write $E(P_T) = P_0(e, s) \exp\{r_f T\}$ and $\sigma = \sigma(s)$.[6]

Our model is in the spirit of Holmström (1979). The agent can costlessly destroy output or inflate volatility σ, and the principal cannot observe the agent's actions. As a consequence, the manager's wealth $W_T = w(P_T)$ only depends on the end-of-period stock price P_T, and the wage scheme $w(.)$ is non-decreasing.

We think of the strategy s as a feasible combination of many different actions that affect, among other things, project choice, mergers and acquisitions, capital structure, or financial transactions. Part of the strategy could be, for instance, an R&D project that increases value and risk. Another part could be financial hedging of some input factor which would reduce value and risk. Due to its richness, we do not model the agent's choice of strategy in detail. Instead we assume that the contract chosen by the firm does not make the CEO risk-seeking, and we show in our empirical analysis below that this assumption always holds.[7] Therefore, the CEO chooses an action that minimizes firm risk σ given expected value $E(P_T)$, or equivalently that maximizes expected value $E(P_T)$ given risk σ. Let $\widetilde{s}(e, \sigma)$ denote the strategy that maximizes expected value $E(P_T)$ given effort e and volatility σ. Then the agent's choice of effort e and strategy s is equivalent to a choice of effort e and volatility σ: $E(P_T) = P_0(e, \widetilde{s}(e, \sigma)) \exp\{r_f T\} = P_0(e, \sigma) \exp\{r_f T\}$. In the remainder of this paper, we therefore work with the reduced form of our model where the agent chooses effort e and volatility σ.

We assume that there is a first-best firm strategy $s^*(e)$ that maximizes firm value (given effort e). Let $\sigma^*(e) := \sigma(s^*(e))$ denote the (minimum) firm risk that is associated with this strategy. If the agent wants to reduce risk to some value below $\sigma^*(e)$, 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

---

[6]In our model, effort only affects expected value but not firm risk whereas strategy affects both value and risk. Other models (e.g. Feltham and Wu, 2001) assume that the agent only chooses effort and that effort affects value *and* risk. The main difference between Feltham and Wu (2001) and our model in this respect is that our model allows the CEO to affect value and risk independently of each other.

[7]More formally, we assume that the CEO's expected utility declines when volatility σ increases. This assumption is intuitive: A risk-averse CEO whose wealth is linked to firm-value is averse to an increase in firm risk σ. Providing risk-taking incentives by making the contract more convex (while keeping effort incentives and the CEO's utility constant) is costly. Therefore, firms will never increase risk-taking incentives beyond the optimal point where the CEO is indifferent to firm risk.

While this assumption is intuitive, we cannot show it formally. The reason is that the costs of an increase in risk-taking incentives given that effort incentives and utility are held constant cannot be written in closed-form. However, our empirical results below are consistent with this assumption. In particular, we find that risk-taking incentives are always costly.

reduces risk but also profits (e.g., costly hedging). In both cases, a reduction in volatility σ leads to a reduction in firm value $E(P_T)$. We therefore assume that $P_0(e, \sigma)$ is increasing and concave in σ as long as $\sigma < \sigma^*(e)$. In the region above $\sigma^*(e)$, firm value $P_0(e, \sigma)$ is weakly decreasing in σ; it is flat if the agent can take additional risk at no costs (e.g., with financial transactions). Finally, we assume that the stock price $P_0(e, \sigma)$ is increasing and concave in e (given volatility σ).

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

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

Here, r_f is the risk-free rate, and $P_0(e, \sigma) = E(P_T(u|e, \sigma)) \exp\{-r_f T\}$ is the expected present value of the end-of-period stock price P_T.[8]

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

$$U(W_T, e) = V(W_T) - C(e) = \frac{W_T^{1-\gamma}}{1 - \gamma} - C(e). \tag{2}$$

If $\gamma = 1$, we define $V(W_T) = \ln(W_T)$. 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 $\underline{U}$.

---

[8]We follow Dittmann and Maug (2007) and assume that either there is no premium for systematic risk or that 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. 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 to earn the risk-premium. Our assumption effectively means that all risk in the model is firm-specific.

2.2 Optimal contract

In order to implement a given effort e^* and level of volatility σ^*, shareholders solve the following optimization problem:

$$\min_{W_T} E\left[W_T(P_T)|e^*, \sigma^*\right] \tag{3}$$

$$\text{subject to } \frac{dW_T(P_T)}{dP_T} \geq 0 \text{ for all } P_T \tag{4}$$

$$E\left[V(W_T(P_T))|e^*, \sigma^*\right] - C(e^*) \geq \underline{U} \tag{5}$$

$$\{e^*, \sigma^*\} \in \text{argmax } \{E\left[V(W_T(P_T))|e, \sigma\right] - C(e)\} \tag{6}$$

Hence, shareholders choose the wage schedule $W_T(P_T)$ that minimizes contracting costs subject to three constraints: The monotonicity constraint (4), the participation constraint (5), and the incentive compatibility constraint (6). We replace (6) with its first-order conditions

$$\frac{dE\left[V(W_T(P_T))|e, \sigma\right]}{de} - \frac{dC}{de} = 0, \tag{7}$$

$$\frac{dE\left[V(W_T(P_T))|e, \sigma\right]}{d\sigma} = 0. \tag{8}$$

We discuss the validity of the first-order approach (i.e. that (6) can indeed be replaced by (7) and (8)) in detail in Appendix A. We call condition (7) the effort incentive constraint and (8) the volatility incentive constraint.

Proposition 1. (Optimal contract): *The optimal contract that solves the shareholders' problem (3), (4), (5), (7), and (8) has the following functional form:*

$$\frac{dV(W_T^*)}{dW_T} = \begin{cases} c_0 + c_1 \ln P_T + c_2 (\ln P_T)^2 & \text{if } ln(P_T) > -\frac{c_1}{2c_2} \\ c_0 - \frac{c_1^2}{4c_2} & \text{if } ln(P_T) \leq -\frac{c_1}{2c_2} \end{cases} \tag{9}$$

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

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

The proof of Proposition 1 and full expressions for the parameters c_0, c_1, and c_2 can be found in Appendix B. To develop an intuition for the optimal contract (10) it is instructive to look first at the

optimal contract without risk-taking incentives. This contract has the form $W_T^\gamma = c_0 + c_1 \ln P_T$ and is globally concave as long as $\gamma \geq 1$ (see Dittmann and Maug, 2007). The comparison shows that risk-taking incentives are provided by the additional quadratic term $c_2(\ln P_T)^2$. 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 (10) is flat below some threshold $\widetilde{P} = \exp\{-\frac{c_1}{2c_2}\}$, convex and increasing for some region above this threshold, and finally concave, because the concavity of the logarithm dominates the convexity of the quadratic term asymptotically.

3   Data set

We use the ExecuComp database to construct approximate CEO contracts at the beginning of the 2006 fiscal year.[9] We first identify all persons in the database who were CEO during the full year 2006 and executive of the same company in 2005. We calculate the base salary ϕ (which is the sum of salary, bonus, and "other compensation" from ExecuComp) from 2006 data, and take information on stock and option holdings from the end of the 2005 fiscal year. We subsume bonus payments under base salary, because previous research has shown that bonus payments are only weakly related to firm performance (see Hall and Liebman, 1998).[10]

We estimate each CEO's option portfolio with the method proposed by Core and Guay (2002) and then 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.[11] In this step, we lose five CEOs for whom we cannot numerically solve this system of

---

[9] We do not perform our analysis for a more recent year for two reasons. First, we cannot construct our sample consistently for 2007, because there was a significant change in the reporting standard in 2006; some firms reported according to the new standard while other firms still used the old standard. Second, we did not choose 2008 or 2009 to avoid using data from the financial crisis.

[10] We do not take into account pension benefits, because they are difficult to compile and because there is no role for pensions in a one-period model. Pensions can be regarded as negative risk-taking incentives (see Sundaram and Yermack, 2007, and Edmans and Liu, 2010), so that we overestimate risk-taking incentives in observed contracts.

[11] Appendix F contains more details about this algorithm. We take into account the fact that most CEOs exercise their

three equations in three unknowns.

We take the firm's market capitalization P_0 from the end of the 2005 fiscal year. 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 2005. We estimate the firm's stock return volatility σ from daily CRSP stock returns over the fiscal year 2006 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 2006.

We estimate 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).[12] To arrive at meaningful wealth estimates, we discard all CEOs who do not have a history of at least five years (from 2001 to 2005) on ExecuComp. During this period, they need not be CEO. This procedure results in a data set with 727 CEOs.

[Insert Table 1 here]

Table 1 provides an overview of our data set. The median CEO owns 0.32% of the stock of his company and has options on an additional 0.92% of the company's stock. The median base salary is \$1.04m, and the median non-firm wealth is \$12.0m. The representative option has a median maturity of 4.4 years and is well in the money with a moneyness (K/P_0) of 72%. 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 grant is in the money for 90% of the CEOs in our sample. In the interest of readability, we call an option with a strike price K that is close to the observed strike price K^d an "at-the-money option." Consequently, we call an option grant "in-the-money" only if its strike price K is lower than the observed strike price K^d.

stock options before maturity by multiplying the maturities of the individual option grants by 0.7 before calculating the representative option (see Huddart and Lang, 1996, and Carpenter, 1998). In these calculations, we use the stock return volatility from ExecuComp and, for the risk-free rate, the U.S. government bond yield with 5-year maturity from January 2006. 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.

[12]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).

We require that all CEOs in our data set are included in the ExecuComp database for the years 2001 to 2006, and this requirement is likely to bias our data set towards surviving CEOs, namely those who are older and richer and who work in bigger and more successful firms. Table 1 Panel B describes the full ExecuComp universe of CEOs in 2006. Compared to this larger sample, our CEOs are, on average, one year older and work in bigger firms (+$450m) with better past performance (1.3% higher return during the past five years). In a robustness check below, we analyze in how far this selection bias affects our results.

The only parameter in our model that we cannot estimate from the data is the manager's coefficient of relative risk aversion γ. We use $\gamma = 3$ in most of our analysis and provide robustness checks for $\gamma = 0.5$ and $\gamma = 6$. This range includes the risk-aversion parameters used in previous research.[13]

4   Measuring Risk-taking Incentives

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.[14] An exception are 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. Depending on whether or not the CEO has too little or too much incentives to take risk, a rise in volatility respectively increases or decreases firm value and, due to the CEO's equity portfolio, 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 (8). This constraint can be rewritten as

$$E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}\frac{dP_T}{d\sigma}\bigg| e, \sigma\right] = 0 \tag{11}$$

---

[13]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.

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

Substituting in the derivative of the stock price P_T with respect to volatility σ from (1) yields

$$\Leftrightarrow E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}\left(\frac{dP_0}{d\sigma}\frac{P_T}{P_0} + P_T\left(-\sigma T + u\sqrt{T}\right)\right)\bigg| e, \sigma\right] = 0. \tag{12}$$

As $dP_0/d\sigma$ is not random, we can rearrange (12) as

$$PPS^{ua}\frac{dP_0}{d\sigma} = -\nu^{ua}, \tag{13}$$

$$\text{where } PPS^{ua} := E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_0}\bigg| e, \sigma\right] = E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}\frac{P_T}{P_0}\bigg| e, \sigma\right] \tag{14}$$

$$\text{and} \qquad \nu^{ua} := E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}P_T\left(-\sigma T + u\sqrt{T}\right)\bigg| e, \sigma\right]. \tag{15}$$

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 firm value P_0 stays constant.

The first order condition (13) equates marginal benefits to marginal costs of an increase in volatility from the agent's point of view. The benefits stem from an increase in firm value $dP_0/d\sigma$ in which the manager participates via his incentive pay PPS^{ua}. The costs are given by the decrease of the manager's utility $-\nu^{ua}$ due to higher volatility. Hence, the agent will take an action if only if its benefits exceed its cost, i.e if

$$PPS^{ua}\frac{dP_0}{d\sigma} > -\nu^{ua} \Leftrightarrow \frac{dP_0}{d\sigma}\frac{1}{P_0} > -\frac{\nu^{ua}}{PPS^{ua}}\frac{1}{P_0}. \tag{16}$$

We therefore define the incentives to avoid risk as

$$\rho := -\frac{\nu^{ua}}{PPS^{ua}}\frac{1}{P_0}. \tag{17}$$

Appendix F contains a step by step user's guide on how to numerically calculate risk avoidance ρ.

Equation (17) defines a hurdle rate: ρ is the required increase in firm value per increase in firm risk that any new project must fulfill in order to be adopted by the CEO. Consider a project that would increase firm risk by one percentage point, e.g., from 30% to 31%, and let $\rho = 2$. Then the agent takes this project only if it increases firm value by at least 2%. All positive NPV projects that generate less than 2% increase in firm value for each percent of additional risk will 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 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 firm value. If $\rho = 0$, the CEO is indifferent to firm risk and will therefore implement all profitable projects irrespective of their riskyness. We refer to ρ as incentives to avoid risk or 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 this means that he gains from taking positive-NPV actions. Second, consider the case where vega is positive, and so the manager has an incentive to take risky projects even if they are negative-NPV. Again, 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.

Table 2, Panel A displays descriptive statistics for the incentives to avoid risk ρ in the observed contract for five values of risk aversion γ. In all cases, risk avoidance ρ is positive for most CEOs; for $\gamma \geq 3$ it is positive for virtually all CEOs. The results in Table 2 are therefore consistent with our assumption that the contracts chosen by the firm do not make CEOs risk-seeking. For $\gamma = 3$, the average ρ is 1.87 and the median is 1.75. This implies that the average CEO in our sample passes up risky positive NPV projects if they increase fim value by less than 1.87% per percentage point of additional volatility. For lower values of risk aversion γ, risk-avoidance is lower. For $\gamma = 0.5$, the average and median ρ are 0.19.

[Insert Table 2 here]

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. It is difficult to judge, however, what a plausible optimal level for ρ is, because the optimal level depends on the availability of risky projects: a firm that has only few risky projects will not benefit much from an increase in risk-taking incentives. Nevertheless, a median ρ of 1.75 for $\gamma = 3$ appears large when taking into account that CEO pay typically makes up only about 1% of firm value (see the median of 'value of contract' and 'firm value' in Table 1). A potential reason is that CEOs are less risk averse (see Graham, Harvey, and Puri, 2009), so that

12

$\gamma < 3$. We still use $\gamma = 3$ as the base case in this paper because it is a conservative choice; the fit of our model to the data improves as γ decreases. Another reason why risk avoidance ρ is high in Table 2, Panel A is that major shareholders might not be well diversified and therefore want to take inefficiently low risk (see Faccio, Marchica, and Mura, 2010).

5 Empirical Results

In this section, we calibrate the optimal contract (10) to the data and evaluate how well it approximates observed contracts. We assume that shareholders want to implement a certain action $\{e^*, \sigma^*\}$ and that they have done so in the observed contract.[15] Under this assumption, we can reformulate the shareholder's optimization problem (3) to (6) as follows (see Appendix D for the derivation):

$$\min_{c_0, c_1, c_2} E\left[W_T^*(P_T | c_0, c_1, c_2)\right] \tag{18}$$

$$\text{subject to } E\left[V(W_T^*(P_T | c_0, c_1, c_2))\right] = E\left[V(W_T^d(P_T))\right] \tag{19}$$

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

$$\rho(W_T^*(P_T | c_0, c_1, c_2)) = \rho(W_T^d(P_T)), \tag{21}$$

where $W_T^d(P_T) = \phi^d + n_S^d P_T + n_O^d \max\{P_T - K^d, 0\}$ is the observed contract (d for "data") that we construct from the data as described in Section 3. Intuitively, we search for the contract $W_T(P_T | c_0, c_1, c_2)$ with shape (10) that achieves three objectives. First it provides the same effort and risk-taking incentives to the agent as the observed contract (conditions (20) and (21)). Second it provides the agent with the same utility as the observed contract (condition (19)), and third it is as cheap as possible for the firm (objective (18)).[16] 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, the observed contract is not efficient according to the model: it is possible to find a cheaper contract that implements the same effort and the same strategy as the observed contract. In this case, either compensation practice is inefficient or the model is incorrect. In both cases, the

---

[15]This calibration method has first been used by Dittmann and Maug (2007). It is the first stage of the two-stage procedure in Grossman and Hart for the effort/volatility level implemented by the observed contract. We cannot repeat this task for alternative effort/volatility levels, because this would require knowledge of the production and the cost function. Therefore we cannot analyze the optimal level of effort or volatility (i.e., the second stage in Grossman and Hart, 1983).

[16]Note that we have as many constraints in problem (18) to (21) as we have parameters, so that there are no degrees of freedom left to minimize costs. Therefore, we solve a system of three equations (19) to (21) in three unknowns for every CEO in our sample. The resulting contract has the optimal shape and therefore must be cheaper than the observed contract.
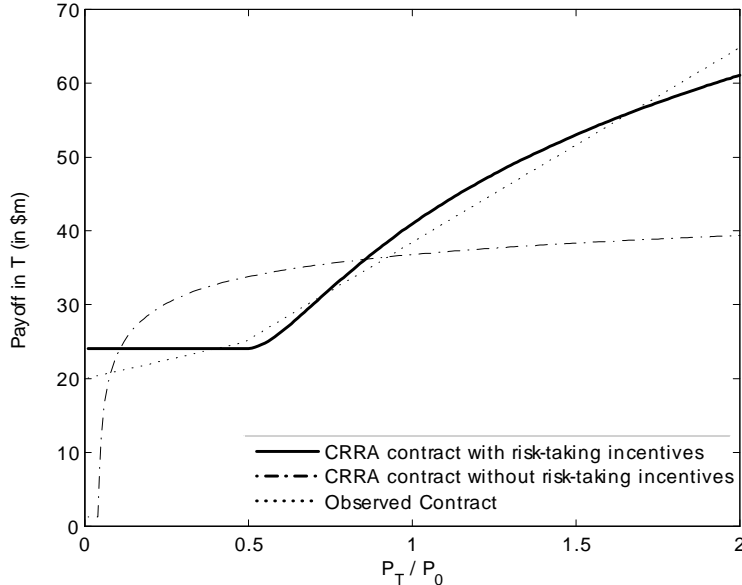
Figure 1: The figure shows end of period wealth W_T for the observed contract (dotted line), the optimal CRRA contract with risk-taking incentives (solid line), and the optimal CRRA contract without risk-taking incentives (dashed line) for a representative CEO whose parameters are close to the median of the sample. The parameters are $\phi = \$1.1$m, $n_S = 0.33\%$, $n_O = 0.57\%$ for the observed contract. Initial non-firm wealth is $W_0 = \$15.6$m. $P_0 = \$2.8$bn, $\sigma = 27.9\%$, and $K/P_0 = 49\%$, $T = 4.2$ years, $r_f = 4.4\%$, $d = 1.8\%$. All calculations are for $\gamma = 3$.

model is not descriptive of the data.

Figure 1 shows our calibration results for a representative CEO.[17] The solid line represents the optimal contract W_T^* that solves the optimization problem (18) to (21), 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 end-of-period stock price P_T which we express as a multiple of the beginning-of-period stock price P_0. The optimal contract with risk-taking incentives protects the CEO from losses. It provides the CEO with a flat wealth of \$24m if the stock price falls below 49% of the initial stock price. Intuitively, limiting the downside for bad outcomes provides better (i.e., cheaper) risk-taking incentives than rewarding good outcomes. In the region between 49% and 70% of the initial stock price, 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 the CEO is, the less interested he is in additional wealth.

---

[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 smallest.

As a benchmark, we also calibrate the optimal contract without risk-taking incentives from Dittmann and Maug (2007); it is the broken line in Figure 1. For this purpose, we solve the optimization problem (18) to (20) without the volatility incentive constraint (21) 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 1 shows that the benchmark contract is globally concave and puts the agent's entire wealth at risk. As a consequence, it makes the agent extremely averse to taking additional risk. For the full sample, Table 2, Panel B shows descriptive statistics for the incentives to avoid risk, ρ, for the benchmark contract. For $\gamma = 3$, average ρ is 9.43 compared to 1.87 in the observed contract.[18] With the benchmark contract, the agent will therefore be willing to increase firm risk by one percentage point only if the additional project increases firm value by at least 9.43%. Note that by construction the RTI contract has the same ρ as the observed contract.

The figure suggests that the model with risk-taking incentives (solid line) fits the observed contract (dotted line) much better than the model without risk-taking incentives (broken 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:

$$D_1 = E\left(\frac{|W_T^*(P_T) - W_T^d(P_T)|}{W_T^d(P_T)}\right). \tag{22}$$

We recognize that the observed contract we construct in Section 3 is a stark simplification of the contracts used in practice, especially because typical contracts contain several grants of stock options with different maturities and different strike prices. Contracts are therefore 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

$$D_2 = E\left(\frac{|W_T^*(P_T) - W_T^{smth}(P_T)|}{W_T^{smth}(P_T)}\right), \tag{23}$$

where $W_T^{smth}(P_T)$ sums up the expected value of all option grants held by the CEO. For a grant that has a maturity larger than T, this is just the Black Scholes value for the remaining maturity

---

[18] For 94% of all CEO-γ combinations, risk-avoidance ρ is higher in the RTI contract than in the observed contract (not shown in the table). The remaining 6% mostly occur for $\gamma = 6$ and are very likely due to to numerical problems, because the benchmark contract is much steeper for small values of P_T for $\gamma = 6$ than it is for $\gamma = 3$. When the contract approaches zero, differences between very small and very large numbers occur in the numerical routines that cannot be handled well numerically. This is also the reason why the 90% quantile of ρ is lower for $\gamma = 6$ than for $\gamma = 0.5$ or $\gamma = 3$. We therefore conclude that risk-taking incentives are always costly in our model and that firms will never choose a contract that makes the CEO risk-seeking.

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 E. For the representative CEO shown in Figure 1, the distance is 5.2% for the contract with risk-taking incentives and 22.2% for the contract without risk-taking incentives. The representative CEO has only one option grant, so both distance measures have the same value in this case.

[Insert Table 3 here]

Table 3, Panel A shows the results for all CEOs in our sample. The left part of the table describes the optimal contract with risk-taking incentives 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 mean and median of a few key variables that describe the contract. These variables include the two distance measures D_1 and D_2 from (22) and (23) 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 $\Pr(P_T \leq -\frac{c_1}{2c_2})$ from equation (10). Second, the inflection quantile is the probability mass below the point where the contract curvature changes from convex to concave.

Table 3 demonstrates that the optimal contract provides the agent with comprehensive downside protection. For $\gamma = 3$, the median minimum wealth is 1.4 times the initial wealth W_0. Only for 0.1% of the CEOs in our sample is the minimum wealth lower than their observed non-firm wealth W_0. The contract pays out the minimum wage for the worst outcomes with a median probability of 16.1%. The median inflection quantile is 32.5%, so that the contract is convex for mediocre performance between the 16.1% quantile and the 32.5% quantile and concave for good performance above the 32.5% quantile.

Table 3, Panel A also shows the savings firms could realize when they switch from the observed contract to the optimal 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 $\gamma = 3$, mean (median) savings are 10.4% (6.9%). The mean distance D_1 between observed

contract and optimal contract is 8.0%, and the mean distance D_2 is 8.6%. For lower values of risk aversion γ, we obtain a better fit: For $\gamma = 0.5$, the average distance D_1 is only 2.5%. Contracts are then convex over a larger range of stock prices from the 1.7% quantile to the 77.7% quantile for the median CEO. Savings from recontracting are smaller for lower values of risk aversion γ, because savings are generated by efficient risk sharing which is less important if the CEO is less risk averse. 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.

The right part of Table 3 displays the results for the benchmark model without risk-taking incentives. This contract does not contain any downside protection, so the CEO can potentially lose all her wealth. Moreover, it is globally concave for all CEOs if $\gamma > 1$, so that the kink quantile and the inflection quantile are both zero. Due to convergence problems, the sample for the two sets of results in Table 3, Panel A is not the same. We therefore report the numbers again 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 28.3% for the benchmark model compared to 8.0% 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 34.5% of contracting costs on the table while the RTI model puts this number at 10.4% only. These numbers suggest that risk-taking incentives play an important role in observed compensation contracts. Observed contracts appear markedly less inefficient when risk-taking incentives are taken into account.

6 Robustness checks

6.1 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 3 with constant absolute risk aversion (CARA), so that $V^{CARA}(W_T) = -\exp(-\eta W_T)$ replaces $V(W_T)$ in equation (2). Taking the first derivative and plugging the result into equation (9) from Proposition 1 yields the following corollary:

Corollary 1. *(Optimal CARA contract): If the agent exhibits constant absolute risk aversion with*

parameter η, the optimal contract has the following functional form:

$$W_T^* = \begin{cases} \frac{1}{\eta} \log \left\{ \eta \left[c_0 + c_1 \ln P_T + c_2 (\ln P_T)^2 \right] \right\} & \text{if } ln(P_T) > -\frac{c_1}{2c_2} \\ \frac{1}{\eta} \log \left\{ \eta \left[c_0 - \frac{c_1^2}{4c_2} \right] \right\} & \text{if } ln(P_T) \leq -\frac{c_1}{2c_2} \end{cases} \tag{24}$$

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. Results are shown in Table 4.

[Insert Table 4 here.]

Table 4 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, it guarantees a minimum payout that is always higher than the CEO's nonfirm wealth, and it is convex for intermediate payouts and concave for good payouts.

6.2 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 demonstrates that younger and less successful CEOs are underrepresented in our data set. We therefore divide our sample in quintiles according to four variables: CEOs' non-firm wealth W_0, CEO age, firm value P_0, and the past five years' stock return. 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.]

The table shows that the model fit is worse for younger and less wealthy CEOs. For the 20% youngest and the 20% least wealthy CEOs, we find an average distance of 11.7% and, respectively, 11.4% compared to 8.0% for the full sample (see Table 3). Given that our sample is biased towards older and more wealthy CEOs, the average distance in the unbiased sample would be somewhat higher than shown in Table 3. We find the opposite effect, however, for past performance: the 20% best-performing firms have an average distance of 10.6%. As we oversample firms with good performance, the average distance in Table 3 should be adjusted downwards. Altogether, the effect of the sample bias on our results is therefore small.

18

7 Application: Optimal strike prices

Figure 1 suggests that the model with risk-taking incentives can explain option contracts much better than the benchmark model without risk-taking incentives. In this section, we therefore analyze the RTI model's implications for optimal strike prices in a standard option contract. We consider contracts that consist of fixed salary ϕ, the number of stock n_S, and the number of options n_O with strike price K:

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

This contract has the same structure as the stylized observed contract that we construct in Section 3 above. For each CEO, we solve the optimization problem (18) to (21) with W_T^{lin} instead of W_T^*, where the principal's choice variables are ϕ, n_S, n_O, and K.[19]

[Insert Table 6 here]

Table 6 describes our results for three values of γ: 0.5, 3, and 6. In all cases, the RTI model predicts that the median CEO does not hold any stock.[20] Instead, the median CEO would have more options (+43% for $\gamma = 3$; compare Table 6 with Table 1) and more base salary (+211%). For 99.2% of the CEOs in our sample, the strike price in the optimal contract is lower than in the observed contract. While the moneyness of the observed contract is 70.1% on average from Table 1, it is 51.7% for the optimal 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 prior to 2006, our results imply that options should have been issued 73.8% (= 51.7%/70.1%) in the money.

---

[19] 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 argue that a good model should not *assume* but rather *generate* positive option holdings and positive salaries. 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 (in most cases we find $n_S^* = 0$, as we discuss shortly), we repeat our calibration with the additional restriction $n_S = 0$ whenever we obtain a solution with $n_S^* > 0$ in the original problem. In almost all cases, the contract with $n_S = 0$ is slightly cheaper than the initially found contract with $n_S^* > 0$. This shows that interior solutions with $n_S^* > 0$ are a numerical artifact. For our empirical analysis we always use the solution with the lowest costs.

[20] The small positive average stock holdings are due to a few CEOs who either don't have any options in the observed contract (i.e., $n_O^d = 0$) or for whom our routine hits the boundaries for the strike price K (see footnote 19).
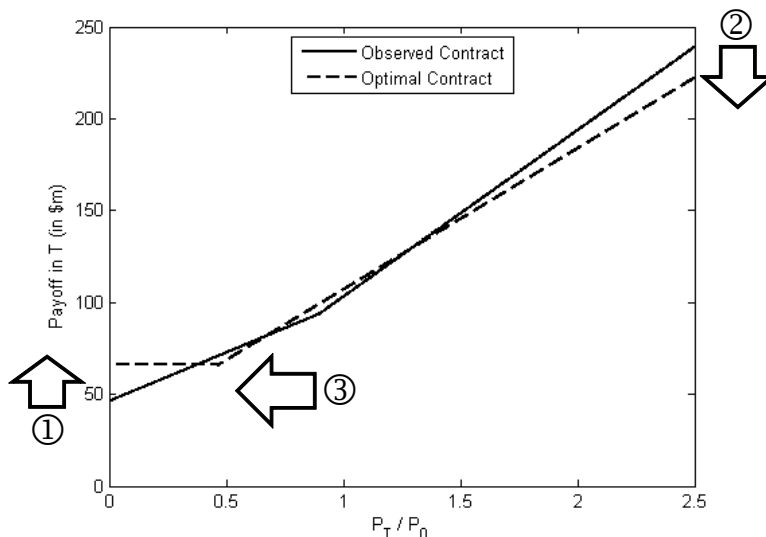
Figure 2: The figure shows end-of-period wealth W_T as a function of end-of-period stock price P_T for the observed contract (solid line) and the optimal piecewise linear contract (dashed line) for one CEO in our sample. The arrows indicate the three main features of the optimal contract relative to 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 $\phi = \$6.3$m, $n_S = 5.97\%$, $n_O = 4.44\%$ for the observed contract. Initial non-firm wealth is $W_0 = \$32.1$m. $P_0 = \$853$m, $\sigma = 25.7\%$, and $K/P_0 = 90\%$, $T = 4.4$ years, $r_f = 4.4\%$, $d = 0.9\%$, and $\gamma = 3$.

The general picture is that the stock and option holdings in the observed contract are replaced by option holdings that are considerably deeper in the money. As options are less valuable than shares, this exchange is accompanied by an increase in 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 optimal contract are limited, however. The median firm would just save 1.9% of contracting costs for $\gamma = 3$ and the average is 4.1%. 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 is 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 (see Footnote 19), it is not a *problem* of our model but rather a *result*.

Figure 2 illustrates our main results. It shows the payout function $W_T^{lin}(P_T)$ of the observed contract and the optimal contract for one CEO in our sample. This CEO is not representative for our sample; for a typical CEO the two contracts are more difficult to distinguish visually. The three arrows in Figure 2 indicate the main features of the optimal contract and help to develop an intuition

for our main result that in-the-money options are a cheaper way to provide incentives than a portfolio of stock and at-the-money options. The first feature of the optimal contract is that it provides for less punishment in the bad states of the world than the observed contract, which improves risk-taking incentives. On the other hand, the optimal contract also gives fewer rewards in the best states of the world (feature 2), which reduces risk-taking incentives. These two effects offset each other, so that the optimal contract provides the same risk-taking incentives as the observed contract. 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 effort incentives and offsets the effect of features (1) and (2). Therefore, the optimal contract also generates the same effort incentives as the observed contract; it merely moves some of the effort incentives from the tails of the distribution to its center. Finally observe that features (1) and (2) make the optimal contract less risky than the observed contract. Therefore the agent demands a lower risk-premium and the optimal contract is cheaper for shareholders. The same effects can be found for the general optimal contract depicted in Figure 1. The RTI-contract provides for less punishments in the bad states of the world and for less rewards in the very best states compared to the observed contract. At the same time, the slope of the RTI contract is steeper in the center of the stock price distribution, i.e. incentives are moved from the tails of the distribution to its center.

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 analyze how the optimal option contract looks like when realistic taxes are taken into account.[21]

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 firm level. However, this rule can be easily circumvented by subjecting in-the-money options to specific performance criteria. We therefore concentrate on the 20% penalty tax from Section 409A and neglect the potential effects of Section 162(m) in the following analysis.[22]

---

[21] Another potential reason why we do not see in-the-money options in the U.S. are the U.S. accounting rules. In-the-money options always had to be expensed while at-the-money options did not need to be expensed prior to 2006. This accounting rule possibly explains the absence of in-the-money options before 2004, the year in which Section 409A was enacted.

[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

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 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-money. We find that in this setting the observed contract turns out to be optimal for 95.7% of all CEOs for whom our algorithm converges (not shown in the tables).[23]

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).[24] A potential reason is that in-the-money options cause some costs that are not included in our model and that justify government intervention. Our results in Table 6 show that the use of in-the-money options is associated with large increases in base salary. These might be difficult to explain to shareholders and the general public, and might cause social unrest and higher wage demands. Moreover, there might be concerns that executives try to influence the strike price of the option grants just as some appear to have done in the recent backdating scandal. A commitment to using only at-the-money options would reduce this rent-seeking activity, and our analysis shows that the costs of such a commitment are low.

8 Optimal contracts when CEOs are loss averse

Our analysis in Section 5 shows that the RTI model can explain observed contracts reasonably well and certainly much better than the benchmark model without risk-taking incentives. 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 reasonably 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.

---

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.

[23] See Hall and Murphy (2000) for an alternative justification of at-the-money strike prices.

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

8.1 The standard loss-aversion model

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

$$V^{LA}(W_T) = \begin{cases} (W_T - W^R)^\alpha & if \quad W_T \geq W^R \\ -\lambda (W^R - W_T)^\beta & if \quad W_T < W^R \end{cases}, \text{ where } 0 < \alpha, \beta < 1 \text{ and } \lambda \geq 1. \quad (25)$$

Here, W^R is the agent's reference wealth level. Payouts above this level are coded as gains, while payouts below are coded as losses. The agent is risk-averse over gains and risk-seeking over losses, and losses receive a higher weight ($\lambda > 1$) than gains. The utility $U^{LA}(W_T, e) = V^{LA}(W_T) - C(e)$ then replaces equation (2). Following Dittmann, Maug, and Spalt (2010), we use $\alpha = \beta = 0.88$ and $\lambda = 2.25$ and parameterize reference wealth W^R by

$$W_{2006}^R = W_0 + \phi_{2005} + \theta \cdot MV(n_{2005}^S, n_{2005}^O, P_{2006}),$$

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 nonfirm 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.

Figure 3 shows the LA contract for $\theta = 0.1$ together with the CRRA-RTI contract for $\gamma = 3$ and the observed contract for the representative CEO. Visual inspection shows that both models fit the observed contract reasonably well. However, there are two important differences: First, while the LA contract is convex over all realistic stock price outcomes, the CRRA-RTI contract is concave for medium and large stock prices. Second, the LA contract features a discontinous jump for very low stock prices from a payout just above the reference point to the lowest possible payout of zero. As a consequence, the LA model approximates the observed contract poorly for very small stock prices, but seems to do a better job than the CRRA-RTI model for high stock prices.

[Insert Table 7 here]

Table 7 displays our results for the LA model for three different values of reference wealth as parameterized by θ. In addition to mean and 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") and the inflection quantile where the contract changes from convex to concave. We find that the

23
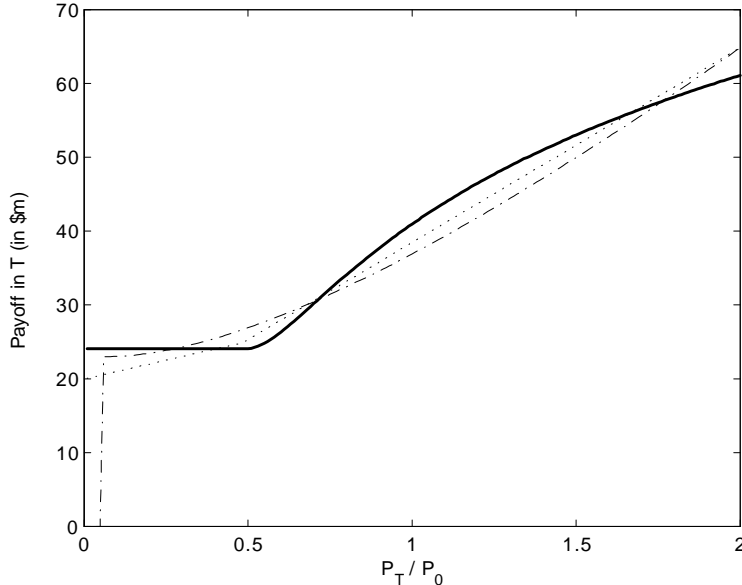
Figure 3: The figure shows end-of-period wealth W_T of three different contracts for the same repre-
sentative CEO as Figure 1. The dotted line shows the observed contract; the solid line displays the
optimal CRRA contract with risk-taking incentives for $\gamma = 3$; and the dashed line shows the optimal
LA contract for $\theta = 0.1$.

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 4.3% for the LA model compared with 6.9% for the CRRA-RTI
model (see Table 3).[25] For higher reference 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 wealth.
For $\theta = 0.5$, the average jump quantile is 3.47% and for $\theta = 0.9$ it is 9.36%. We therefore 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.

8.2   Risk-taking incentives in the loss-aversion model

CEO preferences are different in the loss-aversion model compared to the CRRA model. Hence,
risk-taking incentives differ between the two models. Table 8, Panel A displays descriptive statistics
of risk avoidance ρ in the LA model for the observed contract. A comparison with Table 2, Panel

---

[25]Across all models and all specifications, the CRRA-RTI model with $\gamma = 0.5$ has the best fit. However, we do not
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 shows that risk avoidance in the observed contract is considerably lower if the CEO is loss-averse than if he exhibits constant relative risk aversion. In the LA-model with $\theta = 0.1$, mean and median ρ are both close to zero, and 48.7% of the CEOs have negative ρ, i.e. incentives to take on too much risk. For larger values of θ, ρ increases somewhat but is always much lower than the average 1.87 we find for the CRRA-model with $\gamma = 3$.

[Insert Table 8 here.]

Table 8, Panel B shows similar statistics for ρ in the LA contract. Risk-taking incentives do not differ much between observed contracts and optimal contracts in the LA model. On average, ρ decreases somewhat for $\theta = 0.1$ and $\theta = 0.5$, and increases slightly for $\theta = 0.9$. This is in stark contrast to the CRRA model, where the optimal contract generates severely higher ρ compared to the observed contract (see Table 2). The reason is that the cost effective way to provide effort incentives in the CRRA-model is to punish the agent for very low outcomes, and this policy severely increases risk avoidance. In the LA model, on the other hand, cost effective effort incentives consist not only of sticks but also of carrots in the form of convex payouts for medium and high outcomes. While the sticks reduce effort incentives, the carrots increase them, and the overall effect can go in both directions. As a consequence, our assumption that the contract chosen by the firm does not make the CEO risk-seeking does not hold in general for the LA model.

To analyze risk-taking incentives in the loss-aversion model in more detail, we distinguish six cases, depending on whether or not risk-avoidance is higher in the LA model than in the observed contract and on whether one or both of the risk-avoidance measures are positive. Table 8, Panel C defines these six cases and displays how often each of them applies for the three different values of θ. There are only two cases (cases 1 and 4) where risk-taking incentives are unambiguously worse in the LA model than in the observed contract, so that augmenting the model with risk-taking incentives might improve its fit. In cases 2 and 5, risk-taking incentives are better (i.e. ρ is closer to zero) in the LA model than in the observed contract, so there is no room for improvements.

[Insert Table 9 here.]

The only case that is consistent with our assumptions is case 1. Note that for the CRRA model with $\gamma = 3$, 99.3% of all CEOs fall into this category (see Table 2). For this case, we derive the shape of the optimal LA-RTI contract in Appendix C and then calibrate it to the observed contract for those CEOs where case 1 applies. The results are shown in Table 9 which is structured similarly to

Table 3. 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 6.7% to 3.1% on average. Removing the punishment for poor outcomes increases 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 compared to the LA model. In many cases, the optimal LA-RTI contract has a poor fit, because it is flat at the reference wealth for small and intermediate payouts and takes off with strong convexity only for high payouts. Altogether we therefore conclude that the LA-RTI model does not yield any significant improvements over the LA model. We conclude that risk-taking incentives are less of an issue if managers are loss-averse, because the LA model does not reduce risk-taking incentives nearly as much as the CRRA model.

9   Conclusions

In this paper we analyze a principal-agent model in which the agent not only exerts effort but also determines the firm's strategy and thereby its stock return volatility. In this model, the choice of a more risky firm strategy has two effects on the manager's compensation. The first, obvious effect is that higher volatility makes future payoffs more risky, so that the utility a risk-averse manager derives from restricted stock drops. This effect has already been analyzed extensively in the literature (see Lambert, Larcker and Verrecchia, 1991; Guay, 1999; Carpenter, 2000; Ross, 2004). The second effect that has so far been neglected by the empirical literature is that a more risky firm strategy also affects expected firm value. In a situation where the firm takes inefficiently low risk, risk-taking increases firm value and therefore, via the CEO's equity portfolio, CEO wealth. While this is the relevant situation in equilibrium when the CEO is risk-averse, there is another case that might apply out of equilibrium or for alternative preference specifications, like loss-aversion. Then the firm takes inefficiently high risk and risk-taking reduces firm value and CEO wealth. Therefore, it is not enough to just look at the direct impact of an increase in risk on a manager's compensation package (vega) in order to determine his attitude towards an increase in risk. The indirect effect via a change in firm value and the manager's equity portfolio (delta) must also be taken into account. Our paper provides - to the best of our knowledge - the first empirical analysis of a full principal agent model that takes both effects into account. We also propose a new measure of risk-taking incentives that combines the CEO's preferences and the curvature of the contract and predicts which risky projects the CEO will adopt.

Our model predicts an optimal contract that has a limited downside and a steep slope for inter-

mediate outcomes. It is flat for low performance, increasing and convex for intermediate performance, and increasing and concave for high performance. The optimal contract is therefore reminiscent of a standard bonus scheme that is capped from below as well as from above (see Murphy, 2001, and Healy, 1985). Our calibration results show that the model contract approximates the observed contract well. Across all CEOs, the average distance between the two contracts is 8.0% for a CRRA parameter of 3. In contrast, a model that does not take into account risk-taking incentives differs from the observed contract by 28.8%.

We also calibrate the loss-aversion (LA) model from Dittmann, Maug, and Spalt (2010) to our data and find an average distance of 5.8% for a low reference point. For higher reference points, however, the model is considerably worse than the risk-aversion model with risk-taking incentives (CRRA-RTI). Altogether, it is therefore unclear which model is more successful. The main difference between the two models is that the LA model predicts a discontinuous jump to the lowest possible payout for poor performance while the CRRA–RTI model predicts a flat payout. On the other hand, the LA model is convex over all realistic outcomes whereas the CRRA–RTI model becomes concave for high outcomes. Note that observed contracts are linear for high outcomes, so both models necessarily have an approximation error. We also show that the fit of the LA model does not improve much (and sometimes even gets worse) when risk-taking incentives are taken into account. While risk-taking incentives are neccessary to explain observed contracts in the risk-aversion model, they are not needed in the loss-aversion model.

A limitation of our analysis is that our model is static and considers only two points in time: the time of contract negotiation and the time when the final stock price is realized. Realistically, a bad or unlucky CEO is likely to be replaced if the stock price drops by more than 50%.[26] Such a dismissal has two consequences. First it might affect firm performance if the new CEO is more skilled than the ousted CEO. This effect is beyond the scope of our model, as at least two periods are necessary to describe it. Second, dismissals negatively affect the payout of the ousted CEO, mainly because it reduces the CEO's future employment opportunities. Our model predicts a flat pay for low levels of stock price, so this negative effect of a dismissal is undesirable. Consequently, our analysis can also be interpreted as a justification of severance pay that compensates the manager for his loss in human capital (see Yermack, 2006).

---

[26]Coughlan and Schmidt (1985), Kaplan (1994), and Jenter and Kanaan (2010), among others, analyze the sensitivity of dismissals to past stock price performance.

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 (6) by the two first-order conditions (7) and (8). This approach is only valid if the utility which the agent maximizes has exactly one optimum, and 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) and especially the discussion of their Figure 1). 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 under the optimal contract $W^*(P_T)$ to shirk, 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 are not likely to 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 (10) 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 will not exert any effort and will choose the lowest feasible volatility. Consequently, the first-order approach is violated. Our first condition therefore states that this is not the case. This assumption appears reasonable, because for the median CEO the minimum payout ($1.4m, from Table 3, Panel A for $\gamma = 3$) is only 5.6% of the expected payout ($25.0m, 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 5.6% 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 obtains 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 (4) must hold for every P_T, so that it is actually a continuum of infinitely many 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, (4) can be rewritten as

$$h'(h^{-1}(W_T)) \geq 0. \tag{26}$$

For every W_T, (4) provides one restriction, so the Lagrangian for the differentiation at W_T is:

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

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

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

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

$$g(P_T|e,\sigma) = \frac{1}{P_T \sqrt{2\pi\sigma^2 T}} \exp[-\frac{(\ln P_T - \mu(e,\sigma))^2}{2\sigma^2 T}] \tag{27}$$

with

$$\mu(e,\sigma) = \ln P_0(e,\sigma) + (r_f - \sigma^2/2)T. \tag{28}$$

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

$$g(P_T|e,\sigma) = \lambda_{PC} V_{W_T} g(P_T|e,\sigma) + \lambda_e V_{W_T} g_e(P_T|e,\sigma) + \lambda_\sigma V_{W_T} g_\sigma(P_T|e,\sigma) \tag{29}$$

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

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 (26) is binding, equation (29) defines the Lagrange multiplier λ_{W_T}, and the solution is determined by the binding monotonicity constraint. If (26) is not binding, λ_{W_T} is zero and the first-order condition (29) simplifies with some rearranging to

$$\frac{1}{V_{W_T}(W_T)} = \lambda_{PC} + \lambda_e \frac{g_e}{g} + \lambda_\sigma \frac{g_\sigma}{g}. \tag{30}$$

Consequently, the solution is given by (30) as long as it is monotonically increasing, and flat otherwise.

For the log-normal distribution (27) we get:

$$g_e = g \cdot \frac{\ln P_T - \mu(e,\sigma)}{\sigma^2 T} \cdot \mu_e(e,\sigma) \tag{31}$$

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

$$= g \cdot \frac{[\ln P_T - \mu] \cdot \mu_\sigma \cdot \sigma + [\ln P_T - \mu]^2}{\sigma^3 T} - \frac{g}{\sigma}. \tag{32}$$

Substituting this into the first-order condition (30) yields:

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

From inspection, the optimal wage contract can be written as (9) with parameters c_0, c_1, and c_2:

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

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

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

Equation (10) then follows immediately with $V(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}$. $\blacksquare$

Appendix C: Optimal loss aversion contract

Proposition 2. (Optimal LA contract): *Under the assumptions that (i) the agent is loss-averse as described in (2) and (25) and (ii) the stock price P_T is lognormally distributed as described in (1), the optimal contract $W^*(P_T)$ that solves the shareholders' problem (3), (4), (5), (7), and (8) is:*

$$W_T^{*,LA} = \begin{cases} W^R + [\widetilde{w}(P_T)]^{\frac{1}{1-\alpha}} & \text{if } P_T > \widehat{P} \\ 0 & \text{if } P_T \leq \widehat{P} \end{cases}, \tag{33}$$

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

$$\alpha W^R = \widetilde{w}(P_T) \lambda \left(W^R \right)^\beta + (1-\alpha) \left(\widetilde{w}(P_T) \right)^{\frac{1}{1-\alpha}}. \tag{34}$$

If no solution for $\widehat{P}$ exists to (34), the optimal contract is

$$W_T^{*,LA} = \begin{cases} W^R + [\widetilde{w}(P_T)]^{\frac{1}{1-\alpha}} & \text{if } ln(P_T) > -\frac{c_1}{2c_2} \\ W^R + \left(c_0 - \frac{c_1^2}{4c_2} \right)^{\frac{1}{1-\alpha}} & \text{if } ln(P_T) \leq -\frac{c_1}{2c_2} \end{cases}. \tag{35}$$

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 $P_T < \widehat{P}$. We can therefore

rewrite the optimization problem as:

$$\min_{\widehat{P}, W_T \geq W^R} \int_{\widehat{P}}^{\infty} W_T g(P_T|e,\sigma) dP_T \tag{36}$$

$$s.t. \int_{\widehat{P}}^{\infty} V(W_T) g(P_T|e,\sigma) dP_T + V(0) G(\widehat{P}|e,\sigma) \geq \underline{U} + C(e), \tag{37}$$

$$\int_{\widehat{P}}^{\infty} V(W_T) g_e(P_T|e,\sigma) dP_T + V(0) G_e(\widehat{P}|e,\sigma) \geq C'(e), \tag{38}$$

$$\int_{\widehat{P}}^{\infty} V(W_T) g_\sigma(P_T|e,\sigma) dP_T + V(0) G_\sigma(\widehat{P}|e,\sigma) \geq 0. \tag{39}$$

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:

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

$$- \lambda_\sigma V'(W_T) g_\sigma(P_T|e,\sigma) \tag{40}$$

Setting (40) to zero and solving gives the optimal contract in the gain space as:

$$V'(W_T) = \left[\lambda_{PC} + \lambda_e \frac{g_e(P_T|e,\sigma)}{g(P_T|e,\sigma)} + \lambda_\sigma \frac{g_\sigma(P_T|e,\sigma)}{g(P_T|e,\sigma)} \right]^{-1}. \tag{41}$$

For the Tversky and Kahneman (1992) preferences (25) we can rewrite (41) as:

$$W_T = W^R + \left[\alpha \left(\lambda_{PC} + \lambda_e \frac{g_e(P_T|e,\sigma)}{g(P_T|e,\sigma)} + \lambda_\sigma \frac{g_\sigma(P_T|e,\sigma)}{g(P_T|e,\sigma)} \right) \right]^{\frac{1}{1-\alpha}}. \tag{42}$$

Substituting the relevant expressions for the lognormal distribution from (31) and (32) and rearranging yields

$$W_T = W^R + \left[c_0 + c_1 \ln P_T + c_2 (\ln P_T)^2 \right]^{\frac{1}{1-\alpha}}, \tag{43}$$

where

$$c_0 = \alpha \lambda_{PC} - \alpha \lambda_e \frac{\mu_e \cdot \mu}{\sigma^2 T} - \alpha \lambda_\sigma \left(\frac{\mu \cdot \mu_\sigma}{\sigma^2 T} - \frac{\mu^2}{\sigma^3 T} + \frac{1}{\sigma} \right), \tag{44}$$

$$c_1 = \alpha \lambda_e \frac{\mu_e}{\sigma^2 T} + \alpha \lambda_\sigma \left(\frac{\mu_\sigma}{\sigma^2 T} - \frac{2\mu}{\sigma^3 T} \right), \tag{45}$$

$$c_2 = \frac{\alpha \lambda_\sigma}{\sigma^3 T} \geq 0. \tag{46}$$

Equation (43) provides the shape of the optimal contract for $P \geq \widehat{P}$ - provided that it is monotonic.

The optimal cut-off point $\widehat{P}$. To find $\widehat{P}$ we take the derivative of the Lagrangian with respect to $\widehat{P}$:

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

$$+ \lambda_e \left(V(W(\widehat{P})) - V(0)\right) g_e(\widehat{P}|e,\sigma) + \lambda_\sigma \left(V(W(\widehat{P})) - V(0)\right) g_\sigma(\widehat{P}|e,\sigma) \tag{47}$$

$$= -\left(V(W(\widehat{P})) - V(0)\right) g(\widehat{P}|e,\sigma)\left[\frac{W(\widehat{P})}{V(W(\widehat{P})) - V(0)} - \lambda_{PC} - \lambda_e \frac{g_e\left(\widehat{P}|e,\sigma\right)}{g\left(\widehat{P}|e,\sigma\right)} - \lambda_\sigma \frac{g_\sigma\left(\widehat{P}|e,\sigma\right)}{g\left(\widehat{P}|e,\sigma\right)}\right].$$
$$\tag{48}$$

This derivative of the Lagrangian is zero if the term in squared brackets in (48) is zero. Substituting equation (41) and rearranging yields:

$$\frac{\partial \mathcal{L}}{\partial \widehat{P}} = 0 \Leftrightarrow V(W(\widehat{P})) - V(0) - V'\left(W\left(\widehat{P}\right)\right) W(\widehat{P}) = 0. \tag{49}$$

With Tversky and Kahneman (1992) preferences (25) we obtain:

$$\alpha W(\widehat{P}) - \lambda \left(W^R\right)^\beta \left(W(\widehat{P}) - W^R\right)^{1-\alpha} - \left(W(\widehat{P}) - W^R\right) = 0. \tag{50}$$

With (43) equation (50) becomes:

$$\alpha W^R = \left(c_0 + c_1 \ln \widehat{P} + c_2 (\ln \widehat{P})^2\right) \lambda \left(W^R\right)^\beta + (1-\alpha)\left(c_0 + c_1 \ln \widehat{P} + c_2 (\ln \widehat{P})^2\right)^{\frac{1}{1-\alpha}}. \tag{51}$$

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

As the wage function W_T from (43) is quadratic, the solution to condition (51) 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 (43) as long as this function is monotone increasing. Otherwise, the optimal contract is constant. This proves (35).

Condition (51) 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 (51), and this proves (33). ∎

Appendix D: Calibration method

This appendix shows how the original optimization problem (3), (4), (7), and (8) can be transformed into (18) to (21) which can be calibrated to the data. Our derivations are analogue to those in Dittmann and Maug (2007). We start by rewriting the effort incentive constraint (7) so that the LHS of the equation does not contain any quantities that we cannot compute while the RHS does not contain the wage function (see Jenter (2002)):

$$PPS^{ua}(W_T(P_T)) = E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_0}\right] = \frac{C'(e)}{\frac{dP_0}{de}} \tag{52}$$

Under the null hypothesis that the model is correct, the observed contract fulfills this equation, so that the effort incentive constraint in our calibration problem can be written as (20). For the volatility incentive constraint (8), equations (13) and (17) imply

$$\rho(W_T(P_T)) = \frac{dP_0}{d\sigma}\frac{1}{P_0}. \tag{53}$$

Note that this equation again separates quantities that we cannot compute $(dP_0/d\sigma)$ from quantities that depend on the shape of the optimal contract (ρ). Under our null hypothesis, we therefore obtain (21). For the participation constraint (5), we first note that it must be binding as CEO utility is not downward restricted. If the constraint does not bind, we can shift the wage function downward until it binds. Under the null hypothesis the participation constraint can then be written as (19).

Appendix E: 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

$$W_T^{smth}(P_T) := \phi e^{r_f T} + n_S P_T e^{dT} + \sum_{i=1}^N n_O^i V(T^i, K^i, P_T) e^{r_f(T-T^i)}, \tag{54}$$

where $V(T^i, K^i, P_T) = E\left(\max\left\{P_{T^i} - K^i, 0\right\}|P_T\right)$. 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. The proceed at time T^i from exercising the option is then $V(T^i, K^i, P_T) = E\left(\max\left\{P_{T^i} - K^i, 0\right\}|P_T, P_0\right)$.

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

This Appendix contains formulae for our measure of risk avoidance ρ from (17) that can be readily implemented in a computer program. We start with a few definitions:

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

$$CV = \sigma\sqrt{T},$$

$$TW = (\phi + W_0)\exp\left\{r_f T\right\},$$

$$MD2 = \frac{\ln(K) - \ln(PC)}{CV},$$

With these definitions, we can calculate PPS^{ua} and ν^{ua} as follows:

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

where $f(u)$ is the standard normal density function. Our measure of risk avoidance then follows from (17).

For a CEO with more than one option grant, the option portfolio must first be aggregated into one representative option. We therefore 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 solved 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 multiply T^i by 0.7 to correct for early exercise (see Footnote 11 above).

References

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

Acharya, Viral V., Yakov Amihud, and Lubomir Litov, 2008, Creditor rights and corporate risk-taking, discussion paper, New York University.

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

Billett, Matthew T., David C. Mauer, and Yilei Zhang, 2010, Stockholder and bondholder wealth effects and CEO incentive grants, *Financial Management* 39, 463-587.

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.

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, 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.

DeFusco, Richard A., Robert R. Johnson, and Thomas S. Zorn, 1990, The effect of executive stock option plans on stockholders and bondholders, *Journal of Finance* 45, 617-627.

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.

Edmans, Alex, and Qi Liu, 2010, Inside debt, *Review of Finance*, forthcoming.

Edmans, Alex, and Xavier Gabaix, 2009, Tractability and detail-neutrality in incentive contracting, discussion paper, Wharton School.

Edmans, Alex, Xavier Gabaix, and Augustin Landier, 2008, A multiplicative model of optimal CEO incentives in market equilibrium, *Review of Financial Studies*, forthcoming.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov, 2009, Dynamic incentive accounts, discussion paper, New York University.

Faccio, Mara, Maria-Teresa Marchica, and Roberto Mura, 2010, Large shareholder diversification and corporate risk-taking, discussion paper, Purdue University.

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, 2006, Optimal compensation with induced moral hazard in investment, Working paper, University of Southern Denmark.

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

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.

Healy, Paul M., 1985, The effect of bonus schemes on accounting decisions, *Journal of Accounting and Economics* 7, 85-107.

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

Hemmer, Thomas, Oliver Kim, and Robert E. Verrecchia, 1999, Introducing Convexity into Optimal Compensation Contracts, *Journal of Accounting and Economics* 28, 307-327.

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.

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.

Jenter, Dirk, 2002, Executive compensation, incentives, and risk, Working paper, MIT Sloan School of Management.

Jenter, Dirk and Faid Kanaan, 2010, CEO turnover and relative performance evaluation, *Journal of Finance*, forthcoming.

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.

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, 2007, Motivating innovation, Working paper, MIT Sloan School of Management.

May, Don O., 1995, Do managerial motives influence firm risk reduction strategies?, *Journal of Finance* 50, 1291-1308.

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

Murphy, Kevin J., 2001, Performance standards in incentive contracts, *Journal of Accounting and Economics* 30, 245-278.

Oyer, Paul, 2004, Why do firms use incentives that have no incentive effects?, *Journal of Finance* 59, 1619-1650.

Peng, Lin, and Ailsa Röell, 2009, Managerial incentives and stock price manipulation, discussion paper, Columbia University.

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.

Smith, Gavin S., and Peter L. Swan, 2007, The incentive to 'bet the farm': CEO compensation and major investments, discussion paper, University of New South Wales.

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.

Tchistyi, Alexei, David Yermack, and Hayong Yun, 2007, Negative hedging: Performance sensitive debt and CEO's equity incentives, discussion paper, New York University.

Tufano, Peter, 1996, Who Manages Risk? An Examination of Risk Management Practices in the Gold Mining Industry, *Journal of Finance* 51, 197-1137.

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*, forthcoming.

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

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 727 CEOs from 2006. Panel B describes all 1,490 executives in the ExecuComp universe who are CEO in 2006.

Panel A: Data set with 727 U.S. CEOs

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Stock (%)	n_S	1.83%	4.94%	0.04%	0.32%	4.68%
Options (%)	n_O	1.37%	1.62%	0.14%	0.92%	3.17%
Base Salary ($m)	ϕ	1.64	4.47	0.51	1.04	2.43
Value of Contract ($m)	π_0	159.63	1,700.06	4.58	24.97	172.74
Non-firm Wealth ($m)	W_0	62.8	667.0	2.5	12.0	72.2
Firm Value ($m)	P_0	9,294	22,777	377	2,387	20,880
Strike Price ($m)	K	6,829	19,803	269	1,539	13,799
Moneyness (%)	K/P_0	70.1%	21.7%	41.2%	72.0%	100.0%
Maturity (years)	T	4.6	1.4	2.8	4.4	6.4
Stock Volatility (%)	σ	30.0%	13.4%	16.4%	28.3%	45.5%
Dividend Rate (%)	d	1.24%	2.25%	0.00%	0.63%	3.30%
CEO Age (years)		56.0	6.8	47	56	64
Stock Return 2001-5 (%)		11.8%	15.6%	-5.7%	11.4%	28.7%

Panel B: All 1,490 ExecuComp CEOs in 2006

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Stock (%)	n_S	1.95%	6.26%	0.02%	0.28%	4.22%
Options (%)	n_O	1.26%	1.57%	0.08%	0.79%	2.88%
Base Salary ($m)	ϕ	1.68	4.01	0.48	1.02	2.63
Firm Value ($m)	P_0	8,840	24,760	339	2,091	17,796
CEO Age (years)		55.1	7.1	46	55	64
Stock Return 2001-5 (%)		10.5%	23.2%	-13.8%	9.8%	34.1%

Table 2: Risk avoidance with Constant Relative Risk Aversion (CRRA)

This table displays descriptive statistics for risk avoidance ρ from equation (17) for five different values of the CRRA-parameter γ. Panel A shows results for the observed contract. Panel B displays results for the optimal CRRA-contract that does not take risk-taking into account.

Panel A: Observed contract

γ	Obs.	Mean	Standard Deviation	10% Quantile	Median	90% Quantile	Proportion with $\rho > 0$
0.5	727	0.19	0.39	-0.30	0.19	0.64	70.2%
1	727	0.62	0.56	-0.08	0.59	1.31	87.5%
2	727	1.33	0.86	0.30	1.25	2.43	96.8%
3	727	1.87	1.07	0.60	1.75	3.38	99.3%
6	727	2.91	1.50	1.13	2.68	4.88	99.7%

Panel B: Optimal CRRA-contract without risk-taking incentives

γ	Obs.	Mean	Standard Deviation	10% Quantile	Median	90% Quantile	Proportion with $\rho > 0$
0.5	727	1.32	0.63	0.62	1.26	2.11	99.9%
1	726	2.40	1.12	0.99	2.40	3.71	100.0%
2	727	5.74	18.92	3.64	6.71	8.58	99.9%
3	726	9.43	17.21	6.75	10.34	13.02	99.7%
6	652	12.04	7.25	0.02	15.02	18.77	99.4%

Table 3: Optimal CRRA contracts
with and without risk-taking incentives

This table describes the optimal contracts according to the CRRA-RTI model from equation (10) and the CRRA model from Dittmann and Maug (2007) for three different values of the CRRA parameter γ. The table displays mean and median of six measures that describe the optimal contract. The two distance metrics D_1 and D_2 are definied in equations (22) and (23). *Savings* are the difference in compensation costs between observed contract and optimal contract expressed as a percentage of costs of the observed contract, $(\pi_0^d - \pi_0^*)/\pi_0^d$. *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 probability that the end-of-period stock price P_T is smaller than the point where the wage schedule $W(P_T)$ starts to increase. The *inflection quantile* is the probability that the end-of-period stock price P_T is smaller than the point where the wage scheme turns from convex to concave. Panel A displays these statistics for all CEOs in our sample. The number of observations varies across different values of γ and across the two models due to numerical problems and because we exclude all CEO-γ-combinations for the CRRA-RTI model for which the observed contract implies negative risk-avoidance ρ from equation (17). Panel B shows results for those CEO-γ-combinations where we obtain 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	2.5%	8.0%	13.1%	14.2%	28.8%	36.5%
	median	1.9%	6.9%	10.2%	12.1%	27.9%	30.9%
Distance D_2	mean	5.8%	8.6%	13.1%	12.7%	26.2%	35.6%
	median	4.0%	7.4%	9.7%	10.9%	25.2%	30.3%
Savings	mean	0.1%	10.4%	30.6%	2.1%	34.7%	53.7%
	median	0.0%	6.9%	27.1%	1.1%	32.7%	54.3%
Minimum wealth	mean	3.1	1.7	1.3	0.0	0.0	0.0
	median	1.3	1.4	1.2	0.0	0.0	0.0
	Prop < 1	11.9%	0.1%	0.8%	100.0%	100.0%	100.0%
Kink quantile	mean	4.8%	19.6%	22.4%	0.0%	0.0%	0.0%
	median	1.7%	16.1%	19.5%	0.0%	0.0%	0.0%
Inflection quantile	mean	78.1%	34.9%	31.4%	2.1%	0.0%	0.0%
	median	77.7%	32.5%	29.3%	0.0%	0.0%	0.0%
Observations		388	688	373	727	726	652

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	2.5%	8.0%	13.5%	13.8%	28.3%	27.0%
	median	1.9%	6.9%	10.7%	11.9%	27.5%	25.0%
Distance D_2	mean	5.8%	8.6%	13.5%	13.0%	25.7%	26.7%
	median	4.0%	7.4%	10.2%	11.1%	24.8%	24.8%
Savings	mean	0.1%	10.4%	31.2%	1.7%	34.5%	54.0%
	median	0.0%	6.9%	28.2%	1.0%	32.1%	55.3%
Observations		388	688	334	388	688	334

Table 4: Optimal contracts for CARA utility

This table contains the results from repeating our analysis from Table 3 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 mean and median of six measures that describe the optimal contract. The two distance metrics D_1 and D_2 are definied in equations (22) and (23). *Savings* are the difference in compensation costs between observed contract and optimal contract expressed as a percentage of costs of the observed contract, $(\pi_0^d - \pi_0^*)/\pi_0^d$. *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 probability that the end-of-period stock price P_T is smaller than the point where the wage schedule $W(P_T)$ starts to increase. The *inflection quantile* is the probability that the end-of-period stock price P_T is smaller than the point where the wage scheme turns from convex to concave. The number of observations varies across different values of γ due to numerical problems and because we exclude all CEO-γ-combinations for the CARA-RTI model for which the observed contract implies negative risk-avoidance ρ from equation (17). 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	6.8%	9.3%	12.7%	22.2%	23.1%	24.6%
	median	5.7%	8.9%	12.4%	19.7%	22.6%	24.3%
Distance D_2	mean	9.2%	9.8%	12.8%	20.5%	20.4%	23.1%
	median	7.7%	9.1%	11.9%	18.1%	19.5%	22.7%
Savings	mean	2.4%	15.1%	25.8%	6.3%	27.4%	39.9%
	median	0.9%	12.1%	24.8%	3.7%	26.0%	40.2%
Minimum wealth	mean	2.9	2.2	2.0	0.0	0.0	0.0
	median	1.5	1.4	1.4	0.0	0.0	0.0
	Prop < 1	0.0%	0.0%	0.0%	100.0%	100.0%	100.0%
Kink quantile	mean	20.7%	22.9%	18.2%	0.0%	0.0%	0.0%
	median	17.2%	19.3%	14.7%	0.0%	0.0%	0.0%
Inflection quantile	mean	54.6%	36.6%	26.1%	0.0%	0.0%	0.0%
	median	52.6%	33.7%	22.8%	0.0%	0.0%	0.0%
Observations		279	419	594	279	419	594

Table 5: Model fit for subsamples

This table shows mean distance D_1 from equation (22) for quintiles formed according to four variables: initial non-firm wealth W_0, CEO age, firm value P_0, and the past five year stock return (from the start of 2001 to the end of 2005). 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	Wealth W_0 (in $m)		CEO Age		Firm Value P_0 (in $m)		Stock return 2001-2005	
	Mean	D_1	Mean	D_1	Mean	D_1	Mean	D_1
1	2.6	11.4%	41.9	11.7%	386	8.7%	-18.8%	7.5%
2	6.6	7.8%	48.1	9.4%	1,135	8.5%	3.6%	6.8%
3	12.3	7.5%	52.5	8.1%	2,358	8.1%	11.3%	7.5%
4	26.1	6.8%	57.0	7.0%	5,648	7.2%	18.9%	8.5%
5	270.1	6.8%	64.6	7.6%	32,685	7.8%	43.5%	10.6%
P-Value Q1-Q5	0.0000		0.0040		0.9583		0.0001	

Table 6: Optimal contracts that consist of salary, stock, and options

This table describes the optimal piecewise linear contract for three different values of the CRRA parameter γ. The table displays mean and 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: $(\pi_0^d - \pi_0^*)/\pi_0^d$. The number of observations varies across different values of γ because we exclude all CEO-γ-combinations for which the observed contract implies negative risk-avoidance ρ from equation (17).

		$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$
Salary ϕ^*	mean	5.15	8.19	6.64
	median	1.24	3.23	2.76
Stock n_S^*	mean	0.45%	0.08%	0.20%
	median	0.00%	0.00%	0.00%
Options n_O^*	mean	2.38%	2.22%	1.41%
	median	1.37%	1.32%	1.03%
Moneyness K^*/P_0	mean	51.4%	51.7%	45.0%
	median	52.1%	51.9%	45.3%
	Prop.<K/P0	83.3%	99.2%	99.8%
Savings	mean	0.1%	4.1%	15.4%
	median	0.0%	1.9%	10.9%
Observations		102	632	552

Table 7:

Optimal loss aversion contracts without risk-taking incentives

This table 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. The table displays mean and median of five measures that describe the optimal contract. The two distance metrics D_1 and D_2 are definied in equations (22) and (23). *Savings* are the difference in compensation costs between observed contract and optimal contract expressed as a percentage of costs of the observed contract, $(\pi_0^d - \pi_0^*)/\pi_0^d$. The *jump quantile* is the probability that the end-of-period stock price P_T is smaller than the point where the contract jumps from the lowest possible payout to some payout above the reference wealth. The *inflection quantile* is the probability that the end-of-period stock price P_T is smaller than the point where the wage scheme turns from convex to concave. The number of observations varies across different values of θ due to numerical problems.

		$\theta = 0.1$	$\theta = 0.5$	$\theta = 0.9$
Distance D_1	mean	5.8%	19.0%	31.7%
	median	4.3%	15.9%	29.0%
Distance D_2	mean	6.4%	17.6%	28.6%
	median	4.5%	15.4%	26.3%
Savings	mean	0.8%	4.8%	9.5%
	median	0.1%	3.8%	8.4%
Jump quantile	mean	0.25%	3.47%	9.36%
	median	0.00%	1.80%	7.79%
Inflection quantile	mean	100%	100%	100%
	median	100%	100%	100%
Observations		715	676	586

Table 8: Risk avoidance when managers are loss averse

This table displays descriptive statistics for risk avoidance ρ from equation (17) for three different levels of reference wealth W^R parameterized by θ. Panel A shows results for the observed contract. Panel B displays results for the optimal LA contract that does not take risk-taking into account. Panel C defines six cases for changes in risk avoidance from the observed contract to the optimal LA contract and reports the relative frequency with which these cases apply for each of the three levels of reference wealth.

Panel A: Observed contract

θ	Obs.	Mean	Standard Deviation	10% Quantile	Median	90% Quantile	Proportion with $\rho > 0$
0.1	727	-0.04	0.28	-0.44	0.01	0.27	51.3%
0.5	727	0.27	0.37	-0.24	0.31	0.72	76.2%
0.9	727	0.41	0.38	-0.12	0.46	0.87	84.6%

Panel B: Optimal LA contract

θ	Obs.	Mean	Standard Deviation	10% Quantile	Median	90% Quantile	Proportion with $\rho > 0$
0.1	715	-0.16	0.63	-0.67	-0.22	0.25	30.8%
0.5	676	-0.13	1.01	-1.06	-0.34	0.88	34.0%
0.9	586	0.55	1.38	-1.30	0.62	2.21	71.8%

Panel C: Changes in risk avoidance

θ	Obs.	Case 1: $\rho^{obs} > 0$ $\rho^{LA} \geq \rho^{obs}$	Case 2: $\rho^{obs} > 0$ $\rho^{LA} \in [0, \rho^{obs})$	Case 3: $\rho^{obs} > 0$ $\rho^{LA} < 0$	Case 4: $\rho^{obs} < 0$ $\rho^{LA} \leq \rho^{obs}$	Case 5: $\rho^{obs} < 0$ $\rho^{LA} \in (\rho^{obs}, 0]$	Case 6: $\rho^{obs} < 0$ $\rho^{LA} > 0$
0.1	715	13.0%	10.5%	27.6%	29.5%	12.2%	7.3%
0.5	676	15.2%	11.2%	50.7%	12.9%	2.4%	7.5%
0.9	586	44.0%	17.6%	24.2%	2.7%	1.2%	10.2%

Table 9: Optimal LA contracts with and without risk-taking incentives

This table describes the optimal contracts according to the LA-RTI model from equations (33), (34), and (35) and the LA model from Dittmann, Maug, and Spalt (2010) for three different levels of reference wealth W^R parameterized by θ. The table displays mean and median of five measures that describe the optimal contract. The two distance metrics D_1 and D_2 are definied in equations (22) and (23). *Savings* are the difference in compensation costs between observed contract and optimal contract expressed as a percentage of costs of the observed contract, $(\pi_0^d - \pi_0^*)/\pi_0^d$. The *jump quantile* is the probability that the end-of-period stock price P_T is smaller than the point where the contract jumps from the lowest possible payout to some payout above the reference wealth. The *inflection quantile* is the probability that the end-of-period stock price P_T is smaller than the point where the wage scheme turns from convex to concave. The number of observations is small and varies across different values of θ, because we only consider the CEOs from Case 1 in Table 2, Panel C. In the other cases, either our model assumptions are violated or the optimal LA and LA-RTI contracts are identical. We also lose some observations due to numerical problems.

		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	2.1%	20.6%	43.0%	2.2%	21.8%	37.3%
	median	0.7%	17.4%	37.7%	1.0%	19.4%	37.1%
Distance D_2	mean	2.2%	18.9%	37.4%	2.2%	20.0%	33.5%
	median	0.8%	15.3%	32.3%	1.3%	17.5%	32.8%
Savings	mean	0.4%	7.3%	8.7%	1.1%	8.3%	11.2%
	median	0.0%	7.5%	8.8%	0.0%	8.0%	10.4%
Jump quantile	mean	0.1%	3.1%	5.3%	0.3%	6.7%	14.3%
	median	0.0%	0.0%	0.1%	0.0%	5.4%	13.5%
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		75	85	182	75	85	182