TI 2009-076/2
Tinbergen Institute Discussion Paper

How Important Are Risk-Taking Incentives in Executive Compensation?

Ingolf Dittmann
Ko-Chia Yu

Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

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

Ingolf Dittmann[†] Ko-Chia Yu[‡]

August 24, 2009

Abstract

This paper investigates whether observed executive compensation contracts are designed to provide risk-taking incentives in addition to effort incentives. We develop a stylized principal-agent model that captures the interdependence between firm risk and managerial incentives. We calibrate the model to individual CEO data and show that it can explain observed compensation practice surprisingly well. In particular, it justifies large option holdings and high base salaries. Our analysis suggests that options should be issued in the money. If tax effects are taken into account, the model is consistent with the almost uniform use of at-the-money stock options. We conclude that the provision of risk-taking incentives is a major objective in executive compensation practice.

JEL Classifications: G30, M52

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

---

*We are grateful to Jennifer Carpenter, Pierre Chaigneau, John Core, Alex Edmans, Rüdiger Fahlenbrach, Xavier Gabaix, Robert Gillenkirch, Wayne Guay, Dirk Jenter, Ernst Maug, Oliver Spalt, Jack Stecher, David Walker, David Yermack, Dan Zhang, and seminar participants at the Accounting Research Workshop (Berne), the CEPR Summer Symposium (Gerzensee), NHH Bergen, NYU Stern School of Business, Stanford Graduate School of Business, University of Mannheim, University of Zurich, Vrije Universiteit Amsterdam, and the Wharton School. We acknowledge financial support from NWO through a VIDI grant.

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

[‡]Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands. Email: yu@ese.eur.nl. Tel: +31 10 408 8947.

1 Introduction

The relation between risk and incentives is an important issue in executive compensation that is still not fully understood. Standard agency theory puts forth the informativeness principle according to which firms should use less incentive pay if the stock price is more volatile and therefore less informative of the agent's effort. There is, however, little empirical evidence for the informativeness principle in executive compensation.[1] Another strand of the literature recognizes that the CEO's actions affect firm risk, so that incentives can have an effect on risk. There is ample empirical evidence that CEOs indeed respond to risk-taking incentives and adjust their actions accordingly.[2] Also the reaction of stock and bond prices to first time equity grants implies that investors expect that these grants affect firm risk (see DeFusco, Johnson and Zorn (1990) and Billett, Mauer and Zhang (2006)). What is little understood, however, is whether shareholders provide these risk-taking incentives on purpose. Alternatively, shareholders might only be interested in effort incentives, and the risk-taking incentives documented in the literature could just be a side effect of these effort incentives.

We approach this question with a new calibration method that avoids the endogeneity problems that invariably arise in regressions that include firm risk and managerial incentives. Instead we model the endogeneity between risk and incentives and test the implications of the model. We develop a stylized principal-agent model from standard building blocks where an effort-averse and risk-averse agent chooses his effort and the firm's strategy, and where the strategy affects firm value and risk. This model incorporates not only the notion that the CEO's actions affect firm risk and firm value, but also the informativeness principle that higher firm risk makes the stock price less informative and weakens effort incentives. We calibrate this model to the data on 737 U.S. CEOs and for each generate predictions about the optimal compensation structure, i.e. the optimal mix of base salary, stock, and options, and the optimal strike price. We then compare these predictions with observed data and find that our model can explain observed compensation structures - and in particular the pervasive use of stock options - much better than models that do not take into account risk-taking incentives. We therefore conclude that the provision of risk-taking incentives is a major objective in executive compensation practice.

[1] Prendergast (2002) provides a brief survey of the empirical evidence for the informativeness principle in executive compensation. Three of the eleven papers included in the survey find a significant negative relationship between risk and incentives. Three other papers find a significant positive relationship, and the remaining five papers find no significant relationship.

[2] Tufano (1996) and Knopf, Nam and Thornton (2002) show this 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. Low (2009) analyzes changes in firm risk after a change in takeover legislation.

In our model, CEOs are poorly diversified because a large part of their wealth is linked to the company's share price to provide effort incentives. In the absence of proper risk-taking incentives, CEOs therefore prefer low firm risk and tend to choose a firm strategy that results in inefficiently low risk. They might, for instance, pass up a profitable but very risky project, or they might hedge their firm's risk at some cost. Shareholders can reduce this inefficiency by providing risk-taking incentives. The challenge is to provide risk-taking incentives without impairing effort incentives. 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 ones, i.e. the optimal contract resembles a call option on the firm's stock.

In line with this intuition, our calibrations predict contracts with large option holdings and little or no stock. The optimal strike price is lower than the observed strike price which indicates that options should be issued in the money according to the model. In-the-money options provide incentives for intermediate and high outcomes and they avoid punishing the CEO for bad outcomes. Hence they provide effort and risk-taking incentives at the same time. Our model also predicts higher base salaries than observed, because these must rise as stock is replaced by less valuable options to guarantee the CEO's reservation utility. The savings that can be expected when firms switch from the observed contract to the optimal contract are low and average only 5.3% of total compensation costs.

In an otherwise similar model without risk-taking incentives, Dittmann and Maug (2007) find optimal contracts that save 54% of compensation costs. These contracts involve negative option holdings, much higher stock holdings, and negative salaries, i.e. CEOs are required to invest a sizeable proportion of their wealth into their own firm. Even if alternative preference specifications are taken into account, at least 26% of the CEOs should have negative option holdings or negative fixed salaries (see Dittmann, Maug and Spalt (2009)). Hence, risk-taking incentives can rationalize why the vast majority of CEOs receive positive salaries and option grants, whereas effort-incentives alone cannot explain this stylized fact. We conclude from this comparison that risk-taking incentives play an important role in executive compensation practice.

The U.S. tax system strongly discriminates against in-the-money options.[3] In our calibrations,

[3] According to IRC Section 162(m), in-the-money stock options are not considered as performance-based compensation, so that the "one-million-dollar" rule applies and only up to $1m (including base salary) are deductible on corporate tax returns. Moreover, Section 409A requires that the difference between the stock price and the strike price

the savings from recontracting are much smaller than the additional tax penalties most firms and executives would have to pay if in-the-money options were used. If we include these tax penalties in our model, observed contracts turn out to be optimal for 76% to 94% of all CEOs in our sample (depending on assumptions), so our model is broadly consistent with compensation practice. In this context, our analysis suggests that the current U.S. tax system forces firms to resort to inefficient contract arrangements, because most firms - and especially small firms with poor past performance - could benefit from granting in-the-money options. Moreover, our analysis shows that 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.

Our calibration approach bridges the gap between theoretical and empirical research on executive compensation. It permits us to test the quantitative (and not just the qualitative) implications of our model, and it generates a price tag for any deviation from optimality. Moreover, this approach 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. Our evidence is indirect, but it is free of any endogeneity bias.

Our analysis also gives rise to a new measure of risk-taking incentives. Most researchers use the vega of the manager's portfolio, i.e. its sensitivity to a change in stock return volatility. Our model suggests that the ratio of vega to delta (the sensitivity to a change in stock price) is a superior measure. This measure takes into account not only the direct effect of an increase of volatility on the manager's wealth (which is vega), but also the indirect effect that higher volatility is associated with a higher stock price. An increase in stock price then feeds through to managerial wealth via the manager's incentive pay, i.e. delta.[4]

We also contribute to the discussion 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-

be recognized as income at the time of vesting, rather than on exercise. Thus this rule accelerates income recognition from the exercise date to the vesting date. In addition, Section 409A imposes an additional 20% tax on this income (see Alexander, Hirschey, and Scholz (2007)).

[4]In the second-best optimum, the CEO's utility decreases in firm risk, so that he takes inefficiently low risk (relative to the first-best optimum). If the CEO takes on more risk in this situation, firm value increases. The utility-adjusted vega is negative, so that the ratio of vega to delta increases in vega and in delta. On the other hand, if risk-taking incentives are so high that the CEO's utility increases in risk, vega is positive and the ratio of vega to delta increases in vega and decreases in delta. In this situation, the CEO takes inefficiently high risk. As risk-taking incentives are costly, this latter situation will never be an equilibrium in our model.

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.

There are a few theory papers that also consider both effort-aversion and risk-taking incentives in models of executive compensation.[5] To our knowledge, this paper is the first, however, to calibrate such a model and to test its quantitative implications. While we attribute the existence of options to the provision of risk-taking incentives in this paper, we acknowledge that there are alternative explanations for the use of options in executive compensation.[6] We also contribute to recent literature on calibrations of contracting models by considering a richer and more realistic model than previous papers.[7]

In the next section, we present our model in which the manager must choose effort and the firm's strategy. We derive the appropriate measure of risk-taking incentives in Section 2.2 and explain our calibration approach in Section 2.3. 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 2.4 discusses the construction of our data set. In Section 3, we present our main results on optimal piecewise linear contracts consisting of base salary, stock, and an option grant. Section 4 discusses reasons why in-the-money options are rarely used in practice. In particular, we analyze the impact of U.S. taxes on our calibration results here. In order to better understand our results for piecewise linear contracts, we then theoretically derive the unrestricted shape of the optimal contract in Section 5 and calibrate this shape to the data. Section 6 contains robustness

[5] Lambert (1986) and Core and Qian (2002) consider discrete volatility choices, where the agent must exert effort in order to gather information about the 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 (2008) and Sung (1995) solve models with continuous effort and volatility choice, but Hellwig (2008) assumes that the agent is risk-neutral and Sung (1995) that the principal can observe (and effectively set) volatility.

[6] Oyer (2004) models options as a device to retain employees when recontracting is expensive, and 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. 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. Dittmann, Maug and Spalt (2009) show that options can be explained if managers are loss-averse. With the exception of Dittmann, Maug and Spalt (2009), none of these models has been calibrated to data, and some models are too stylized to be calibrated at all.

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

checks, and Section 7 concludes. The appendix explores the validity of the first-order approach and contains the proofs.

2 The model and its calibration

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 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 $\sigma = \sigma(s)$.[8]

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, etc. Due to its richness, we do not model the agent's choice of strategy in detail. Instead we recognize that a risk-averse agent with a wage contract $w(P_T)$ that is increasing in P_T will always choose 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 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

[8] 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.

profitable projects (e.g. an R&D project), or he takes an additional action that 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: if the agent can costlessly take on more risk in financial markets, it is flat; otherwise, a higher value of σ also leads to a distortion of the agent's actions and thereby to a lower firm value. Finally, we assume that the stock price $P_0(e, \sigma)$ is increasing and concave in e (given volatility σ).

Our model is in the spirit of Holmström (1979): The principal cannot observe the agent's actions e and σ, so the manager's wage W_T only depends on the end-or-period stock price P_T.[9] We use risk-neutral pricing and 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.[10]

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

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

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 $C''(e) > 0$. 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

[9] The ex-post volatility can obviously be estimated from stock returns, but these are only realizations of the ex-ante distribution whose volatility the CEO selects. Moreover, volatility is not exclusively determined by the CEO's management strategy. If the CEO has other means to drive up volatility (e.g. by frequent contradictory announcements), total observed volatility can be manipulated and can be higher than the fundamental volatility the CEO selects in our model.

[10] Risk-neutral pricing 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.

that give him expected utility $\overline{U}$. The shareholders' optimization problem then is:

$$\max_{W_T, e, \sigma} E\left[P_T - W_T(P_T) | e, \sigma\right] \tag{3}$$

$$\text{subject to} \quad E\left[V(W_T(P_T)) | e, \sigma\right] - C(e) \geq \overline{U} \tag{4}$$

$$\text{and} \qquad \{e, \sigma\} \in \text{argmax } \left\{E\left[V(W_T(P_T)) | e, \sigma\right] - C(e)\right\} \tag{5}$$

We replace the incentive compatibility constraint (5) with its first-order conditions:

$$\frac{dE\left[V(W_T(P_T)) | e, \sigma\right]}{de} - \frac{dC}{de} = 0 \tag{6}$$

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

We discuss the validity of the first-order approach (i.e. that (5) can indeed be replaced by (6) and (7)) in detail in Appendix A. We call condition (6) the effort incentive constraint and (7) the volatility incentive constraint.

2.2 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.[11] 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 empirical literature on risk-taking incentives: higher volatility leads to higher firm value (as more valuable risky projects are adopted) and via the pay-for-performance sensitivity (the delta of the manager's equity portfolio) to higher managerial utility. Consequently, there are two ways to provide risk-taking incentives: increasing (utility adjusted) vega or increasing (utility adjusted) delta. 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 (7). This

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

constraint can be rewritten as

$$E\left[\left.\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}\frac{dP_T}{d\sigma}\right| e,\sigma\right]=0 \tag{8}$$

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

$$\Leftrightarrow E\left[\left.\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}\left(\frac{dP_0}{d\sigma}\frac{P_T}{P_0}+P_T\left(-\sigma T+u\sqrt{T}\right)\right)\right| e,\sigma\right]=0. \tag{9}$$

As $dP_0/d\sigma$ is not random, we can rearrange (9) as

$$PPS^{ua}\frac{dP_0}{d\sigma}=-\nu^{ua}, \tag{10}$$

$$\text{where } PPS^{ua}:=E\left[\left.\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_0}\right| e,\sigma\right]=E\left[\left.\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}\frac{P_T}{P_0}\right| e,\sigma\right] \tag{11}$$

$$\text{and} \qquad \nu^{ua}:=E\left[\left.\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_T}P_T\left(-\sigma T+u\sqrt{T}\right)\right| e,\sigma\right]. \tag{12}$$

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 (10) equals 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. Rewriting the first order condition (10) yields our proposed measure for total risk-taking incentives:

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

RTI is indeed a measure of risk-taking incentives: a manager with higher RTI will choose higher volatility σ, because $P_0(\sigma)$ is concave and $dRTI/d\sigma>0$. The first-best solution is $RTI=0$. Then the manager is indifferent to firm risk and will choose the optimal firm strategy. If $RTI<0$, risk-taking incentives are inefficiently low (relative to the first-best solution); if $RTI>0$, they are inefficiently high. Thus while risk-taking incentives always increase in vega ν^{ua}, they increase in PPS^{ua} only if $RTI<0$.

With a risk-averse manager, the second-best RTI will always be negative. Risk-taking incentives increase the uncertainty of the agent's payoff and are therefore costly. Consequently, the firm will increase risk-taking incentives only up to the point where the marginal costs of additional incentives are equal to the marginal benefit of an increase in firm risk. In this second-best situation, RTI increases in vega ν^{ua} and in the pay-for-performance sensitivity PPS^{ua}. Stock grants have a positive effect on PPS^{ua} and a negative effect on ν^{ua}, and the combined effect on RTI is negative. The same is true for options as long as their strike price is low (see Ross, 2004).

2.3 Calibration method

We cannot calibrate the full optimization problem to the data, because this requires knowledge (or estimates) of the production function $P_0(e, \sigma)$ and of the cost function $C(e)$. We therefore resort to the subproblem of finding a new contract with a given shape that achieves three objectives. Firstly it provides the same effort and risk-taking incentives to the agent as the observed contract. Secondly it provides the agent with the same utility as the observed contract, and thirdly it is as cheap as possible for the firm. This subproblem is the first stage of the two-stage procedure in Grossman and Hart (1983), where they search for the cheapest contract that implements a given level of effort. In our case, this is the level of effort that is implemented by the observed contract. 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 investment choices as the observed contract. In this case, either compensation practice is inefficient or the model is incorrect. In both cases, the model is not descriptive of the data.

We only calculate the cost-effective contract 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)). Our method analyzes the optimal structure of compensation only.

We start by rewriting the effort incentive constraint (6) 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)):

$$PPS^{ua}(W_T(P_T), e, \sigma, \gamma) = E\left[\frac{dV(W_T)}{dW_T}\frac{dW_T}{dP_0}\bigg| e, \sigma\right] = \frac{C'(e)}{\frac{dP_0}{de}} \tag{14}$$

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 becomes:

$$PPS^{ua}(W_T^*(P_T), e, \sigma, \gamma) = PPS^{ua}(W_T^d(P_T), e, \sigma, \gamma) \tag{15}$$

Here W_T^* denotes the new (cost minimizing) contract and W_T^d denotes the observed contract (d for "data").

We can reformulate the participation constraint (4) and the volatility incentive constraint (7) in a similar way:

$$E\left[V(W_T^*(P_T))|e, \sigma, \gamma\right] \geq E\left[V(W_T^d(P_T))|e, \sigma, \gamma\right], \tag{16}$$

$$RTI(W_T^*(P_T), e, \sigma, \gamma) = RTI(W_T^d(P_T), e, \sigma, \gamma). \tag{17}$$

For our calibration approach to work, we also need to restrict the shape of the optimal contract, so that it depends on only a few parameters. In Section 5, we derive the optimal contract shape which depends on three parameters and we calibrate this to the data. In the next section, we calibrate a piecewise linear contract that consists of fixed salary ϕ, the number of shares n_S, and the number of options n_O with strike price K:

$$W_T^{lin}(P_T) = (W_0 + \phi)\exp\{r_f T\} + n_S P_T + n_O \max\{P_T - K, 0\}. \tag{18}$$

With W_0 we denote the manager's initial non-firm wealth, i.e. all wealth that is not invested in stock or options of his own firm. We express the number of shares n_S and the number of options n_O as a percentage of outstanding shares, so that $0 \leq n_S \leq 1$. Our numerical optimization problem is to minimize the costs of the new contract, $E\left(W_T^{lin}(P_T)|e^d, \sigma^d\right)$, subject to the constraints (15), (16), and (17). We have four parameters to minimize costs over: ϕ, n_S, n_O, and K.

2.4 Data set

We use the ExecuComp database to construct approximate CEO contracts at the beginning of the 2006 fiscal year. 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)).[12]

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 (in fact at (18)) 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.[13] In this step, we lose five CEOs for whom we cannot numerically solve this system of three equations in three unknowns.

We take the firm's market capitalization P_0 from the end of 2005. 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). 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 737 CEOs.

[Insert Table 1 here]

[12] 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 (2007)), so that we overestimate risk-taking incentives in observed contracts.

[13] We take into account the fact that most CEOs exercise their 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 $T = 10$ as these are typical values for newly granted options.

Table 1 Panel A provides an overview of our data set. The median CEO owns 0.3% of the stock of his company and has options on an additional 1% of the company's stock. The median base salary is \$1.1m, and the median non-firm wealth is \$11.1m. The representative option has a median maturity of five 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.

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 own somewhat more options (+0.1%). They work in bigger firms (+\$500m) with better past performance (1.25% higher return during the past five years). We conclude that our sample is subject to a moderate survivorship bias. We investigate this bias by separately analyzing subsamples with more successful and less successful CEOs in Section 6.

The only parameter in our model that we cannot estimate from the data is the manager's coefficient of relative risk aversion γ. We therefore repeat our analysis for six different risk-aversion parameters ranging from $\gamma = 0.5$ (low risk-aversion) to $\gamma = 8$ (strong risk-aversion). This range includes the risk-aversion parameters used in previous research. We regard values of γ below 1 as unrealistically low as they imply implausible private portfolio decisions: with $\gamma < 1$, the CEO would like to borrow heavily and invest much more than his entire wealth in the stock market.

3 Optimal piecewise linear contracts

In this section we present our main empirical results. For each CEO in our sample, we numerically calculate the cheapest piecewise linear contract that provides the manager with the same utility and the same incentives as the observed contract. We call this cheaper contract the "optimal contract" and compare it with the observed contract.

More formally, we minimize $E\left(W_T^{lin}(P_T)\right)$ subject to the participation constraint (16) and the two incentive compatibility constraints (15) and (17). 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.[14]

We also need to restrict the strike price K, because options and shares become indistinguishable if K approaches zero, and the problem becomes 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.[15]

[Insert Table 2 here]

Table 2, Panel A contains our calibration results for six values of the risk-aversion parameter γ, ranging from 0.5 to 8. For low values of risk-aversion we lose some of our 737 observations, because risk-taking incentives from (13) are positive.[16] The column *Observations* displays the remaining observations after CEOs with positive ν^{ua} have been deleted, and the column *Converged* shows the number of CEOs for which our numerical routine was successful. In addition, the table describes the four contract parameters ϕ, n_S, n_O, and K of the calibrated optimal contract, and the percentage savings the firm could realize by switching from the observed contract W_T^d to the optimal contract W_T^* , i.e.

$$savings = \left[E\left(W_T^d(P_T)\right) - E\left(W_T^*(P_T)\right)\right] / E\left(W_T^d(P_T)\right). \tag{19}$$

Optimal contracts differ systematically from observed contracts regarding the CEO's stock holdings. While observed contracts nearly always contain stock holdings, 99% of all CEOs would not receive any shares according to the optimal contract for $\gamma = 3$. Instead, the strike price of their option holdings would be much lower: the median strike price is 51% of the share price compared to

[14] We do not allow for negative stockholdings, because compensation could then become non-monotonic in stock price. We discuss non-monotonic contracts in Section 5 and Appendix A.

[15] 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.

[16] As long as the agent is risk-averse, our model predicts negative RTI in equilibrium (see the discussion at the end of Section 2.2). Therefore, a positive RTI directly rejects our model assumptions. We interpret the fact that $RTI > 0$ for many CEOs for $\gamma \leq 1$ as a confirmation that these levels of risk-aversion are unrealistically low. Note that, for the more reasonable value $\gamma = 3$, virtually all the CEOs in our sample have negative RTI.

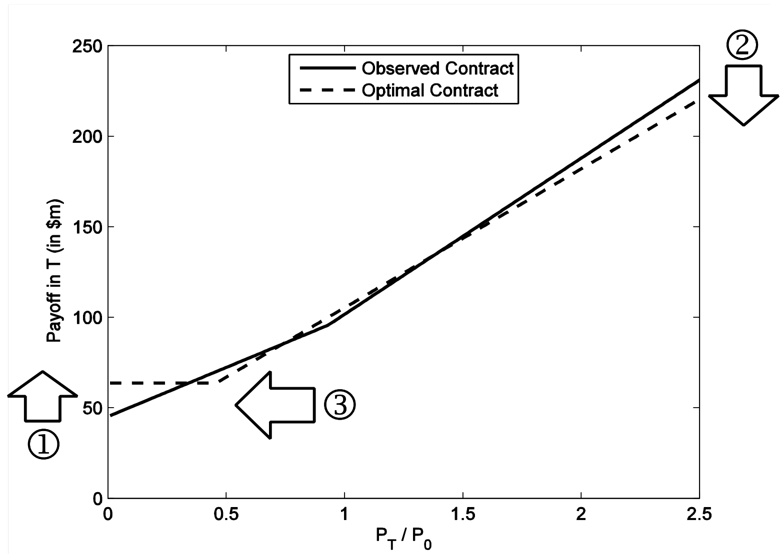

Figure 1: The figure shows end-of-period wealth W_T as a function of end-of-period stock price P_T for the observed contract (solid line) and the optimal piecewise linear contract (dashed line) for one CEO in our sample. The arrows indicate the three main features of the optimal contract relative to 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, $n_S = 5.97\%$, $n_O = 4.45\%$ for the observed contract. Initial non-firm wealth is $W_0 = \$30.3$m. $P_0 = \$853$m, $\sigma = 25.6\%$, and $K/P_0 = 90\%$, $T = 4.9$ years, $r_f = 4.4\%$, $d = 0.9\%$. All calculations are for $\gamma = 3$.

72% for the observed contract. While average and median option holdings are higher for the optimal contract with $\gamma = 3$, this is not uniformly so for all CEOs. Instead, we find that the *sum* of stock and options is always smaller in the optimal contract than in the observed contract (not shown in the table). Therefore, the optimal contract is less steep than the observed contract in the best states of the world.

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 model predicts that median base salaries (for $\gamma = 3$) should nearly triple from \$1.1m to \$3.2m. For $\gamma \geq 1$, optimal base salaries and option holdings are virtually always positive. Hence, a model with effort and risk-taking incentives can explain these stylized facts far better than models that account for effort incentives only. In those models, at least 25% of the CEOs should receive no options or a negative fixed salary (see Dittmann and Maug (2007) and Dittmann, Maug and Spalt (2009)).

Figure 1 illustrates our main results. It shows the payout function $W_T(P_T)$ 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 1 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 for the optimal contract than for the observed contract, and the optimal contract is cheaper for shareholders.

However, the savings generated by switching to the optimal contract are limited. For $\gamma = 3$, the median firm would just save 2.6% of its compensation costs (the average is 5.3%, see Table 2, Panel A). The savings in the case shown in Figure 1 are 2.8%. 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 15), it is not a *problem* of our model but rather a *result*.

While 98.8% of the CEOs in our sample would not receive any stock if firms implemented the optimal contract, there are still 1.2% who would. A more detailed analysis (not shown in the tables) shows that there are two reasons for these positive stockholdings. A few CEOs have no options in their observed contract, so that it is not possible to construct an alternative contract with all the three features highlighted in Figure 1. For other CEOs, our optimization routine hits the boundary $K/P_0 = 20\%$ or $K/P_0 = 10\%$, so that we have a corner solution with positive stock holdings. Beyond these two cases, we find no true interior solutions with $n_S^* > 0$, except for $\gamma = 0.5$. We therefore conclude that, within our model, in-the-money options are generally preferable to a portfolio of at-the-money options and stock.

Table 2, Panel B reproduces the results from Panel A for those 282 CEOs for which our algorithm converges for all $\gamma \geq 1$. This table shows that, as γ increases, the optimal contract features fewer stock options, lower strike prices, and lower base salaries. Therefore, the contract becomes flatter and less convex as γ increases. Savings are considerable for high levels of risk-aversion and negligible for $\gamma = 1$. This finding is not surprising as savings stem from improved risk-sharing, which is more important if CEOs are more risk-averse.

4 Taxes and the popularity of at-the-money options

The low savings from recontracting shown in Table 2 imply that observed compensation practice is consistent with our model if there is an effect (possibly even small) in favor of shareholdings or at-the-money options that we did not account for in our model. In this section, we review a few potential reasons why at-the-money options are so popular in compensation practice.

The U.S. tax system strongly discriminates against in-the-money options (see Footnote 3). 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. Walker (2009) argues that this rule "is probably the measure that most strongly discourages explicit grants of in-the-money options." 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.[17]

To illustrate the effect of taxes, we first consider a representative CEO whose parameters are closest to the median values shown in Table 1.[18] The observed contract of this representative CEO consists of $1.1m base salary, $7.9m stock, and at-the-money options with a Black-Scholes value of

[17] 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. These accounting reasons probably explain the absence of in-the-money options before 2004, the year in which Section 409A was enacted.

[18] 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 relative difference for each CEO and select the CEO for whom this maximum difference is smallest.

$12.1m. Our model proposes instead $3.9m base salary, no stock, and in-the-money options with a value of $17.0m. This contract would generate savings of $0.2m or 1% of total compensation costs, but the CEO would have to pay additional taxes of $3.4m (= 20% · $17m) in expectation, so that a portfolio of stock and at-the-money options is cheaper than in-the-money options if taxes are taken into account.

In order to investigate this tax effect more systematically, we repeat our numerical analysis for $\gamma = 3$ with the 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 93.7% of all CEOs for whom our algorithm converges (not shown in the tables).

This tax analysis does not take into account that part of the shares held by an executive might not be restricted but held voluntarily. If these are replaced by in-the-money options, the executive would have to sell them and buy in-the-money options from the proceeds. As these options are bought from private wealth, they would not be subject to the 20% penalty tax.[19] In the above example, all the shares held by the representative CEO are unrestricted. If he sells them and invests the proceeds of $7.9m into options, only $9.1m (= $17m −$7.9m) are subject to the penalty tax, resulting in a penalty of $1.82m which still exceeds the benefits from recontracting ($0.2m). For the full sample, we find that the optimal contract remains optimal for 75.5% of all CEOs under these assumptions. For the remaining 24.5% the optimal contract is identical to the optimal contract without taxes (see Table 2), except that more options are awarded to compensate the CEO for the tax payment.

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).[20] A potential reason is that the rest of the world generally tends to follow the U.S. when it comes to executive compensation and especially executive stock options. Alternatively, one can argue that in-the-money options cause some costs that are not included in our model and that justify government intervention. Our results in Table 2 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. Alternatively, there might be concerns that executives

[19] It is not obvious that this second way to include taxes in our model is necessarily the more accurate one. Unrestricted shares can also be seen as the result of restricted stock awarded in previous periods. If in-the-money options instead of restricted stock had been issued in the previous periods, the tax penalty would have applied.

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

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 could reduce this rent-seeking activity, and our analysis shows that the costs of such a commitment are low.

5 Optimal nonlinear contracts

A limitation of our analysis so far is that we only consider piecewise linear contracts with one kink (i.e. one option grant). In reality however, CEOs have many different option grants with different strike prices, and we could augment our model by allowing for a second option grant with a different strike price. Given the numerical difficulties in the case with one option, however, we do not consider including a second option grant (and thereby two additional choice variables) as a fruitful strategy. Instead, we now turn to the general non-linear contract. We theoretically derive the shape of the optimal contract, parameterize it, and then numerically search for the cheapest contract with this theoretically optimal shape. As any piecewise linear contract will be an approximation to this general contract, we can infer the shape of any piecewise linear contract from this general contract. Moreover, the savings of the general contract are an upper bound for the savings that can be achieved by any piecewise linear contract.

5.1 Theoretical shape of the optimal contract

In this subsection we solve the shareholders' problem and establish the shape of the optimal contract (which is non-linear).

Proposition 1. ***(Optimal general contract):*** *The optimal contract that solves the shareholders' problem (3), (4), (6), and (7) has the following functional form:*

$$W_T^{\gamma} = \alpha_0 + \alpha_1 \ln P_T + \alpha_2 (\ln P_T)^2, \tag{20}$$

where α_0, α_1, and α_2 depend on the distribution of P_T and the Lagrange multipliers of the optimization problem, with $\alpha_2 \geq 0$.

The full expressions for the parameters α_0, α_1, and α_2 can be found in Appendix B together with the proof of Proposition 1. The optimal contract in the standard principal agent model with effort aversion alone (i.e. without volatility incentive constraint (7)) has the same functional form (20) with $\alpha_2 = 0$. If $\alpha_2 = 0$ and $\gamma \geq 1$, the wage function is globally concave and therefore cannot explain option contracts.

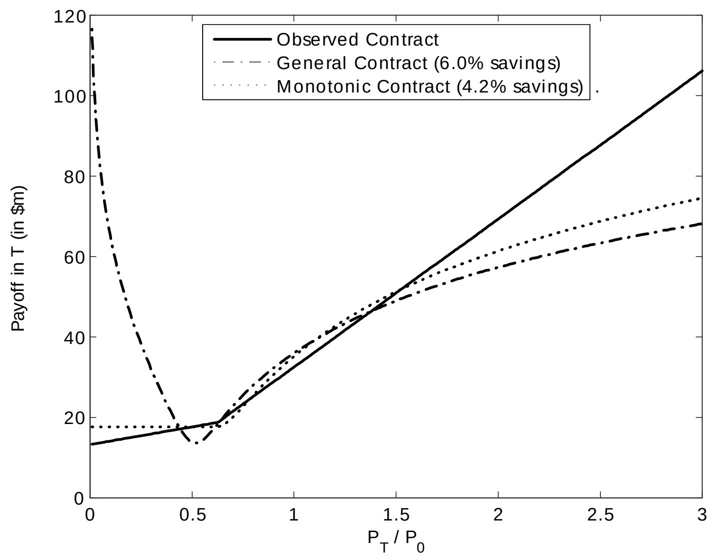

Figure 2: The figure shows end of period wealth W_T for the observed contract (solid line), the optimal general contract (dashed line) and the optimal monotonic contract (dotted line) for a representative CEO whose parameters are close to the median of the sample. The parameters are $\phi = \$1.1\text{m}$, $n_S = 0.26\%$, $n_O = 0.92\%$ for the observed contract. Initial non-firm wealth is $W_0 = \$9.0\text{m}$. $P_0 = \$3.1\text{bn}$, $\sigma = 18.4\%$, and $K/P_0 = 63\%$, $T = 6.3$ years, $r_f = 4.4\%$, $d = 1.7\%$. All calculations are for $\gamma = 3$.

In Appendix B, we also show that the optimal contract is non-monotonic as long as $\alpha_2 > 0$. The agent's final wealth W_T decreases for low stock prices, reaches a minimum, and then increases again.[21] Moreover, the function is convex for low stock prices and eventually becomes concave for high stock prices. Figure 2 depicts the optimal general contract (broken line) for the representative CEO (see Footnote 18 for the choice of the CEO and Section 2.3 for the calibration method). In addition, the figure shows the observed piecewise linear contract (solid line) and the optimal monotonic contract (dotted line) that we will introduce shortly.

The shape of the optimal general contract not only differs markedly from observed contracts, but it is also grossly counterintuitive. To understand this shape, we need to go back to the first-order approach which is violated by the general contract (20) (We show this more formally in Appendix A.). Our optimization routine ensures that the observed choice (e^d, σ^d) remains a *local* optimum in the manager's decision calculus, but it does not guarantee that it stays the *global* optimum. The new global optimum is obvious: choose effort e and volatility σ as low as possible, so that the end-of-

[21] For extremely low stock prices, W_T will exceed the firm value P_T which might not be possible under limited liability. We could add a condition $W_T \leq P_T$ to the optimization problem which would add a jag at the lowest stock prices in Figure 2. As everything else remains unchanged, we work with the simpler, more intuitive problem without this restriction. Note also that the firm could use an insurance to overcome the limited liability restriction.

period firm value P_T is as close to zero as possible. Hence, the "optimal" general contract provides the agent with incentives to destroy the firm.[22] Moreover when justifying our assumptions on $P(\sigma)$ at the beginning of Section 2, we assumed that, given the level of volatility, the manager always chooses a strategy that maximizes firm value. This assumption is also violated by the general contract (20).

We address these problems by restricting the contract to be monotonic, so that it cannot generate perverse incentives.[23] We therefore assume that the wage schedule $W_T(P_T)$ is differentiable almost everywhere and introduce the additional monotonicity constraint:[24]

$$\frac{dW_T(P_T)}{dP_T} \geq 0 \text{ for all } P_T \tag{21}$$

Proposition 2. ***(Optimal monotonic contract):*** *The optimal contract that solves the shareholders' problem (3), (4), (6), (7), and (21) has the following functional form:*

$$W_T^{\gamma} = \begin{cases} \alpha_0 + \alpha_1 \ln P_T + \alpha_2 (\ln P_T)^2 & \textit{if } ln(P_T) > -\frac{\alpha_1}{2\alpha_2} \\ \alpha_0 - \frac{\alpha_1^2}{4\alpha_2} & \textit{if } ln(P_T) \leq -\frac{\alpha_1}{2\alpha_2} \end{cases} \tag{22}$$

where α_0, α_1, and α_2 depend on the distribution of P_T and the Lagrange multipliers of the optimization problem, with $\alpha_2 > 0$.

The monotonic contract (22) is flat where the general contract is decreasing, while the two shapes are identical where the general contract is increasing. Even though the shape is identical, the parameters α_0, α_1, and α_2 will be different for the two contracts. As compensation is reduced in bad states of the world where the general contract is decreasing, compensation must be increased in good states of the world to satisfy the participation constraint. See Figure 2 for an illustration for the representative CEO. A comparison of Figure 2 with Figure 1 shows that the monotonic contract exhibits the same three features that are highlighted in Figure 1 for the piecewise linear contract. Relative to the observed contract, the optimal contract punishes very poor outcomes less, rewards

[22] Note that the general contract (20) is globally optimal if the expected stock price $E(P_T)$ is bounded sufficiently far away from zero no matter what action the manager chooses. Practically, this means that the manager cannot destroy the firm for certain, even though realizations of P_T close to zero remain possible. This assumption is defensible for some firms (e.g. utilities), but not for others (e.g. internet start-ups). Core and Qian (2002) and Flor, Frimor, and Munk (2006) also find U-shaped optimal contracts. For Core and Qian (2002) this is not a problem, because they only allow two effort and two volatility choices. Flor, Frimor, and Munk (2006) restrict the contract shape to be monotonic to avoid this problem.

[23] The obvious way to look at the second-order conditions in order to find assumptions under which the first-order approach holds is not feasible, because the second-order conditions are too complicated. See Appendix A for a further discussion of the validity of the first-order approach.

[24] The differentiability assumption simplifies the formulation of the restriction and the proof. It is not needed for the result in Proposition 2.

very good outcomes less, and provides more effort incentives in the center of the distribution where the optimal contract is steeper than the observed contract.

By construction, the monotonic contract (22) does not provide the perverse incentives to destroy the firm and therefore appears preferable to the general contract (20). Nevertheless, we still cannot guarantee that the first-order approach is valid here. For the proof of Proposition 2, we therefore still need the assumption that the first-order approach holds. In Appendix A, we numerically analyze the validity of the first-order approach and establish that for most CEOs in our sample there are reasonable assumptions so that the CEOs will not deviate from the observed choice of effort and volatility under the optimal monotonic contract.

5.2 Empirical results

We now turn to calibrating the optimal general contract (20) and the optimal monotonic contract (22). These contract shapes are defined by the three parameters α_0, α_1, and α_2. As we also have three constraints (15), (16), and (17), there are no degrees of freedom left to minimize costs. Therefore, we just solve a system of three equations in three unknowns for every CEO in our sample. The contract we obtain in this way is always cheaper than the observed contract, because the new contract has the optimal functional form.

[Insert Table 3 here.]

Table 3 shows our numerical results for the optimal general contract (Panel A) and the optimal monotonic contract (Panel B) for six values of the risk-aversion parameter γ. We do not tabulate the parameters α_0, α_1, and α_2, as they cannot be interpreted independently of each other. Instead, the table shows the manager's minimum wealth ($\min W_T^*(P_T)$), the location of the kink where the wage function changes from flat to increasing, the inflection point where the wage function changes from convex to concave, and the savings that can be generated by switching from the observed contract to the optimal contract. Our discussion of Table 3 concentrates on the optimal monotonic contract (Panel B) as the optimal general contract violates the first-order approach except under very strong assumptions (see Appendix A and Footnote 22).

Table 3, Panel B shows that for $\gamma = 3$, firms can save on average 13% (median: 9%) of the compensation costs if they replace the observed contract with the optimal monotonic contract. These average savings are considerably higher than the 4% savings for our representative CEO in Figure 2. This observation illustrates that calibrations to average or "typical" values can be misleading and that calibrations to individual data are necessary. Similarly to the piecewise linear contract described in Table 2, savings increase as the risk-aversion parameter γ increases in Table 3, because improved risk-sharing becomes more important if the manager is more risk-averse.

The location of the kink is on average 56% of the beginning-of-period stock price, i.e. the manager receives his minimum pay if the stock price drops by $1 - 56\% = 44\%$ or more. The average inflection point is 75% for $\gamma = 3$, so the wage scheme is convex for any stock price P_T below $0.75P_0$, and concave above this point. Even though the monotonic contract allows for a convex region, the contract is concave over the range of the most likely outcomes. 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. Consequently, the inflection point decreases and the concavity region increases as the CEOs' risk aversion increases.

The optimal monotonic contract provides complete downside protection. Table 3, Panel B displays descriptive statistics for the CEO's minimum wealth $\min(W_T^*)$ relative to the beginning-of-period wealth W_0 for this contract. It shows that none of the CEOs in our sample would have to invest part of their personal wealth in the firm as long as $\gamma \geq 1$. For $\gamma = 3$, the average minimum wealth is 2.2 times initial non-firm wealth W_0, while the median minimum wealth is 1.4 times initial wealth. Even after extremely poor performance, the typical CEO in our sample will increase his personal wealth by 40%. This result is in stark contrast to optimal contracts from more standard principal-agent models with effort incentives only (see Dittmann and Maug (2007)), where managers are required to put a large part of their wealth at risk. Intuitively, limiting the downside for bad outcomes provides better (i.e. cheaper) risk-taking incentives than rewarding good outcomes.

By construction, the savings from the optimal general contract (Table 3, Panel A) are higher than those from the optimal monotonic contract (Panel B). The average difference is 3.5% ($= 16.4\% - 12.9\%$) for $\gamma = 3$, so the average costs of restricting the optimal contract to be monotonic are small. A comparison of the two panels of Table 3 further yields that the location of the minimum payout and the inflection point are both higher for the monotonic contract. This contract features a larger convex region that is needed to provide the same incentives as the general contract. Also, the CEOs' minimum wealth is much higher for the monotonic contract than for the general contract, where 41% of the CEOs can lose some (but never more than 40%) of their wealth for $\gamma = 3$.

The optimal piecewise linear contracts analyzed in Section 3 are an approximation to the shape of the optimal monotonic contract. Optimal piecewise linear contracts feature no stock holdings and options that are deep in the money. This approximation generates (on average for $\gamma = 3$) savings of 5.3%, and it can be improved upon by allowing for additional option grants with different strike prices. The optimal contract will then feature no stock, a large long position in an option grant that is deep in the money and many smaller short positions in option grants with higher strike prices. In this way, a piecewise linear contract can approximate the concave shape of the optimal contract for medium and high stock prices. The upper bound for the savings that can be realized by such a piecewise linear contract is 12.9%, which are the savings of the optimal monotonic contract from Table 3, Panel B.

6 Robustness Checks

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 4 displays for these subsamples the average savings as a percentage of pay that firms could realize by switching to the optimal piecewise linear contract. The last line shows the p-value of the Wilcoxon test that average savings are identical in the first and the fifth quintile.

[Insert Table 4 here.]

The table shows that savings are considerably higher for younger and especially less wealthy CEOs. With constant relative risk-aversion, higher wealth implies lower absolute risk-aversion and consequently fewer gains from efficient risk-sharing. The table also demonstrates that smaller firms and those with poor past performance would benefit more from recontracting. Their CEOs typically have options that are less in-the-money or even out-of-the-money. Therefore, the payout pattern of their options differs more from that of their stock holdings than it does for more successful CEOs. In our model, savings are generated by replacing the portfolio of stock and options with an option grant that is "intermediate" in the sense that its strike price lies between the strike price of the original option and zero, which is the "strike price" of stock. The scope for these savings is larger, if stock and options in the observed contract differ more from one another, i.e. if the strike price of the original option is high. This suggests that our full sample results are biased downwards and that the average savings in the unbiased sample would be somewhat higher than the 5.3% shown in Table 2.

Wealth robustness check CEO wealth is not observable and we can therefore work only with a rough approximation. In order to see to what extent our results depend on our wealth estimates, we repeat our analysis after multiplying the wealth estimate of all CEOs by a factor M that ranges from 0.5 to 2. Table 5 displays the results for $\gamma = 3$.

[Insert Table 5 here.]

A comparison of Table 2, Panel A and Table 6 shows that an increase in wealth W_0 has a similar effect as a decrease in the risk aversion parameter γ. With constant relative risk-aversion, higher wealth implies lower absolute risk-aversion. This leads to more options, a higher strike price, and lower savings. In absolute terms, however, the variation of our results across different wealth multipliers M is small. We therefore conclude that the imprecision in our wealth estimates is unlikely to bias our results significantly. Our qualitative results are certainly not affected.

CEO preferences 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). In order to see whether our results are robust to alternative assumptions on CEO risk aversion, we repeat our analysis from Table 2 with constant absolute risk aversion (CARA), so that $V^{CARA}(W_T) = -\exp(-\rho W_T)$ replaces $V(W_T)$ in equation (2). To maintain comparability with our previous results, we calculate the coefficient of absolute risk aversion ρ from γ so that both utility functions exhibit the same risk-aversion at the expected end-of-period wealth, i.e. we set $\rho = \gamma/(W_0 + \pi_0)$, where π_0 is the market value of the manager's contract (i.e., the costs of the contract to the firm). Table 6 displays the results for six different values of γ.

[Insert Table 6 here.]

The results are quite similar to those for CRRA in Table 2, Panel A. With CARA preferences, the strike price is somewhat higher than with CRRA preferences: for $\gamma = 3$ the strike price averages 52.1% for CARA instead of 50.5% for CRRA. Savings from recontracting are higher for CARA than for CRRA for low values of risk-aversion ($\gamma < 3$) while the opposite holds for high values of risk-aversion ($\gamma > 3$). We conclude that our results continue to hold for CARA utility.

7 Conclusions

In this paper we analyze a principal-agent model in which the agent does not only exert effort but also determines the firm's strategy and thereby its stock return volatility. In this model, the choice of a more risky 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 increases expected firm value. The reason is that the first-best solution, where the optimal management strategy is chosen irrespective of its risk, is not achievable. In the second-best solution, the manager passes up some profitable but risky projects as these would reduce his utility, or he adopts some unprofitable but safe projects that increase his utility. If the firm's strategy is adjusted and becomes more risky in this second-best environment, more profitable and less unprofitable projects will be adopted and firm value increases. Therefore, it is not sufficient to only consider the direct impact of an increase in risk on a manager's compensation package (vega) to determine his attitude towards an increase in risk. The indirect effect via an increase in firm value and the manager's equity incentives (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.

When we look at piecewise linear contracts consisting of fixed salary, stock, and an option grant, we find optimal contracts that look very different from observed compensation practice. According to the model, managers should not receive any stock but instead in-the-money options and higher fixed salary. However, the savings generated by switching to this optimal contract are low and average only 5.3%. This suggests that observed compensation practice is close to the optimum and that a slight preference of shareholders for stock, for at-the-money options, or against an increase in base salary renders observed compensation practice efficient. One such effect included in our model in a robustness check is the extra tax that must be paid by the firm and the CEO if options are issued in the money. These tax penalties are prohibitive for most firms, i.e. they render the observed contract efficient if they are taken into account. But even in the absence of such taxes, the observed contract can easily be optimal if firms have a preference not to increase base salaries and are willing to forgo the 5.3% savings. In times of an increasingly hot public debate on executive compensation, such an upward restriction on base salaries appears plausible.[25]

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

A limitation of our main analysis is its restriction to a single option grant (with a single strike price). In order to understand optimal contracts with more than one option grant, we derive and estimate the general monotonic contract that is not restricted to be piecewise linear. Any piecewise linear contract with a given number of option grants will be an approximation to this general monotonic contract. We find that the optimal monotonic contract pays a flat wage for low outcomes and is increasing and eventually concave over medium and high outcomes. Therefore, it can be implemented by a high fixed salary (twice the observed salary for the median CEO), long option holdings with low, in-the-money strike price, and short option holdings with higher strike prices. Alternatively, it can be approximated by fixed salary and a linear bonus scheme with an upper bound on the bonus payout (see Healy, 1985). Such a contract would save up to 12.9% for the average firm.

Another 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 (2006), 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 (5) by the two first-order conditions (6) and (7). 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. The literature has developed a few approaches to check or ensure the validity of the first-order approach, but none of these existing solutions works for our model.[27] We therefore develop a new approach here.

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 a given alternative contract 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 for a particular contract if the agent shirks under this contract.

We first recognize that the agent's choice of effort e and volatility σ can be rephrased as a choice

[27] For the one-dimensional effort aversion problem, Rogerson (1985) and Jewitt (1988) derive a number of sufficient conditions for the first-order approach to hold. These results cannot readily be extended to the two-dimensional case where the agent also chooses volatility. Moreover, these conditions only work if the contract shape is not restricted, so they would be useful only for the general contract (20). Hellwig (2008) solves a principal-agent model with effort and volatility choice without resorting to the first-order approach by assuming that the agent is risk-neutral. This approach does not generalize to risk-averse agents, however, and we cannot use Hellwig's results, because risk-aversion is a central point in our argument. If the agent were risk neutral, any contract that provides the necessary incentives would be optimal, because principal and agent would always agree on the value of a contract. As a consequence, there would be infinitely many optimal contract shapes. For the one-dimensional effort aversion problem, Dittmann and Maug (2007) derive a sufficient condition that can be evaluated numerically. We have derived similar conditions for our model, but it turns out that they are nearly always violated. Lambert (1986) shows that the first-order approach holds in his stylized model with two feasible volatility levels and three feasible firm value outcomes. In our opinion, such a model is not rich enough to be meaningfully calibrated to observed data. Finally, Armstrong, Larcker, and Su (2007) work with strong assumptions on the cost function $C(e)$ and production function $P_0(e)$. They calibrate these functions to the data and calculate the optimal contract without using the first-order approach. We are skeptical that cost functions and production functions can be inferred from observable data, so we do not follow this approach.

of expected firm value P_0 and volatility σ, because $P_0 = P_0(e, \sigma)$ is monotonically increasing in e for every σ. The following proposition provides a condition when the agent does not deviate from the observed effort-volatility choice.

Proposition 3. ***(First-order approach):*** *Let W_T^d denote the observed contract that implements the observed volatility σ^d and the observed firm value P_0^d. Consider an alternative, cheaper contract W_T^* of a given shape that solves our optimization problem (i.e. it minimizes $E\left(W_T^*(P_T)\right)$ subject to the participation constraint (16) and the two incentive compatibility constraints (15) and (17)).*

1. *If*

$$E\left[V(W_T^*)|P_0', \sigma'\right] < E\left[V(W_T^d)|P_0^d, \sigma^d\right], \text{ for all feasible } \{P_0', \sigma'\} \text{ with } P_0' < P_0^d, \tag{23}$$

   there exist pairs $\{C(e), P_0(e, \sigma)\}$ of cost functions and production functions such that the agent does not choose a combination of effort and volatility that results in a lower expected firm value P_0 under the cheaper contract W_T^*.

2. *A choice $\{P_0', \sigma'\}$ with $P_0' < P_0^d$ is infeasible, if the agent prefers it to $\{P_0^d, \sigma^d\}$ under the observed contract. A necessary condition for $\{P_0', \sigma'\}$ being infeasible is*

$$E\left[V(W_T^d)|P_0', \sigma'\right] > E\left[V(W_T^d)|P_0^d, \sigma^d\right] \tag{24}$$

   for all or a subset of the pairs $\{C(e), P_0(e, \sigma)\}$ identified in part (1).

Part 1 of Proposition 3 provides a condition under which there are assumptions that guarantee that the first-order approach is not violated. If condition (23) holds at some point $\{P_0', \sigma'\}$, the agent's utility from money is higher at the observed choice $\{P_0^d, \sigma^d\}$ than at the alternative choice $\{P_0', \sigma'\}$. Therefore, the agent will not deviate from the observed choice, if the additional effort costs that $\{P_0^d, \sigma^d\}$ involves relative to $\{P_0', \sigma'\}$ do not reverse inequality (23). This is the case if the cost function is not too steep. Further below in this Appendix, we therefore check condition (23) numerically in order to establish whether the first-order approach is violated or (possibly) not.

Part 1 of Proposition 3 gives an indication what type of assumptions must be made to ensure the validity of the first-order approach. In addition to assumptions on cost function $C(e)$ and production function $P_0(e, \sigma)$, we need assumptions on the set of outcomes $\{P_0, \sigma\}$ that are feasible. Based on information contained in the observed contract, Part 2 of Proposition 3 identifies outcomes that are infeasible and must therefore lie outside the feasibility set. For example, the combination $\{P_0^d, 0\}$, where volatility is reduced to zero while the expected stock price stays constant, is obviously preferred by the agent to the observed combination $\{P_0^d, \sigma^d\}$. However, any combination $\{P_0^d, \sigma\}$ with $\sigma < \sigma^d$ must be infeasible, because the agent would have chosen it if it had been feasible. Part 2 of Proposition 3 extends this reasoning to all values $P_0' < P_0^d$. Our argument is reminiscent of the revealed preferences approach, although we do not infer utility but rather the feasibility of an action.

In order to check condition (23) numerically, we have to make an assumption as to what combinations $\{P_0', \sigma'\}$ are feasible. We assume that there is for each P_0' a lower bound $\sigma^{mf}(P_0')$ (the minimum feasible volatility), so that $\{P_0', \sigma'\}$ is feasible if and only if $\sigma' > \sigma^{mf}(P_0')$. Part 2 of Proposition 3 provides a sufficient condition for a combination $\{P_0', \sigma'\}$ to be infeasible. As the agent's utility decreases in volatility σ, we can derive a lower bound for the minimum feasible volatility $\sigma^{mf}(P_0')$ from this condition:

$$\underline{\sigma}^{mf}(P_0') = \min\left\{\sigma \geq 0 | E(V(W_T^d)|P_0', \sigma) \leq E(V(W_T^d)|P_0^d, \sigma^d)\right\}. \tag{25}$$

This lower bound rules out combinations $\{P_0', \sigma'\}$ which we know must be infeasible from the observed contract. Clearly, a point that lies just above this lower bound is not automatically feasible. If such a point were feasible, the observed contract would not be robust to slight perturbations, e.g., in the agent's cost function. While we cannot rule this out on a theoretical basis as our model does not allow for such complications, we feel that this is unlikely to occur in practice. We therefore set the minimum feasible volatility to

$$\sigma^{mf}(P_0') = w\underline{\sigma}^{mf}(P_0') + (1 - w)\sigma^d, \tag{26}$$

where $w \in [0, 1]$. We use five different weights w in our numerical analysis: $w = 1$, 0.95, 0.9, 0.75, and 0.5. $w = 1$ is the most conservative assumption that only rules out those combinations $\{P_0', \sigma'\}$ that are certainly infeasible. For lower values of w we make increasingly stronger assumptions that reduce the number of feasible combinations.

We also need an upper bound for the feasible volatility in our numerical analysis. We arbitrarily use $2\sigma^d$, i.e., we assume that the manager cannot increase the firm's stock return volatility by more than 100%. This assumption is innocuous, because violations of the first-order approach typically occur at low levels of volatility. For each CEO in our sample, we generate a grid with 2,500 points that is given by 50 equally spaced points in the interval $(0, P_0^d]$ for P_0' and 50 equally spaced points in the interval $(\sigma^{mf}(P_0'), 2\sigma^d]$ for σ', where $\sigma^{mf}(P_0')$ is the minimum feasible sigma from (26). We check condition (23) for each grid point and calculate the proportion of CEOs for whom condition (23) is violated for at least one grid point. Table 7 reports the "proportion with certain violation" for three values of γ (1, 3, and 5). For this proportion of the CEOs in our sample, the first-order approach is violated. For the remaining CEOs, there are cost and production functions such that the first-order approach is not violated.

[Insert Table 7 here.]

The table shows that for $\gamma = 3$ and $w = 1$, the first-order approach is certainly violated for 82.4% of the CEOs in our sample under the optimal piecewise-linear contract. For the monotonic and the general contracts, the respective probabilities are 92.3% and 99.6%. As argued above, $w = 1$ is a very conservative assumption, and for lower values of w the proportion of CEOs for which the first-order approach is certainly violated falls for the linear and the monotonic contracts (but not for the general contract). For $w = 0.75$, the first-order approach is certainly violated only for one third of the CEOs in our sample for the piecewise linear contract and for half of the CEOs for the monotonic contract. For $w = 0.5$, the proportion with certain violations drops to below 10% for these two contracts. A comparison across different values of risk-aversion shows that violations are less frequent if risk-aversion is low: for $w = 0.75$ and the monotonic contract, the proportion of CEOs with certain violation is only 12.1% for $\gamma = 1$ compared to 51.6% for $\gamma = 3$ and 63.9% for $\gamma = 5$. In contrast, the first-order approach is virtually always violated for the general contract, independently of the value of w.

Finally, we investigate whether the results reported in this paper change if we consider only those CEOs for whom the first-order approach is not certainly violated. We recalculate Tables 2 and 3 for two different assumptions on the minimum feasible volatility from equation (26) given by $w = 0.75$ and $w = 0.9$ (results not reported in the tables). We find lower savings but otherwise similar results. For $\gamma = 3$ and $w = 0.75$, average savings are 2.7% for the linear contract (compared to 5.3% in the full sample) and 10.4% for the monotonic contract (compared to 12.9% in the full sample). For $w = 0.9$, the savings are 1.5% and 8.9%, respectively. Hence, the optimal contract predicted by our model is "more likely" to violate the first-order approach if the predicted savings are high. This is intuitive, because an optimal contract with low savings by construction must be very similar to the observed contract. As we assume that the observed contract does not violate the first-order approach, this will also be true for the optimal contract.

To conclude, this appendix demonstrates that the first-order approach can be justified for the monotonic and the piecewise-linear contract, but not for the general contract. We show that the first-order approach is valid under realistic (although possibly strong) assumptions for most CEOs in our sample, and we show that our qualitative results are not affected if we exclude those CEOs for whom the first-order approach is violated.

Appendix B: Proofs

Proof of Proposition 1: The Lagrangian is

$$\begin{aligned} L = & \int_0^\infty \left[P_T - W_T\right] g(P_T|e,\sigma) dP_T + \lambda_{PC} \left(\int_0^\infty V(W_T, e) g(P_T|e,\sigma) dP_T - C(e) - \overline{U} \right) \\ & + \lambda_e \left(\int_0^\infty V(W_T) g_e(P_T|e,\sigma) dP_T - \frac{dC}{de} \right) + \lambda_\sigma \int_0^\infty V(W_T) g_\sigma(P_T|e,\sigma) dP_T, \end{aligned} \tag{27}$$

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

$$g(P_T|e,\sigma) = \frac{1}{P_T\sqrt{2\pi\sigma^2 T}} \exp[-\frac{(\ln P_T - \mu(e,\sigma))^2}{2\sigma^2 T}] \tag{28}$$

with

$$\mu(e,\sigma) = \ln P_0(e,\sigma) + (r_f - \sigma^2/2)T. \tag{29}$$

g_e and g_σ are the derivatives of $g(.)$ with respect to e and σ. We differentiate (27) with respect to W_T and set this derivative equal to zero:

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

Some rearranging yields:

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

For the log-normal distribution (28) we get:

$$\begin{aligned} g_e &= g \cdot \frac{\ln P_T - \mu(e,\sigma)}{\sigma^2 T} \cdot \mu_e(e,\sigma) \\ g_\sigma &= g \cdot \frac{[\ln P_T - \mu(e,\sigma)] \cdot \mu_\sigma(e,\sigma) \cdot \sigma^2 T + [\ln P_T - \mu(e,\sigma)]^2 \sigma T}{(\sigma^2 T)^2} - \frac{g}{\sigma} \\ &= g \cdot \frac{[\ln P_T - \mu] \cdot \mu_\sigma \cdot \sigma + [\ln P_T - \mu]^2}{\sigma^3 T} - \frac{g}{\sigma}. \end{aligned}$$

Substituting this into the first-order condition (30) yields (together with the assumption of constant relative risk aversion (2)):

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

From inspection, the optimal wage contract can be written as (20) with parameters α_0, α_1, and α_2:

$$\begin{aligned}
\alpha_0 &= \lambda_{PC} - \lambda_e \frac{\mu_e \cdot \mu}{\sigma^2 T} - \lambda_\sigma \left(\frac{\mu \cdot \mu_\sigma}{\sigma^2 T} - \frac{\mu^2}{\sigma^3 T} + \frac{1}{\sigma} \right), \\
\alpha_1 &= \lambda_e \frac{\mu_e}{\sigma^2 T} + \lambda_\sigma \left(\frac{\mu_\sigma}{\sigma^2 T} - \frac{2\mu}{\sigma^3 T} \right), \\
\alpha_2 &= \lambda_\sigma \frac{1}{\sigma^3 T} \geq 0.
\end{aligned}$$

Lemma 1. ***(Shape of the general contract):***

(1) The agent receives the lowest payout $\min\{W_T(P_T)\}$ *at the stock price* $\underline{P} = \exp\left\{-\frac{\alpha_1}{2\alpha_2}\right\}$*. The wage function decreases monotonically for* $P < \underline{P}$ *and increases monotonically for* $P > \underline{P}$*.*

(2) If $\gamma \geq 1$*, the wage function is concave for* $P > \exp\left\{1 - \frac{\alpha_1}{2\alpha_2}\right\}$*.*

Proof of Lemma 1: The first derivative of the wage function (20) with respect to the end-of-period stock price P_T is

$$\frac{dW_T}{dP_T} = \frac{1}{\gamma} W_T^{1-\gamma} \cdot 2\alpha_2 \cdot (\ln P_T + \frac{\alpha_1}{2\alpha_2}) \cdot \frac{1}{P_T}. \tag{31}$$

Therefore, the first derivative is zero if and only if $P_T = \exp(-\frac{\alpha_1}{2\alpha_2})$. The second derivative of the wage function W_T is

$$\frac{d^2 W_T}{dP_T^2} = \frac{1-\gamma}{\gamma^2} W_T^{1-2\gamma} \cdot \left[2\alpha_2 \cdot \left(\ln P_T + \frac{\alpha_1}{2\alpha_2} \right) \cdot \frac{1}{P_T} \right]^2 + \frac{1}{\gamma} W_T^{1-\gamma} \cdot 2\alpha_2 \cdot \frac{1}{P_T^2} \left(1 - \ln P_T - \frac{\alpha_1}{2\alpha_2} \right) \tag{32}$$

$$= \frac{1}{\gamma} W_T^{1-\gamma} \cdot 2\alpha_2 \cdot \frac{1}{P_T^2} \left[\frac{1-\gamma}{\gamma} 2\alpha_2 W_T^{-\gamma} \left(\ln P_T + \frac{\alpha_1}{2\alpha_2} \right)^2 + \left(1 - \ln P_T - \frac{\alpha_1}{2\alpha_2} \right) \right]. \tag{33}$$

For $P_T = \exp(-\frac{\alpha_1}{2\alpha_2})$, the second derivative is positive, which proves statement (1). If $\gamma \geq 1$ and $\ln P_T + \frac{\alpha_1}{2\alpha_2} > 1$, the second derivative (33) is negative, so the wage function is concave in this region, and this proves statement (2).

Proof of Proposition 2: Note that the monotonicity constraint (21) 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, (21) can be rewritten as

$$h'(h^{-1}(W_T)) \geq 0. \tag{34}$$

For every W_T, (21) provides one restriction, so the Lagrangian for the differentiation at W_T is:

$$
\begin{aligned}
L_{W_T} = &\int_0^\infty [P_T - W_T]\, g(P_T|e,\sigma)dP_T + \lambda_{PC}\left(\int_0^\infty V(W_T,e)g(P_T|e,\sigma)dP_T - C(e) - \overline{U}\right) \\
&+ \lambda_e \left(\int_0^\infty V(W_T)g_e(P_T|e,\sigma)dP_T - \frac{dC}{de}\right) + \lambda_\sigma \int_0^\infty V(W_T)g_\sigma(P_T|e,\sigma)dP_T \\
&+ \lambda_{W_T} h'(h^{-1}(W_T)).
\end{aligned}
$$

The first-order condition then is

$$
\begin{aligned}
g(P_T|e,\sigma) = \lambda_{PC} V_{W_T} g(P_T|e,\sigma) + \lambda_e V_{W_T} g_e(P_T|e,\sigma) + \lambda_\sigma V_{W_T} g_\sigma(P_T|e,\sigma) \\
+ \lambda_{W_T} \frac{h''(h^{-1}(W_T)}{h'(h^{-1}(W_T)}.
\end{aligned}
\tag{35}
$$

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 (as before). If the constraint (34) is binding, equation (35) defines the Lagrange multiplier λ_{W_T}, and the solution is determined by the binding monotonicity constraint. If (34) is not binding, λ_{W_T} is zero and the first-order condition (35) simplifies to the first-order condition from Proposition 1, that is equation (30). Consequently, the solution is the same as long as it is monotonically increasing, and flat otherwise.

Proof of Proposition 3, Part 1: Consider the observed contract W_T^d that implements the effort e^d and the observed volatility σ^d. Let $\{e', \sigma'\}$ be an alternative feasible choice for the manager that results in a lower stock price. Under the assumption that our model is correct and that the agent's observed choice $\{e^d, \sigma^d\}$ is indeed optimal, we have (from (5)):

$$
E\left[V(W_T^d)|e',\sigma'\right] - C(e') \le E\left[V(W_T^d)|e^d,\sigma^d\right] - C(e^d), \tag{36}
$$

The effort e is not observable, but given the function $P_0(e,\sigma)$, we can infer the effort from $P_0(e,\sigma)$ and σ. Hence, we can rewrite equation (36) as

$$
E\left[V(W_T^d)|P_0',\sigma'\right] - C(P_0',\sigma') \le E\left[V(W_T^d)|P_0^d,\sigma^d\right] - C(P_0^d,\sigma^d). \tag{37}
$$

Inequality (37) must hold for every feasible alternative choice $\{P_0', \sigma'\}$ with $P_0' < P_0^d$.

We now turn to one of the optimal contracts W_T^* from our calibrations. This can be any of the three optimal contracts that we analyze in the previous sections (piecewise linear, monotonic, or general contract). Consider a choice $\{P_0', \sigma'\}$ with $P_0' < P_0^d$ and

$$
E\left[V(W_T^*)|P_0',\sigma'\right] < E\left[V(W_T^d)|P_0^d,\sigma^d\right]. \tag{38}
$$

We need to distinguish two cases.

Case 1: $\{P_0', \sigma'\}$ is associated with a lower effort than $\{P_0^d, \sigma^d\}$, so that $C(P_0^d, \sigma^d) - C(P_0', \sigma') > 0$. We consider pairs of the cost function $C(.)$ and production function $P(.,.)$ such that

$$C(P_0^d, \sigma^d) - C(P_0', \sigma') < E\left[V(W_T^d)|P_0^d, \sigma^d\right] - E\left[V(W_T^*)|P_0', \sigma'\right]. \tag{39}$$

Rearranging yields

$$E\left[V(W_T^*)|P_0', \sigma'\right] - C(P_0', \sigma') < E\left[V(W_T^d)|P_0^d, \sigma^d\right] - C(P_0^d, \sigma^d) \tag{40}$$

$$= E\left[V(W_T^*)|P_0^d, \sigma^d\right] - C(P_0^d, \sigma^d), \tag{41}$$

where the second line follows from the fact that our calibrated contracts provide the manager with the same utility as the observed contract (see equation (16)). Hence, under these assumptions on $C(.)$ and $P_0(.,.)$, the agent will prefer $\{P_0^d, \sigma^d\}$ to $\{P_0', \sigma'\}$ and the first-order approach is (at this point) not violated.

Case 2: $\{P_0', \sigma'\}$ is associated with a weakly higher effort than $\{P_0^d, \sigma^d\}$, so that $C(P_0^d, \sigma^d) - C(P_0', \sigma') \leq 0$. Then (40) and (41) follow immediately from (38), so that the agent will prefer $\{P_0^d, \sigma^d\}$ to $\{P_0', \sigma'\}$ and the first-order approach is (at this point) not violated.

If the cost function $C(.)$ is constant (or increasing and convex but sufficiently close to being constant), equation (39) holds for all feasible combinations $\{P_0', \sigma'\}$. Hence, if $E\left[V(W_T^*)|P_0', \sigma'\right] < E\left[V(W_T^d)|P_0^d, \sigma^d\right]$ for all feasible $\{P_0', \sigma'\}$, there are assumptions on $C(.)$ and $P_0(.,.)$, such that the first-order approach is not violated.

Part 2: If (37) is violated for an alternative choice $\{P_0', \sigma'\}$ with $P_0' < P_0^d$, the agent would prefer it to $\{P_0^d, \sigma^d\}$. The fact that we observe $\{P_0^d, \sigma^d\}$.then indicates that $\{P_0', \sigma'\}$ is infeasible. If $\{P_0', \sigma'\}$ is associated with a lower effort than $\{P_0^d, \sigma^d\}$ (Case 1), then a sufficient condition for $\{P_0', \sigma'\}$ being infeasible is that

$$E\left[V(W_T^d)|P_0', \sigma'\right] > E\left[V(W_T^d)|P_0^d, \sigma^d\right]. \tag{42}$$

If $\{P_0', \sigma'\}$ is associated with a higher effort than $\{P_0^d, \sigma^d\}$ (Case 2), then (42) is a sufficient condition for $\{P_0', \sigma'\}$ being infeasible only if the cost function and production function are such that

$$C(P_0', \sigma') - C(P_0^d, \sigma^d) < E\left[V(W_T^d)|P_0', \sigma'\right] - E\left[V(W_T^d)|P_0^d, \sigma^d\right]. \tag{43}$$

Equations (42) and (43) then imply

$$E\left[V(W_T^d)|P_0', \sigma'\right] - C(P_0', \sigma') > E\left[V(W_T^d)|P_0^d, \sigma^d\right] - C(P_0^d, \sigma^d), \tag{44}$$

so that the agent would prefer $\{P_0', \sigma'\}$ to $\{P_0^d, \sigma^d\}$ under the observed contract, so $\{P_0', \sigma'\}$ must be infeasible. The additional restriction on the cost function and the production function in (43) holds if the cost function $C(.)$ is "flat enough" (i.e. increasing and convex but sufficiently close to being constant). This is the same requirement on $C(.)$ as in Part 1 of the proof.

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.

Armstrong, Christopher S., David F. Larcker, and Che-Lin Su, 2007, Stock options and chief executive officer compensation, discussion paper, Stanford 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, 2006, Stockholder and bondholder wealth effects and CEO incentive grants, discussion paper, University of Iowa.

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, 2009, Sticks or carrots: Optimal CEO compensation when managers are loss-averse, *Journal of Finance*, forthcoming.

Edmans, Alex, 2007, Inside debt, discussion paper, Wharton School.

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.

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.

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., 2008, A reconsideration of the Jensen-Meckling model of outside finance, forthcoming: *Journal of Financial Intermediation*.

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, 2006, CEO turnover and relative performance evaluation, Working paper, MIT Sloan School of Management.

Jewitt, Ian, 1988, Justifying the First-Order Approach to Principal-Agent Problems, *Econometrica* 56, 1177-1190.

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.

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).

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.

Prendergast, Canice, 2002, The tenuous trade-off between risk and incentives, *Journal of Political Economy* 110, 1071-1102.

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.

Rogerson, William P., 1985, The First-Order Approach to Principal-Agent Problems, *Econometrica* 53, 1357-1367.

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

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

Panel A: Data set with 737 U.S. CEOs

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Stock (%)	n_S	1.76%	4.85%	0.04%	0.31%	3.96%
Options (%)	n_O	1.40%	1.62%	0.15%	0.96%	3.19%
Base Salary ($m)	ϕ	1.60	4.29	0.50	1.07	2.43
Non-firm Wealth ($m)	W_0	64.9	671.5	2.3	11.1	64.1
Firm Value ($m)	P_0	9,347	23,296	366	2,418	19,614
Strike Price ($m)	K	6,929	20,209	236	1,556	12,853
Moneyness (%)	K/P_0	70.6%	21.1%	42.1%	71.8%	99.9%
Maturity (years)	T	5.2	1.6	3.6	5.0	6.6
Stock Volatility (%)	σ	30.3%	13.6%	16.5%	28.5%	45.8%
Dividend Rate (%)	d	1.37%	3.96%	0.00%	0.66%	3.38%
CEO Age (years)		55.9	6.8	47	56	64
Stock Return 2001-5 (%)		11.8%	15.5%	-5.7%	11.5%	28.8%

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: Optimal piecewise linear contracts

This table describes the optimal piecewise linear contract. 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. In addition, it shows the fraction of CEOs with non-positive salaries ($\phi^* \leq 0$), the fraction of CEOs with zero stock holdings ($n_S^* = 0$), and the fraction of CEOs with non-positive option holdings ($n_O^* \leq 0$). 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 last row shows the corresponding values of the observed contract. Panel A shows the results for six different values of the parameter of risk aversion γ. The number of observations varies across different values of γ because we exclude all CEO-γ-combinations for which the observed contract implies positive risk-taking incentives *RTI* from equation (13). Panel B displays the results for those 282 CEOs for whom our numerical routine converges for all γ between 1 and 8.

Panel A: Results for all 737 CEOs

Risk Aversion	Obs.	Con-verged	Base Salary ϕ^* ($m)			Stock n_S^*			Options n_O^*			Moneyness K^*/P_0		Savings	
			Mean	Median	Prop≤0	Mean	Median	Prop=0	Mean	Median	Prop≤0	Mean	Median	Mean	Median
0.5	537	120	1.44	1.22	13.3%	0.62%	0.00%	80.8%	2.39%	1.60%	5.83%	55.7%	57.5%	0.11%	0.04%
1.0	665	394	4.73	2.46	0.0%	0.02%	0.00%	98.5%	2.49%	1.55%	1.52%	54.4%	54.5%	0.41%	0.16%
2.0	720	613	7.90	3.21	0.0%	0.04%	0.00%	98.9%	2.51%	1.39%	0.82%	53.2%	54.2%	2.09%	0.86%
3.0	735	654	7.68	3.15	0.0%	0.10%	0.00%	98.8%	2.15%	1.33%	0.46%	50.5%	50.9%	5.32%	2.56%
5.0	735	604	6.56	2.81	0.0%	0.26%	0.00%	95.9%	1.60%	1.14%	0.00%	44.7%	44.8%	14.01%	9.31%
8.0	737	442	4.66	2.08	0.2%	0.28%	0.00%	88.0%	1.14%	0.86%	0.23%	39.8%	39.8%	26.46%	22.20%
Data	737	N/A	1.60	1.07	0.0%	1.76%	0.31%	1.2%	1.40%	0.96%	5.16%	70.6%	71.8%	N/A	N/A

Panel B: Results for 282 CEOs with numerical results for all levels of risk-aversion

Risk Aversion	Fixed Salary ϕ^* ($m)			Stock n_S^*			Options n_O^*			Moneyness K^*/P_0		Savings	
	Mean	Median	Prop≤0	Mean	Median	Prop=0	Mean	Median	Prop≤0	Mean	Median	Mean	Median
1.0	4.08	2.18	0.0%	0.00%	0.00%	100.0%	2.54%	1.86%	0.00%	54.1%	55.0%	0.48%	0.20%
2.0	3.85	2.11	0.0%	0.00%	0.00%	100.0%	2.38%	1.77%	0.00%	50.6%	51.1%	3.16%	1.61%
3.0	3.64	2.03	0.0%	0.00%	0.00%	100.0%	2.16%	1.65%	0.00%	47.0%	47.3%	8.12%	5.15%
5.0	3.27	1.83	0.0%	0.15%	0.00%	97.5%	1.60%	1.23%	0.00%	40.6%	41.1%	20.05%	17.37%
8.0	2.76	1.65	0.4%	0.23%	0.00%	87.2%	1.11%	0.86%	0.00%	34.4%	32.9%	33.97%	35.44%
Data	1.27	0.94	0.0%	0.70%	0.31%	1.8%	1.92%	1.47%	0.00%	68.2%	68.3%	N/A	N/A

Table 3: Optimal general and optimal monotonic contracts

This table describes the optimal general contract (from equation (20)) in Panel A and the optimal monotonic contract (from equation (22)) in Panel B for six different values of the risk-aversion parameter γ. The table displays the mean and median of the savings, the inflection point, the location of the minimum wage (in Panel A) and the location of the kink (in Panel B). 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 inflection point is the end-of-period stock price P_T where the wage scheme turns from convex to concave. The location of the minimum (Panel A) is the end-of-period stock price P_T where the agent receives the smallest wage. The location of the kink (Panel B) is the end-of-period stock price P_T where the wage schedule $W(P_T)$ starts to increase. The inflection point, the location of the minimum, and the location of the kink are expressed as percentage of the beginning-of-period stock price P_0. The table also shows descriptive statistics for the minimum payout $\min(W_T^*)$ scaled by the observed non-firm wealth W_0. The number of observations varies across different values of γ because we exclude all CEO-γ-combinations for which the observed contract implies positive risk-taking incentives RTI from equation (13).

Panel A: Optimal general contract

Risk Aversion	Obs.	Con-verged	Savings		Location of minimum		Inflection point		Minimum wealth $min(W_T^*)/W_0$				
			Mean	Median	Mean	Median	Mean	Median	Mean	Median	Min.	Max.	Freq. < 1
0.5	537	517	0.23%	0.05%	23%	23%	185%	153%	2.4	1.2	0.0	201.3	19%
1.0	665	632	1.63%	0.51%	35%	35%	96%	95%	2.6	1.2	0.4	228.3	22%
2.0	720	687	8.01%	4.48%	40%	40%	64%	63%	1.8	1.1	0.5	170.6	34%
3.0	735	557	16.44%	12.87%	40%	39%	55%	55%	1.4	1.1	0.6	84.8	41%
5.0	735	274	31.67%	30.18%	36%	34%	46%	45%	1.2	1.1	0.8	3.6	32%
8.0	737	178	47.99%	51.20%	29%	24%	37%	33%	1.1	1.1	0.9	3.6	25%

Panel B: Optimal monotonic contract

Risk Aversion	Obs.	Con-verged	Savings		Location of kink		Inflection point		Minimum wealth $min(W_T^*)/W_0$				
			Mean	Median	Mean	Median	Mean	Median	Mean	Median	Min.	Max.	Freq. < 1
0.5	537	335	0.2%	0.0%	27%	23%	205%	158%	3.0	1.3	0.0	227.8	11%
1.0	665	622	0.9%	0.3%	45%	43%	121%	116%	3.1	1.4	1.0	249.4	0%
2.0	720	711	5.4%	3.0%	55%	54%	85%	85%	2.9	1.4	1.0	242.1	0%
3.0	735	665	12.9%	9.2%	56%	57%	75%	76%	2.2	1.4	1.0	186.6	0%
5.0	735	523	30.0%	29.3%	50%	50%	61%	60%	1.5	1.3	1.0	25.5	0%
8.0	737	398	46.1%	47.9%	44%	43%	52%	51%	1.3	1.2	1.0	4.3	0%

Table 4: Savings from recontracting for subsamples

This table shows average savings 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. Savings are the difference in compensation costs between observed contract and optimal piecewise linear contract expressed as a percentage of costs of the observed contract, $(\pi_0^d - \pi_0^*)/\pi_0^d$. The last row shows the p-value of the two-sample Wilcoxon signed rank test that the average savings are 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	**Savings**	**Mean**	**Savings**	**Mean**	**Savings**	**Mean**	**Savings**
1	2.2	10.0%	46.2	7.3%	381	8.7%	-9.1%	9.7%
2	5.4	5.7%	51.5	5.3%	1,122	5.5%	4.9%	5.0%
3	10.3	4.6%	55.1	4.5%	2,462	4.3%	11.2%	4.0%
4	21.5	4.4%	57.9	5.7%	6,406	4.0%	17.2%	3.9%
5	246.3	2.0%	63.4	4.1%	33,935	4.1%	32.8%	4.0%
P-Value Q1-Q5		0.0000		0.0001		0.0001		0.0000

Table 5: Wealth robustness check

This table contains the results from repeating our analysis from Table 2 when we multiply our wealth estimates by the factor M, for $M \in \{0.5, 0.75, 1.5, 2.0\}$. The risk-aversion parameter γ is set equal to 3. 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. In addition, it shows the fraction of CEOs with non-positive salaries ($\phi^* \leq 0$), the fraction of CEOs with zero stock holdings ($n_S^* = 0$), and the fraction of CEOs with non-positive option holdings ($n_O^* \leq 0$). 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 the multiplier M because we exclude all CEO-M-combinations for which the observed contract implies positive risk-taking incentives *RTI* from equation (13).

Multipl. M	Obs.	Con-verged	Base Salary ϕ^* ($m)			Stock n_S^*			Options n_O^*			Moneyness K^*/P_0		Savings	
			Mean	Median	Prop≤0	Mean	Median	Prop=0	Mean	Median	Prop≤0	Mean	Median	Mean	Median
0.50	735	650	7.21	3.00	0.00%	0.18%	0.00%	96.3%	2.05%	1.25%	0.77%	47.9%	48.3%	8.43%	4.59%
0.75	735	650	7.98	3.09	0.00%	0.20%	0.00%	97.4%	2.14%	1.30%	0.77%	49.5%	49.9%	6.56%	3.37%
1.00	735	654	7.68	3.15	0.00%	0.10%	0.00%	98.8%	2.15%	1.33%	0.46%	50.5%	50.9%	5.32%	2.56%
1.50	729	625	7.30	3.21	0.00%	0.18%	0.00%	97.4%	2.13%	1.36%	0.64%	52.1%	53.0%	3.92%	1.81%
2.00	721	608	7.83	3.20	0.00%	0.09%	0.00%	98.2%	2.25%	1.41%	0.49%	52.5%	53.5%	3.22%	1.43%

Table 6: Piecewise linear contracts when CEOs have CARA utility

This table contains the results from repeating our analysis from Table 2 under the assumption that the CEO has CARA utility. For six different values of γ, we calculate the CEO's coefficient of absolute risk aversion ρ as $\rho = \gamma / (W_0 + \pi_0)$, 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 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. In addition, it shows the fraction of CEOs with non-positive salaries ($\phi^* \leq 0$), the fraction of CEOs with zero stock holdings ($n_S^* = 0$), and the fraction of CEOs with non-positive option holdings ($n_O^* \leq 0$). 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 positive risk-taking incentives *RTI* from equation (13).

Risk Aversion	Obs.	Con-verged	Base Salary ϕ^* ($m)			Stock n_S^*			Options n_O^*			Moneyness K^*/P_0		Savings	
			Mean	Median	Prop≤0	Mean	Median	Prop=0	Mean	Median	Prop≤0	Mean	Median	Mean	Median
0.5	626	422	5.46	2.71	0.00%	0.19%	0.00%	98.1%	2.31%	1.51%	1.90%	55.1%	56.3%	0.54%	0.17%
1.0	705	572	6.61	3.15	0.00%	0.02%	0.00%	98.8%	2.26%	1.40%	1.05%	55.6%	56.5%	1.15%	0.45%
2.0	733	640	6.98	3.20	0.00%	0.29%	0.00%	97.3%	2.13%	1.32%	2.03%	54.1%	55.2%	2.99%	1.34%
3.0	735	639	7.09	3.15	0.00%	0.15%	0.00%	98.1%	2.04%	1.24%	1.25%	52.1%	53.1%	5.34%	2.96%
5.0	735	627	7.06	3.04	0.00%	0.25%	0.00%	96.2%	1.79%	1.12%	0.80%	48.6%	49.3%	10.37%	6.92%
8.0	737	556	5.85	2.57	0.00%	0.42%	0.00%	93.4%	1.44%	0.89%	0.72%	45.3%	45.5%	18.56%	16.11%

Table 7: Validity of the first-order approach

This table displays for each of the three contracts (linear, monotonic, and general) and for three levels of risk aversion the proportion of CEOs for whom the first-order approach is certainly violated, i.e. for whom condition (23) is violated. This is done separately for five different assumptions (parameterized by w) regarding what levels of volatility are feasible. The minimum feasible volatility is given by equation (25). As w decreases from 1 to 0, the lower bound on volatility rises and our assumptions regarding the feasible levels of volatility become stronger. For the remaining CEOs (for whom condition (23) is not violated) there are assumptions on the cost function and the production function such that the CEOs do not deviate from the effort and investment choices implemented by the observed contract.

Lower bound on volatility (w)	Contract type	Proportion with certain violation		
		$\gamma = 1$	$\gamma = 3$	$\gamma = 5$
1.00	linear	76.4%	82.4%	83.1%
	monotonic	54.2%	92.3%	96.4%
	general	97.0%	99.6%	95.6%
0.95	linear	59.9%	72.5%	73.2%
	monotonic	47.4%	89.9%	95.4%
	general	97.0%	99.6%	95.6%
0.90	linear	50.3%	60.4%	63.7%
	monotonic	39.2%	86.3%	92.7%
	general	97.0%	99.6%	95.6%
0.75	linear	27.9%	34.6%	35.3%
	monotonic	12.1%	51.6%	63.9%
	general	97.0%	99.6%	95.6%
0.50	linear	12.9%	8.9%	7.6%
	monotonic	4.0%	6.5%	8.4%
	general	97.0%	99.6%	95.6%