Providing Effort and Risk-Taking Incentives: 
The Optimal Compensation Structure for CEOs*

Ingolf Dittmann† Ko-Chia Yu‡

October 18, 2008

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, calibrate it individually to the data of 737 U.S. CEOs, and show that it can explain
observed compensation practice surprisingly well. In particular, it justifies large option holdings
and high base salaries. Our analysis also 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 John Core, Alex Edmans, Rüdiger Fahlenbrach, Wayne Guay, Dirk Jenter, Ernst Maug, Oliver
Spalt, Jack Stecher, David Yermack, Dan Zhang, and seminar participants at NHH Bergen, the University of Mannheim,
the University of Zurich, and Vrije Universiteit Amsterdam. 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@few.eur.nl. Tel: +31 10 408 1283.
‡Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands. Email: yu@few.eur.nl. Tel:
+31 10 408 8947.
Providing Effort and Risk-Taking Incentives: The Optimal Compensation Structure for CEOs

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, calibrate it individually to the data of 737 U.S. CEOs, and show that it can explain observed compensation practice surprisingly well. In particular, it justifies large option holdings and high base salaries. Our analysis also 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
1 Introduction

The relationship between risk and incentives in executive compensation is an important topic 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.\footnote{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.} 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.\footnote{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) and Smith and Swan (2007) for corporate acquisitions.} 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 risk-taking incentives on purpose or whether risk-taking incentives are just an unintended side effect of effort incentives.

We investigate the importance of risk-taking incentives in executive compensation by modelling the endogeneity between risk and incentives. We develop a stylized principal-agent model from standard building blocks where an effort-averse agent chooses effort and stock price volatility and where volatility affects firm value. 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 generate for each CEO 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 practice 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.

The model is driven by the fact that most CEOs are poorly diversified because a large part of their wealth is linked to the company’s share price to provide effort incentives. CEOs therefore have incentives to take inefficiently low risk. They might, for instance, pass up a profitable, but very risky project, or they might hedge their firms’ risk at some cost. Shareholders can reduce this inefficiency
by providing risk-taking incentives, so that the CEO does not pass up the profitable risky project and does not hedge firm risk. 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). The best way to provide effort and risk-taking incentives therefore is to reward good outcomes and not to punish bad outcomes, i.e. the optimal contract 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 base salaries must rise as stock is replaced by less valuable options in order to guarantee the CEO’s reservation utility. The savings firms can expect when they 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 salary (see Dittmann, Maug and Spalt (2008)). 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. The savings from recontracting we find in our calibrations are much smaller than the additional tax penalties firms and executives would have to pay if in-the-money options were used. If we include these tax penalties in our model, all observed contracts are found to be optimal, so our model is consistent with compen-
sation 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. This approach also allows us to circumvent the endogeneity problem that would invariably arise in cross-sectional regressions that involve risk and incentives. Instead, 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 an increase in volatility raises the stock price. An increase in stock price then feeds through to managerial wealth via the manager’s incentive pay, i.e. delta.

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-taking incentives need to be provided. Our paper shows that options are indeed part of an optimal contract. Options can be detrimental to risk-taking incentives, but they 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. To our knowledge, this paper is the first, however, to calibrate

---

4Lambert (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
such a model and to test its quantitative implications. We also contribute to the recent literature on calibrations of contracting models by considering a richer and more realistic model than previous papers.\footnote{See Dittmann and Maug (2007), Gabaix and Landier (2008), Edmans, Gabaix, and Landier (2008), and Dittmann, Maug, and Spalt (2008).}

In the next section, we present our model in which the manager must choose effort and return volatility. 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 one option grant. In order to understand these results better, we then theoretically derive the unrestricted shape of the optimal contract in Section 4 and calibrate this shape to the data. Section 5 discusses and explores the validity of the first-order approach, and Section 6 concludes. All proofs and derivations can be found in the Appendix.

2 The model and its calibration

We develop a stylized principal-agent model that recognizes that a manager not only exerts effort in order to increase the expected value of the firm, but also makes choices that affect the stock return volatility. We base our analysis on the standard Holmström (1979) effort aversion model with lognormal stock prices and constant relative risk aversion and extend this model by assuming that the manager (agent) also chooses the volatility \( \sigma \) of the firm’s stock return \((P_T - P_0)/P_0\). Our main assumption is that the relationship between end-of-period stock price and the manager’s volatility choice is hump-shaped, or, more precisely, that \( E(P_T(\sigma)) \) is increasing and concave in \( \sigma \) up to some maximum point and decreasing thereafter. Intuitively, if volatility is inefficiently low and if the manager receives additional incentives to increase volatility, he will do so by adopting new value-increasing projects that have not been undertaken before because of their riskiness. He will also refrain from safe but value-destroying actions, such as costly hedging or diversifying mergers.

In the remainder of this introductory text, we substantiate this intuition and justify our assumption of a hump-shaped \( E(P_T(\sigma)) \). In a more detailed model, the manager would first choose an effort 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 (2005) 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.
level and then decide on the firm’s strategy, where a strategy consists of any feasible combination of many different actions that affect, for instance, project financing, mergers and acquisitions, the capital structure, or financial transactions. For our purposes, any strategy can be represented by a pair of stock return volatility $\sigma$ and expected end-of-period stock price $E(P_T)$, and we assume for tractability that all feasible strategies span a compact set in the $E(P_T)$-$\sigma$-space. If the manager prefers high stock prices over low stock prices for any volatility level $\sigma$ - which is always the case if the contract is monotonically increasing in the end-of-period stock price - then he will choose a strategy from the frontier $\mathcal{P}(\sigma) = \max(E(P_T)|\sigma)$ of the compact shape spanned by all strategies. This frontier naturally has a maximum $(\sigma^*, \mathcal{P}(\sigma^*))$ where the value-maximizing strategy is chosen, irrespective of stock return volatility. For levels of volatility below $\sigma^*$, firm value increases in volatility. Here, some value and risk increasing actions (e.g. an R&D project) are not taken while value-destroying but risk-decreasing actions (e.g. hedging) are adopted. On the other hand, firm value decreases in volatility for $\sigma > \sigma^*$, because increasingly risky but value-destroying actions are adopted while comparatively safe and value-increasing actions are forfeited. Our final assumption it that $\mathcal{P}(\sigma)$ is concave for $\sigma < \sigma^*$. For simplicity, we use $E(P_T(\sigma))$ instead of $\mathcal{P}(\sigma)$ to refer to the relationship between volatility and firm value from now on.

2.1 Model

Our model is in the spirit of Holmström (1979): At time 0 a risk-neutral principal (the shareholders) offers a contract to the risk-averse agent (manager). The manager signs the contract and chooses his actions, i.e. effort $e \in [0, \infty)$ and volatility $\sigma \in (0, \infty)$. At the end of the contracting period, at time $T$, the stock price $P_T(e, \sigma)$ is realized and the manager receives his wage $W_T(P_T)$. As the manager’s effort and volatility choice are not observable, the agent’s wage cannot directly depend on these quantities.\(^\text{6}\)

We use risk-neutral pricing and assume that the end-of-period stock price $P_T$ is lognormally distributed:

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

Here, $r_f$ is the risk-free rate, and $P_0(e, \sigma) = E(P_T)\exp\{-r_fT\}$ is the expected present value of

\(^{\text{6}}\)The ex-post volatility can clearly be estimated from stock returns, but these stock returns 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.
the end-of-period stock price $P_T (u, e, \sigma)$.\footnote{We work with risk-neutral pricing, 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 would value these shares possibly 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.} We assume that markets are informationally efficient, so $P_0 (e, \sigma)$ is equal to the stock price at the beginning of the contracting period. $P_0 (e, \sigma)$ is an increasing and concave function in effort $e$ and volatility $\sigma$, i.e. $\frac{dP_0 (e, \sigma)}{de} > 0$, $\frac{d^2 P_0 (e, \sigma)}{de^2} < 0$, and $\frac{d^2 P_0 (e, \sigma)}{d\sigma^2} < 0$. The last two assumptions are non-standard and discussed extensively at the beginning of Section 2.

The manager’s utility is additively separable in wealth and effort and has constant relative risk aversion with parameter $\gamma$ with respect to wealth:

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

(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 $\sigma$ affects the manager’s utility indirectly via the stock price distribution and the utility function $V (.)$. Finally, we assume that the manager has outside employment opportunities that give him expected utility $\overline{U}$. The shareholders’ optimization problem then is:

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

subject to $E [V (W_T (P_T)) | e, \sigma] - C (e) \geq \overline{U}$ 

and $\{e, \sigma\} \in \arg\max \{E [V (W_T (P_T)) | e, \sigma] - C (e)\}$

(5)

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

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

(6)

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

(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 Section 5. We call condition (6) effort incentive constraint and (7) volatility incentive constraint.
2.2 Measuring risk-taking incentives

In the empirical literature on executive compensation, risk-taking incentives are usually measured with 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 (see, among others, Guay (1999), Rajgopal and Shevlin (2002), Knopf, Nam and Thornton (2002), and Coles, Daniel and Naveen (2006)). An exception are Lambert, Larcker and Verrecchia (1991) who work with what we call the "utility adjusted vega", i.e. the partial derivative of the manager’s expected utility with respect to stock return volatility. However, there is another effect of volatility on managerial utility that - to the best of our knowledge - has been ignored in the empirical literature on risk-taking incentives: 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 constraint can be rewritten as

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

(8)

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

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

(9)

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

\[ PPS_{\text{ua}} \frac{dP_0}{d\sigma} = -\nu_{\text{ua}}, \]  

(10)

where \( PPS_{\text{ua}} := E \left[ \frac{dV(W_T) dW_T}{dP_T} dP_T \right] e, \sigma \) = \( E \left[ \frac{dV(W_T) dW_T P_T}{dP_T} \right] e, \sigma \) 

(11)

and \( \nu_{\text{ua}} := E \left[ \frac{dV(W_T) dW_T}{dP_T} P_T \left( -\sigma T + u \sqrt{T} \right) \right] e, \sigma \).  

(12)

Here, \( PPS_{\text{ua}} \) is the utility adjusted pay-for-performance sensitivity, or the utility adjusted delta, which measures by how much the manager’s expected utility increases for a marginal stock price increase. Likewise, \( \nu_{\text{ua}} \) 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: equation (13) implies that a manager with higher \( RTI \) will choose a higher volatility \( \sigma \), because \( P_0(\sigma) \) is concave. In the first-best solution, \( RTI = 0 \). 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. So while risk-taking incentives always increase in vega \( \nu^{ua} \), 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 \( \nu^{ua} \) and in the pay-for-performance sensitivity \( PPS^{ua} \).

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 sub-problem of finding a new contract with a given shape that achieves three objectives: (1) It provides the same effort and risk-taking incentives to the agent as the observed contract. (2) It provides the agent with the same utility as the observed contract. (3) It is as cheap for the firm as possible. This 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, the given level of effort 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 substantially differs from the observed contract, 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 do not consider the second stage of the Grossman and Hart (1983) procedure where the optimal effort level is found by searching across all pairs of effort and costs of implementing this effort. For this stage, estimates of the production function and the cost function are needed, which we cannot produce without extremely restrictive additional assumptions. This means that we cannot say anything about the optimal level of effort or volatility. Our method only analyzes the optimal structure of compensation.

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)):

$$PPS^{ua}(W_T(P_T), \epsilon, \sigma, \gamma) = E \left[ \frac{dV(W_T)}{dP_0} \right] \mid \epsilon, \sigma, \sigma' = \frac{dP_0}{d\epsilon}$$ (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:

$$PPS^{ua}(W^*_T(P_T), \epsilon, \sigma, \gamma) = PPS^{ua}(W^*_T(P_T), \epsilon, \sigma, \gamma)$$ (15)

Here $W^*_T$ denotes the new (cost minimizing) contract and $W^d_T$ denotes the observed contract ($d$ for "data").

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

$$E \left[ V(W^*_T(P_T)) \mid \epsilon, \sigma, \gamma \right] = E \left[ V(W^d_T(P_T)) \mid \epsilon, \sigma, \gamma \right], \quad (16)$$

$$RTI(W^*_T(P_T), \epsilon, \sigma, \gamma) = RTI(W^d_T(P_T), \epsilon, \sigma, \gamma). \quad (17)$$

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 4, we derive the optimal contract shape which depends on three parameters and we calibrate this optimal shape to the data. In the next section, we calibrate a piecewise linear contract that consists of fixed salary $\phi$, the number of shares
\( n_S \), and the number of options \( n_O \) with strike price \( K \):
\[
W_{T}^{\text{lin}}(P_{T}) = (W_{0} + \phi) \exp\{r_{f}T\} + n_{S}P_{T} + n_{O}\max\{P_{T} - K, 0\}.
\]
(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, \( E\left(W_{T}^{\text{lin}}(P_{T})|e^{d},\sigma^{d}\right) \), subject to the constraints (15), (16), and (17). We have four parameters to minimize costs over: \( \phi, n_{S}, n_{O}, \) and \( K \).

2.4 Data set

We use the ExecuComp database in order 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 \( \phi \) (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)).

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 in order 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. In this step, we lose five CEOs for whom we cannot numerically solve

---

8 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 (2008)), so that we overestimate risk-taking incentives in observed contracts.

9 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.
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 $\sigma$ from daily CRSP stock returns over the fiscal year 2006 and drop all firms with less 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 5-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 1-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 5 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.

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 5 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, that is older and richer CEOs 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 address this bias by analyzing subsamples with more successful and less successful CEOs separately.

The only parameter in our model that we cannot estimate from the data is the manager’s coefficient of relative risk aversion \( \gamma \). 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.

### 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 "optimal contract" and compare it with the observed contract.

More formally, we minimize \( E(W^{\text{fin}}_T(P_T)) \) 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 \( \phi \), but we require that \( n_O > -n_S \exp\{dT\} \) and \( \phi > -W_0 \) in order to prevent negative payouts. Negative option holdings or negative salaries are rarely seen in practice, but they are certainly possible. A negative salary would imply that the firm requires the CEO to invest this amount of his private wealth in firm equity. We argue that a good model should not assume but rather generate positive option holdings and positive salaries.

We 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.\(^{10} \)

\[ \text{[Insert Table 2 here]} \]

\(^{10}\)In many cases, the objective function in our problem is extremely 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 artefact. For our empirical analysis we always use the solution with the lowest costs.
Table 2 contains our calibration results for six values of the risk-aversion parameter $\gamma$, 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.\footnote{As long as the agent is risk-averse, our model predicts a 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 corroboration 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 a negative $RTI$.} The column Observations displays the remaining observations after CEOs with positive $v^{ua}$ 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 $\phi$, $n_S$, $n_O$, and $K$ of the calibrated optimal contract, and the percentage savings the firm could realize by switching from the observed contract $W^d_T$ to the optimal contract $W^*_T$, i.e.

$$savings = \left[ E \left( W^d_T(P_T) \right) - E \left( W^*_T(P_T) \right) \right] / E \left( W^d_T(P_T) \right).$$ (19)

Optimal contracts differ markedly from observed contracts, especially regarding the CEOs’ 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: For $\gamma = 3$, the median strike price is 51% of the share price compared to 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 (not shown in the table). For some CEOs, option holdings are lower in the optimal contract compared to the observed contract.

The general picture is that the stock holdings 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 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 (2008)).

The contracts described in Table 2 provide the same effort and risk-taking incentives as the observed contracts at a lower cost to the firm. The savings firms could realize are limited, though. For $\gamma = 3$, the median firm would just save 2.6% of its compensation costs (the average is 5.3%) by switching from the observed contract to the optimal contract. 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. Numerically, this means that the objective function is rather flat which makes it more difficult to obtain convergence (also see Footnote 10).

The small savings also imply that observed compensation practice is consistent with our model in the presence of a (possibly even small) effect in favor of shareholdings or at-the-money options that we did not account for in our model. The U.S. tax system strongly discriminates against in-the-money options (see Footnote 3). In particular, only $1m of the sum of base salary and in-the-money options is tax deductible at the firm level, and the executive must pay a 20% penalty tax (on top of all other taxes) on the intrinsic value of the option at the time the option becomes exercisable. To quantify these tax effects, we consider the representative CEO whose parameters are closest to the median values shown in Table 1.\textsuperscript{12} The observed contract of this representative CEO consists of a base salary of $1.1m, $7.9m in stock, and at-the-money options with a Black-Scholes value of $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.\textsuperscript{13} If we assume for simplicity that the vesting period is equal to the contracting period of 6.3 years and that the CEO exercises the options on the vesting day, the firm would have to pay corporate taxes on an additional $19.8m (= $3.9m + $17m - $1.1m) if they had implemented the optimal contract.\textsuperscript{14} With a tax rate of 35% for the firm, the firm would pay an additional $6.9m and the CEO in expectation an additional $3.4m ( = 20% · $17m). While the optimal contract is $0.2m cheaper than the observed contract, it leads to additional tax payments of $10.3m. In order to investigate this tax effect more systematically, we repeat our numerical analysis for \( \gamma = 3 \) with these two tax penalties. We assume that these taxes have to be paid if and only if the strike price is lower than the observed strike price, so we assume that all options in the observed contract were issued at-the-money. In this setting, the observed contract turns out to be optimal for all CEOs for whom our algorithm converges. Effectively, the U.S. tax system prohibits the use of in-the-money stock options.

We do not include taxes in our base model for three reasons. First, modelling taxes in a one-

\textsuperscript{12}For each parameter (observed salary \( \phi^d \), observed stock holdings \( n_S^d \), observed option holdings \( n_O^d \), wealth \( W_0 \), firm size \( P_0 \), stock return volatility \( \sigma \), 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.

\textsuperscript{13}It is interesting to note that the savings for the representative CEO are much lower than the median or mean savings. This observation demonstrates that a calibration to average or "typical" values can be misleading and that calibrations to individual data are necessary.

\textsuperscript{14}In our one-period model, we cannot allow for tax events or option exercises at a time other than the end of the period. Typical vesting periods are shorter than 6.3 years.
period world is only possible under quite restrictive additional assumptions on the vesting period and the manager’s exercise policy. Second, our numerical calibrations are only stable when we consider one option grant with one strike price, because it is difficult to distinguish between two option grants if the strike prices are close to each other. When modelling the tax penalty for in-the-money options, however, it is crucial to allow the firm to issue some options in the money and others at or out of the money. And third, realistic taxes would greatly increase the complexity of the model and would obscure the economic intuition.

The U.S. tax code can explain why we do not see any in-the-money options in the United States, but in-the-money options are also scarce in other countries. The reason could be that there are other costs to higher base salaries or in-the-money options that are not included in our model. For example, there might be a cost to increasing base salaries as this might be difficult to explain to shareholders and the general public. Alternatively, there might be concerns that executives try to influence the strike price of the option grants just as they appear to do via the so-called backdating of option grants. A commitment to using only at-the-money options could reduce this rent-seeking activity, and our analysis shows that the costs of this commitment are low.15

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 there is no scope to replace stock and at-the-money options for in-the-money options. 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 \( \gamma = 0.5 \), where objective functions are extremely flat due to the manager’s low risk-aversion. We therefore conclude that, within our model, in-the-money options are generally preferable to a portfolio of at-the-money options and stock.

[Insert Table 3 here.]

It is difficult to compare our calibration results across different values of \( \gamma \) in Table 2, because convergence rates vary considerably across \( \gamma \). In Table 3, we therefore reproduce Table 2 for those 282 CEOs for which our algorithm converges for all \( \gamma \geq 1 \). This table shows that, as \( \gamma \) increases, the optimal contract features fewer stock options, a lower strike price, and lower base salaries. So the contract becomes flatter and less convex as \( \gamma \) increases. Savings are considerable for high levels

15 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 needed not to be expensed prior to 2006.
of risk-aversion and negligible for $\gamma = 1$. This finding is not surprising as the savings stem from improved risk-sharing, which is more important if CEOs are more risk-averse.

[Insert Table 4 here.]

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 contract. The last line shows the p-value of the Wilcoxon test that average savings are identical in the first and the fifth quintile. 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 less gains from efficient risk-sharing. The table also demonstrates that smaller firms and firms with poor past performance would benefit much more from recontracting. Their CEOs typically have options that are less in-the-money or even out-of-the-money. So the payout pattern of their options differs more from the payout pattern of their stock holdings than it does for more successful CEOs. As a result there is more scope for savings from replacing this portfolio with in-the-money options only. These results suggest that our full sample results are biased downwards, so the average savings in the unbiased samples would be somewhat higher than the 5.3% shown in Table 2.

4 Optimal nonlinear contracts

A limitation of our analysis so far is that we look at stylized piecewise linear contracts with only one kink (i.e. only 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.

4.1 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:

\[ W_T^\gamma = \alpha_0 + \alpha_1 \ln P_T + \alpha_2 (\ln P_T)^2, \]  

where \( \alpha_0, \alpha_1, \text{and } \alpha_2 \) 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 \( \alpha_0, \alpha_1, \text{and } \alpha_2 \) can be found in the Appendix together with the proof of Proposition 1. Note that the quadratic term in (20) is the only difference to the problem without risk-taking incentives; in the standard principal agent model with effort aversion alone, we obtain the same functional form with \( \alpha_2 = 0 \). For \( \gamma \geq 1 \), the quadratic term is the only source of convexity in (20). With \( \alpha_2 = 0 \) and \( \gamma \geq 1 \), the wage function is globally concave and therefore cannot explain option contracts.

In the appendix, 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.\(^{16}\) Moreover, the function is convex for low stock prices and eventually becomes concave for high stock prices. Figure 1 further below depicts the optimal general contract for a representative CEO.

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). Our optimization routine ensures that the observed choice \((e^d, \sigma^d)\) 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

\(^{16}\)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 1. 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.
and volatility \( \sigma \) as low as possible, so that the end-of-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.\(^{17}\) 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.\(^{18}\) We therefore assume that the wage schedule \( W_T(P_T) \) is differentiable almost everywhere and introduce the additional monotonicity constraint:\(^{19}\)

\[
\frac{dW_T(P_T)}{dP_T} \geq 0 \text{ for all } P_T
\]  

(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:

\[
W_T^\gamma = \begin{cases} 
\alpha_0 + \alpha_1 \ln P_T + \alpha_2 (\ln P_T)^2 & \text{if } \ln(P_T) > -\frac{\alpha_1}{4\alpha_2} \\
\alpha_0 - \frac{\alpha_1^2}{4\alpha_2} & \text{if } \ln(P_T) \leq -\frac{\alpha_1}{4\alpha_2}
\end{cases}
\]  

(22)

where \( \alpha_0, \alpha_1, \text{and } \alpha_2 \) 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 \( \alpha_0, \alpha_1, \text{and } \alpha_2 \) 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 in order to satisfy the participation constraint.

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

\(^{17}\) 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 in order to avoid this problem.

\(^{18}\) 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. Another heuristic way to remove the perverse incentives to "destroy" the firm is to assume that 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). Under this assumption, the non-monotonic contract (20) is optimal.

\(^{19}\) The differentiability assumption simplifies the formulation of the restriction and the proof. It is not needed for the result in Proposition 2.
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. We will discuss this assumption and its validity in more detail in Section 5.

4.2 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 $\alpha_0$, $\alpha_1$, and $\alpha_2$. 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.

We start our discussion by considering the representative CEO whose parameters are closest to the median values shown in Table 1 (see Footnote 12). Figure 1 depicts three contracts for this representative CEO: the observed piecewise linear contract (solid line), the optimal general contract (broken line), and the optimal monotonic contract (dotted line). The horizontal axis shows the stock price at the end of the contracting period $P_T$ as a multiple of the beginning of period stock price $P_0$. The vertical axis displays the manager’s end-of-period wealth $W_T$ in million dollars. The optimal general contract pays out the smallest wage if $P_T$ is about 50% of $P_0$, and this payout is about the same as the observed fixed salary which is the minimal payout for the observed contract. For stock prices smaller than $0.5P_0$, the end of period wealth is decreasing in the stock price. If we impose monotonicity, the minimum payout increases and becomes much higher than the observed fixed salary. Also, the payouts for good outcomes where $P_T > 1.5P_0$, are higher for the monotonic contract than for the general contract. This insures that the agent receives the same expected utility from the two contracts.

Table 5 shows our full-sample results for the optimal general contract for six values of the risk-aversion parameter $\gamma$. We do not tabulate the parameters $\alpha_0$, $\alpha_1$, and $\alpha_2$, as they cannot be interpreted independently of each other. Instead, the table describes the manager’s minimum wealth ($\min W^*_\gamma(P_T)$) expressed as a percentage of initial non-firm wealth $W_0$, and the location of the minimum wage ($\arg \min W^*_\gamma(P_T)$) expressed as a percentage of the initial stock price $P_0$. In addition, the table shows mean and median of the savings (19) and the inflection point, i.e. the stock price $P_T$ at which the contract shape changes from convex to concave.

[Insert Table 5 here.]
Panel A of Table 5 shows that for $\gamma = 3$, firms can save on average 16% (median: 13%) of the compensation costs if they replace the observed contract with the optimal general contract. These average savings are considerably higher than the savings of 6% for our representative CEO in Figure 1 (see Footnote 13). Similarly to the piecewise linear contract described in Table 2, savings increase as the risk-aversion parameter $\gamma$ increases in Table 5, because improved risk-sharing becomes more important if the manager is more risk-averse.

The location of the minimum payout is on average 40%, i.e. the manager receives his minimum pay if the stock price dropped by $1 - 40\% = 60\%$. If the price drops further, pay increases again. The average inflection point is 55% for $\gamma = 3$. So the wage scheme is convex for any stock price $P_T$ below $0.55P_0$, and concave above this point. So even though the general contract allows for a convex region, the contract is concave over the range of the most likely outcomes. The inflection point decreases and the concavity region increases as the CEOs become more risk averse. This is intuitive as very risk-averse CEOs care little about high payouts in good states of the world.

Panel B of Table 5 describes the effect on the manager’s wealth of a switch from the observed contract to the optimal general contract. For $\gamma = 3$, the average minimum payout is 138% (median: 105%) of the CEO’s non-firm wealth $W_0$. In contrast to the piecewise linear contract, 41% of the
CEOs can lose some (but never more than 40%) of their wealth. So even though some of their wealth is at risk for some CEOs, these contracts feature considerable downside protection - especially in comparison to contracts from more standard principal-agent models with effort incentives only that put the manager’s entire wealth at risk.

Table 6 shows our results for the optimal monotonic contract. It is organized like Table 5. By construction, the savings from the monotonic contract are lower than the savings from the general contract. 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 Tables 4 and 5 further yields that the location of the minimum payout and the inflection point are both higher for the monotonic contract than for the general contract. So the monotonic contract will feature 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, and none of the CEOs is required to put any private wealth at risk (for $\gamma \geq 1$). On average for $\gamma = 3$, the CEOs’ minimum payout is more than twice their initial non-firm wealth.

The optimal piecewise linear contracts summarized in Table 2 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 5.5%, 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 6.

5 Discussion of the first-order approach

Like most of the literature on principal agent models, 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. Under the weak assumptions of our model,
this sufficient condition does not hold, and it is very well possible that the first-order approach is violated. In this section, we briefly review different approaches in the literature to ensure the validity of the first-order approach and argue that none of these approaches works for our model. We then develop a new approach that allows us to quantify the severity of this problem, i.e. the "likelihood" that the first-order approach is indeed violated. Finally, we numerically calculate this measure for each of the three contract types discussed in this paper (piecewise linear, monotonic, and general contract).

While many papers simply assume that the first-order approach is valid (see, among others, Hirshleifer and Suh (1992), Feltham and Wu (2001), or Flor, Frimor, and Munk (2006)), the literature has developed a number of ways to tackle this problem. We briefly discuss these approaches here:

- 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 of little use for the monotonic or the piecewise-linear contract.

- 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.

- Armstrong, Larcker, and Su (2007) consider the standard effort aversion model and make strong assumptions on the cost function \( C(e) \) and production function \( P(e) \). Then they calibrate these
functions to the data and calculate the optimal contract without using the first-order approach. We doubt that cost functions and production functions can be sensibly inferred from observable data, so we do not follow this approach.

As none of these approaches works for our model, we derive weaker conditions for whether or not the first-order approach is violated in the following proposition. Based on these results, we then develop a measure that indicates how likely a violation of the first-order approach is.

**Proposition 3. (First-order approach):** Let $W^d_T$ denote the observed contract that implements the observed volatility $\sigma^d$, the observed firm value $P^d$, and the unobservable effort $e^d$. Consider an alternative, cheaper contract $W^*_T$ that solves our calibration problem (i.e. it minimizes $E(W^*_T(P_T))$ subject to the participation constraint (16) and the two incentive compatibility constraints (15) and (17)). Further assume that the following regularity condition holds for both contracts $W^d_T$ and $W^*_T$: For every $\{e', \sigma'\}$ with $P' = P(e', \sigma') < P^d$ and $e' > e^d$ there exists a $\{e'', \sigma''\}$ with $P(e'', \sigma'') = P'$ and $e'' < e^d$ such that the agent prefers $\{e'', \sigma''\}$ over $\{e', \sigma'\}$.

(i) The first-order approach is violated if

$$E[V(W^*_T)|P', \sigma'] > E[V(W^d_T)|P^d, \sigma^d], \text{ for any feasible } \{P', \sigma'\} \text{ with } P' < P^d. \quad (23)$$

(ii) The first-order approach is valid if

$$E[V(W^*_T)|P', \sigma'] \leq E[V(W^d_T)|P', \sigma'], \text{ for all feasible } \{P', \sigma'\} \text{ with } P' < P^d. \quad (24)$$

Figure 2 provides a graphical representation of the results of Proposition 3. It depicts the expected utility from money $E(V(W_T))$ depending on today’s stock price $P_0$ for two contracts: the observed contract $W^d_T$ and the optimal contract $W^*_T$. All quantities in Figure 2 can be calculated if we know the risk-aversion parameter $\gamma$. For expositional reasons, we suppress the agent’s choice of volatility which would give rise to a third dimension as expected utility from money also depends on this volatility. Note that the graph does not show the agent’s expected utility, which is $E(V(W_T)) - C(e)$, because we do not know the cost function $C(e)$.

---

20 This regularity condition implies that a lower stock price $P' < P^d$ is always associated with a lower effort $e' < e^d$. Given that effort is costly but volatility is not, this assumption is likely to be satisfied by all well-behaved contracts and especially the observed contract $W^d_T$ that is monotonous. We discuss the impact of this assumption on our results in the next footnote.
The first-order approach is always violated if the broken line exceeds the dotted line for at least one \( P' < P^d \). On the other hand, the first-order approach is always valid if the broken line lies always below the solid line for all \( P' < P^d \). For intermediate cases (as shown in the Figure), the first-order approach is violated for some cost functions and valid for other cost functions.

The first part of Proposition 3 states that the first-order approach is violated if the optimal contract \( W^*_T \) exceeds the dotted line for any \( P' < P^d \). In that case, the agent receives a higher wage for a lower effort level, so that he will deviate from the observed effort level for sure. According to the second part of Proposition 3, the first-order approach is valid if the optimal contract \( W^*_T \) (i.e. the broken line) lies below the observed contract (the solid line) for all \( P' < P^d \). Then the optimal contract results in a lower utility than the observed contract for all effort-volatility combinations that result in a lower stock price. As the manager did not choose such a value-reducing action under the observed contract, he has even less reason to do so under the optimal contract. Note that we only consider alternative effort choices that result in lower stock prices, because stock price increases are desirable and need not be ruled out.

As shown in Figure 2, it is possible that the optimal contract fulfills neither of the two conditions from Proposition 3. In that case, the first-order approach is fulfilled for some cost functions \( C(e) \) and violated for others. Consider the effort level \( e' \) that results in the stock price \( P' = P(e') \). If the optimal contract lies below the point A, the first-order approach is valid (at this point) irrespective of the cost function. If the optimal contract lies above point C, the first-order approach is violated.
for all feasible cost functions. In the example shown in the figure, the optimal contract \( W_2^* \) (i.e. point B) lies between the points A and C, so that the first-order approach is violated for steeper cost functions and valid for flatter cost functions. If we work with the uninformative prior that cost functions are equally distributed in the sense that any point between A and C is equally likely, then the probability that the optimal contract \( W_2^* \) violates the first-order approach at \( P' \) is \( AB/AC \). We therefore define the probability of violation by

\[
\text{Prob}_V(P', \sigma') = \begin{cases} 
0 & \text{if } E[V(W_2^d)|P', \sigma'] \leq E\left[V(W_2^d)|P', \sigma'\right] \\
1 & \text{if } E[V(W_2^d)|P', \sigma'] > E\left[V(W_2^d)|P^d, \sigma^d\right] \\
Z & \text{otherwise,}
\end{cases}
\]

where

\[
Z = \frac{CE\left(E[V(W_2^d)|P', \sigma']\right) - CE\left(E\left[V(W_2^d)|P', \sigma'\right]\right)}{CE\left(E\left[V(W_2^d)|P^d, \sigma^d\right]\right) - CE\left(E\left[V(W_2^d)|P', \sigma'\right]\right)}.
\]

In order to arrive at numbers that can be readily interpreted, we transform the utilities in dollar values in (25) by calculating their certainty equivalents.

Before we can calculate this probability of violation, we have to determine which combinations \( \{P', \sigma'\} \) are feasible. For example, the agent will clearly prefer \( \{P^d, 0\} \) over \( \{P^d, \sigma^d\} \), but the combination \( \{P^d, 0\} \) is not feasible. Under the assumption that the agent behaves rationally, we know that any combination \( \{P', \sigma'\} \) with \( P' < P^d \) and \( E(V(W_2^d|P', \sigma')) \geq E(V(W_2^d|P^d, \sigma^d)) \) is not feasible. If it were feasible, the agent would have chosen it, because it leads to a higher reward for a lower effort level. We generate a grid with 2,500 points that is given by 50 equally spaced points in the interval \( (0, P^d) \) for \( P' \) and 50 equally spaced points in the interval \( (0, 2\sigma^d) \) for \( \sigma' \).\(^{21}\)

For our first feasibility set \( FS_1 \), we delete any point for which \( E(V(W_2^d|P', \sigma')) \geq E(V(W_2^d|P^d, \sigma^d)) \) and calculate the probability of violation \( \text{Prob}_V(P', \sigma') \) from equation (25) for the remaining points. We also consider two alternative feasibility sets based on the generalized feasibility condition

\[
\frac{CE\left(E\left[V(W_2^d)|P^d, \sigma^d\right]\right) - CE\left(E\left[V(W_2^d)|P', \sigma'\right]\right)}{CE\left(E\left[V(W_2^d)|P', \sigma'\right]\right)} > \delta.
\]

For our first feasibility set \( FS_1 \), the threshold \( \delta \) is equal to zero. For \( FS_2 \) we set \( \delta = 5\% \), and for \( FS_3 \) we set \( \delta = 10\% \). The first feasibility set \( FS_1 \) is very conservative as it assumes that a point is

\(^{21}\) We assume that the manager can reduce the firm’s stock return volatility to (nearly) zero, but that he cannot increase it by more than 100%. We need some upper bound for our approach and arbitrarily choose an increase of 100%. This assumption is innocuous, because violations of the first-order approach are more likely for low values of volatility.
feasible as soon as $E(V(W^q|P',\sigma'))$ is just below $E(V(W^q|P^d,\sigma^d))$. Note that these expected utilities do not include the costs of effort $C(e)$ and that the lower firm value $P'$ is associated with a lower effort level and, consequently, lower costs. Therefore, a point $\{P',\sigma'\}$ where (26) is positive but close to zero is likely to be infeasible, although it could be feasible for a sufficiently flat cost function. This reasoning is taken into account in the two less conservative feasibility sets $FS_2$ and $FS_3$.22

Table 7 shows descriptive statistics for the maximum probability of violation (25), where the maximum is calculated across all feasible grid points. The table displays the proportion of CEOs where the maximum violation is one, i.e. where the first-order approach (given the assumptions on feasibility) is violated for sure. In addition it shows the average maximum violation given that it is smaller than one, and the median maximum violation. We find that the general contract (20) always violates the first-order approach, irrespective of the feasibility set. These violations occur for low levels of firm value $P'$, i.e. for low effort. The reason is the U-shaped wage function that pays high wages for poor outcomes. For the linear contract (18) and the monotonic contract (22), the first order approach is also violated for most CEOs if the conservative feasibility set 1 is considered. In contrast to the general contract, these violations occur for comparatively high firm values $P'$ and low levels of volatility (not shown in the table). When we consider the less conservative feasibility sets 2 and 3, these violations are substantially lower. For the feasibility set 3, the median maximum probability of violation is 29% for the linear contract and 38% for the monotonic contract, and for only 9% (19%, respectively) of the CEOs the first-order approach is violated for sure. We also repeat the analysis shown in Table 7 for other values of risk-aversion and find similar results (not reported). Generally, the probability of violation increases with the risk-aversion parameter $\gamma$. We conclude that the first-order approach is a valid assumption for the piecewise linear and the monotonic contract for a large proportion of the CEOs in our data set.

6 Conclusions

In this paper we analyze a principal-agent model in which the agent does not only exert effort but also determines the stock return volatility. In this model, an increase in volatility has two effects on

22 If the regularity condition does not hold for the observed contract $W^q$, a combination $\{P',\sigma'\}$ might be feasible even though (26) with $\delta = 0$ holds. If the condition does not hold for the new candidate contract $W^\star$, the first-order approach might not be violated even though condition (23) from Proposition 3 (i) holds. In the latter case, our approach might identify too many violations.
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 an increase in volatility 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, however, 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 volatility is raised in this second-best environment, more profitable and less unprofitable projects will be adopted and firm value increases. Therefore, it is not enough to just look at the direct impact of an increase in stock price risk on a manager’s compensation package (vega) in order 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 one 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 a typical firm can realize from switching to this optimal contract are low and average only 5.3%. These low potential savings suggest 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 that we include in our model in a robustness check are the extra taxes the firm and the CEO have to pay if options are issued in the money. These tax penalties are prohibitive, 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.\footnote{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 how optimal contracts with more than one option grant look like, 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 approximated 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. Such a contract would save up to 12.9% for the average firm.

Like any other realistic model in the literature, our analysis uses the first-order approach (see Section 5 for a detailed discussion). We cannot prove in general that this approach is valid, so it might be the case that the optimal contract found by our numerical routine does not implement the observed effort and volatility choices but instead inferior choices. In Section 5, we construct a measure of the probability that the first-order approach holds and show that the general monotonic contract and especially the piecewise linear contract are unlikely to violate the first-order approach for most CEOs in our sample. Another indication that the first-order approach is not violated is the fact that we never find more than one local optimum in our calibrations when we repeat them for different start values. Finally, our results are still relevant if the first-order approach is violated, because they provide an upper bound on potential savings. Given that the potential savings are low on average (5.3% for the piecewise linear contract and 12.9% for the monotonic contract), firms might stick to the observed contracts as they do not want to run the risk that managers make inferior effort or volatility choices.

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%. 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)).

\[24\] Coughlan and Schmidt (1985), Kaplan (1994), and Jenter and Kanaan (2006), among others, analyze the sensitivity of dismissals to past stock price performance.
A Proofs

Proof of Proposition 1: The Lagrangian is

\[
L = \int_0^\infty [P_T - W_T] g(P_T|e,\sigma) dP_T + \lambda_{PC} \left( \int_0^\infty V(W_T,e) g(P_T|e,\sigma) dP_T - C(e) - \bar{U} \right) + \lambda_e \left( \int_0^\infty V(W_T) g_e(P_T|e,\sigma) dP_T - \frac{dC}{de} \right) + \lambda_\sigma \int_0^\infty V(W_T) g_\sigma(P_T|e,\sigma) dP_T,
\]

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

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

with

\[
\mu(e,\sigma) = \ln P_0 + (r_f - \sigma^2/2)T.
\]

\( g_e \) and \( g_\sigma \) are the derivatives of \( g(.) \) with respect to \( e \) and \( \sigma \). 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:

\[
\frac{1}{V_{W_T}(W_T)} = \lambda_{PC} + \lambda_e \frac{g_e}{g} + \lambda_\sigma \frac{g_\sigma}{g}.
\]

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

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

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

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

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 $\alpha_0$, $\alpha_1$, and $\alpha_2$ unrelated to $P_T$:

$$\alpha_0 = \lambda_P \gamma_C - \lambda_e \frac{\mu_e}{\sigma^2 T} \cdot \mu_t - \lambda_\sigma \left( \frac{\mu_\sigma}{\sigma^2 T} - \frac{\mu^2}{\sigma^3 T} + \frac{1}{\sigma} \right),$$

$$\alpha_1 = \lambda_e \frac{\mu_e}{\sigma^2 T} + \lambda_\sigma \left( \frac{\mu_\sigma}{\sigma^2 T} - \frac{2\mu}{\sigma^3 T} \right),$$

$$\alpha_2 = \lambda_\sigma \frac{1}{\sigma^3 T} \geq 0.$$ 

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

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

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

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

So 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

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

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

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

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

For every $W_T$, (21) provides one restriction, so the the Lagrangian for the differentiation at $W_T$ is:
While there is one multiplier \( \lambda_{W_T} \) for each value of \( W_T \), the other three multipliers \( \lambda_{PC}, \lambda_e, \) and \( \lambda_\sigma \) are the same across all values of \( W_T \) (as before). If the constraint (34) is binding, equation (35) defines the Lagrange multiplier \( \lambda_{W_T} \), and the solution is determined by the binding monotonicity constraint. If (34) is not binding, \( \lambda_{W_T} \) 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:** Consider the observed contract \( W^d_T \) that implements the effort \( e^d \) and the observed volatility \( \sigma^d \). 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)):

\[
E \left[ V(W^d_T) | e', \sigma' \right] - C(e') \leq E \left[ V(W^d_T) | e^d, \sigma^d \right] - C(e^d),
\]

(36)

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

\[
E \left[ V(W^d_T) | P', \sigma' \right] - C(P', \sigma') \leq E \left[ V(W^d_T) | P^d, \sigma^d \right] - C(P^d, \sigma^d).
\]

(37)

Rearranging (37) yields

\[
E \left[ V(W^d_T) | P', \sigma' \right] \leq E \left[ V(W^d_T) | P^d, \sigma^d \right] - C(P^d, \sigma^d) + C(P', \sigma').
\]

(38)

Inequality (38) must hold for every alternative choice \( \{P', \sigma'\} \) with \( P' < P^d \).

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). If this contract does not violate the first-order approach, it will implement the same managerial choice \( \{e^d, \sigma^d\} \) as the observed contract, i.e. the equivalent of (38) must hold for this optimal contract for all \( \{P', \sigma'\} \) with \( P' < P^d \):

\[
E \left[ V(W^*_T)|P', \sigma' \right] \leq E \left[ V(W^*_T)|P^d, \sigma^d \right] - C(P^d, \sigma^d) + C(P', \sigma') \tag{39}
\]

\[
= E \left[ V(W^*_T)|P^d, \sigma^d \right] - C(P^d, \sigma^d) + C(P', \sigma'), \tag{40}
\]

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)). According to equation (38), \( E \left[ V(W^*_T)|P^d, \sigma' \right] \) is a lower bound for the right-hand side of equation (40), so a sufficient condition for (40) is (24).

Analogous to (39), the first-order approach is violated if there is a \( P' < P^d \) and a \( \sigma' > 0 \) such that

\[
E \left[ V(W^*_T)|P', \sigma' \right] > E \left[ V(W^*_T)|P^d, \sigma^d \right] - C(P^d, \sigma^d) + C(P', \sigma') > E \left[ V(W^*_T)|P^d, \sigma^d \right]. \tag{41}
\]

The second inequality in (41) follows from the fact that the cost function \( C(e) \) and the production function \( P(e, \sigma) \) are both monotonically increasing in \( e \), so that \( C(P^d, \sigma^d) > C(P', \sigma') \). Part (ii) of Proposition 3 then follows from the fact that the new contract \( W^*_T \) provides the agent with the same utility as the observed contract \( W^d_T \) (see equation (16)).
References


[14] Dittmann, Ingolf, Ernst Maug, and Oliver Spalt, 2008, Sticks or carrots: Optimal CEO compensation when managers are loss-averse, ECGI discussion paper.


Table 1: Description of the dataset

This table displays mean, median, standard deviation, and the 10% and 90% quantile of 12 variables. 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 who are CEO in 2006 according to the ExecuComp database.

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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>10% Quantile</th>
<th>Median</th>
<th>90% Quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock (%) $n_S$</td>
<td>1.76%</td>
<td>4.85%</td>
<td>0.04%</td>
<td>0.31%</td>
<td>3.96%</td>
</tr>
<tr>
<td>Options (%) $n_O$</td>
<td>1.40%</td>
<td>1.62%</td>
<td>0.15%</td>
<td>0.96%</td>
<td>3.19%</td>
</tr>
<tr>
<td>Base Salary ($m) $\phi$</td>
<td>1.60</td>
<td>4.29</td>
<td>0.50</td>
<td>1.07</td>
<td>2.43</td>
</tr>
<tr>
<td>Non-firm Wealth ($m) $W_0$</td>
<td>64.85</td>
<td>671.49</td>
<td>2.25</td>
<td>11.11</td>
<td>64.14</td>
</tr>
<tr>
<td>Firm Value ($m) $P_0$</td>
<td>9,347</td>
<td>23,296</td>
<td>366</td>
<td>2,418</td>
<td>19,614</td>
</tr>
<tr>
<td>Strike Price ($m) $K$</td>
<td>6,929</td>
<td>20,209</td>
<td>236</td>
<td>1,556</td>
<td>12,853</td>
</tr>
<tr>
<td>Moneyness (%) $K/P_0$</td>
<td>70.56%</td>
<td>21.12%</td>
<td>42.07%</td>
<td>71.85%</td>
<td>99.93%</td>
</tr>
<tr>
<td>Maturity (years) $T$</td>
<td>5.16</td>
<td>1.64</td>
<td>3.59</td>
<td>4.96</td>
<td>6.57</td>
</tr>
<tr>
<td>Stock Volatility (%) $\sigma$</td>
<td>30.30%</td>
<td>13.63%</td>
<td>16.46%</td>
<td>28.48%</td>
<td>45.77%</td>
</tr>
<tr>
<td>Dividend Rate (%) $d$</td>
<td>1.37%</td>
<td>3.96%</td>
<td>0.00%</td>
<td>0.66%</td>
<td>3.38%</td>
</tr>
<tr>
<td>CEO Age (years)</td>
<td>55.93</td>
<td>6.84</td>
<td>47.00</td>
<td>56.00</td>
<td>64.00</td>
</tr>
<tr>
<td>Stock Return 2001-5 (%)</td>
<td>11.76%</td>
<td>15.55%</td>
<td>-5.74%</td>
<td>11.47%</td>
<td>28.76%</td>
</tr>
</tbody>
</table>

### Panel B: All 1,490 ExecuComp CEOs in 2006

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>10% Quantile</th>
<th>Median</th>
<th>90% Quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock (%) $n_S$</td>
<td>1.95%</td>
<td>6.26%</td>
<td>0.02%</td>
<td>0.28%</td>
<td>4.22%</td>
</tr>
<tr>
<td>Options (%) $n_O$</td>
<td>1.26%</td>
<td>1.57%</td>
<td>0.08%</td>
<td>0.79%</td>
<td>2.88%</td>
</tr>
<tr>
<td>Base Salary ($m) $\phi$</td>
<td>1.68</td>
<td>4.01</td>
<td>0.48</td>
<td>1.02</td>
<td>2.63</td>
</tr>
<tr>
<td>Firm Value ($m) $P_0$</td>
<td>8,840</td>
<td>24,760</td>
<td>339</td>
<td>2,091</td>
<td>17,796</td>
</tr>
<tr>
<td>CEO Age (years)</td>
<td>55.06</td>
<td>7.08</td>
<td>46.00</td>
<td>55.00</td>
<td>64.00</td>
</tr>
<tr>
<td>Stock Return 2001-5 (%)</td>
<td>10.51%</td>
<td>23.18%</td>
<td>-13.82%</td>
<td>9.83%</td>
<td>34.14%</td>
</tr>
</tbody>
</table>
Table 2: Optimal piecewise linear contracts

This table describes the optimal piecewise linear contract with one option grant. The table displays mean and median of the four contract parameters: base salary $\phi^*$, 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: $(\pi_0 - \pi_0^*)/\pi_0$. Results are shown for six different values of the parameter of risk aversion $\gamma$. The last row shows the corresponding values of the observed contract. The number of observations varies across different values of $\gamma$ because we exclude all CEO-$\gamma$-combinations for which the observed contract implies positive risk-taking incentives $RTI$ from equation (13).

<table>
<thead>
<tr>
<th>Risk Aversion</th>
<th>Obs. Converged</th>
<th>Base Salary $\phi^*$ (Sm) Mean Median Prop≤0</th>
<th>Stock $n_S^*$ Mean Median Prop=0</th>
<th>Options $n_O^*$ Mean Median Prop≤0</th>
<th>Moneyness $K^*/P_0$ Mean Median</th>
<th>Savings Mean Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>537 120</td>
<td>1.44 1.22 13.33%</td>
<td>0.62% 0.00% 80.83%</td>
<td>2.39% 1.60% 5.83%</td>
<td>55.7% 57.5%</td>
<td>0.11% 0.04%</td>
</tr>
<tr>
<td>1.0</td>
<td>665 394</td>
<td>4.73 2.46 0.00%</td>
<td>0.02% 0.00% 98.48%</td>
<td>2.49% 1.55% 1.52%</td>
<td>54.4% 54.5%</td>
<td>0.41% 0.16%</td>
</tr>
<tr>
<td>2.0</td>
<td>720 613</td>
<td>7.90 3.21 0.00%</td>
<td>0.04% 0.00% 98.86%</td>
<td>2.51% 1.39% 0.82%</td>
<td>53.2% 54.2%</td>
<td>2.09% 0.86%</td>
</tr>
<tr>
<td>3.0</td>
<td>735 654</td>
<td>7.68 3.15 0.00%</td>
<td>0.10% 0.00% 98.78%</td>
<td>2.15% 1.33% 0.46%</td>
<td>50.5% 50.9%</td>
<td>5.32% 2.56%</td>
</tr>
<tr>
<td>5.0</td>
<td>735 604</td>
<td>6.56 2.81 0.00%</td>
<td>0.26% 0.00% 95.86%</td>
<td>1.60% 1.14% 0.00%</td>
<td>44.7% 44.8%</td>
<td>14.01% 9.31%</td>
</tr>
<tr>
<td>8.0</td>
<td>737 442</td>
<td>4.66 2.08 0.23%</td>
<td>0.28% 0.00% 88.01%</td>
<td>1.14% 0.86% 0.23%</td>
<td>39.8% 39.8%</td>
<td>26.46% 22.20%</td>
</tr>
<tr>
<td>Data</td>
<td>737 N/A</td>
<td>1.60 1.07 0.00%</td>
<td>1.76% 0.31% 1.22%</td>
<td>1.40% 0.96% 5.16%</td>
<td>70.6% 71.8%</td>
<td>N/A N/A</td>
</tr>
</tbody>
</table>
Table 3: Comparative statics for optimal piecewise linear contracts

This table describes the optimal piecewise linear contract with one option grant for those 282 CEOs for whom our numerical routine converges for all \(\gamma\) between 1 and 8. The table displays mean and median of the four contract parameters: base salary \(\phi^*\), 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 expressed as a percentage of costs of the observed contract, \((\pi_0^d - \pi_0^d)/\pi_0^d\). Results are shown for five different values of the parameter of risk aversion \(\gamma\). The last row shows the corresponding values of the observed contract.

<table>
<thead>
<tr>
<th>Risk Aversion</th>
<th>Fixed Salary (\phi^*) ($m)</th>
<th>Stock (n_S^*)</th>
<th>Options (n_O^*)</th>
<th>Moneyness (K^*/P_0)</th>
<th>Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Median Prop(\leq 0)</td>
<td>Mean Median Prop=0</td>
<td>Mean Median Prop(\leq 0)</td>
<td>Mean Median Prop(\leq 0)</td>
<td>Mean Median</td>
</tr>
<tr>
<td>1.0</td>
<td>4.08 2.18 0.00%</td>
<td>0.00% 0.00% 100.00%</td>
<td>2.54% 1.86% 0.00%</td>
<td>54.1% 55.0% 0.00%</td>
<td>0.48% 0.20%</td>
</tr>
<tr>
<td>2.0</td>
<td>3.85 2.11 0.00%</td>
<td>0.00% 0.00% 100.00%</td>
<td>2.38% 1.77% 0.00%</td>
<td>50.6% 51.1% 0.00%</td>
<td>3.16% 1.61%</td>
</tr>
<tr>
<td>3.0</td>
<td>3.64 2.03 0.00%</td>
<td>0.00% 0.00% 100.00%</td>
<td>2.16% 1.65% 0.00%</td>
<td>47.0% 47.3% 0.00%</td>
<td>8.12% 5.15%</td>
</tr>
<tr>
<td>5.0</td>
<td>3.27 1.83 0.00%</td>
<td>0.15% 0.00% 97.52%</td>
<td>1.60% 1.23% 0.00%</td>
<td>40.6% 41.1% 0.00%</td>
<td>20.05% 17.37%</td>
</tr>
<tr>
<td>8.0</td>
<td>2.76 1.65 0.35%</td>
<td>0.23% 0.00% 87.23%</td>
<td>1.11% 0.86% 0.00%</td>
<td>34.4% 32.9% 0.00%</td>
<td>33.97% 35.44%</td>
</tr>
<tr>
<td>Data</td>
<td>1.27 0.94 0.00%</td>
<td>0.70% 0.31% 1.77%</td>
<td>1.92% 1.47% 0.00%</td>
<td>68.2% 68.3% 0.00%</td>
<td>N/A N/A</td>
</tr>
</tbody>
</table>


Table 4: Subsample analysis of savings from recontracting

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 \( \gamma \) is set equal to 3. 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 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.

<table>
<thead>
<tr>
<th>Quin-tile</th>
<th>Wealth ( W_0 ) (in Sm)</th>
<th>CEO Age</th>
<th>Firm Value ( P_0 ) (in Sm)</th>
<th>Stock return 2001-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Savings</td>
<td>Mean</td>
<td>Mean Savings</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td>2.2 10.0%</td>
<td>46.2 7.3%</td>
<td>381 8.7%</td>
<td>-9.1%</td>
</tr>
<tr>
<td>2</td>
<td>5.4 5.7%</td>
<td>51.5 5.3%</td>
<td>1,122 5.5%</td>
<td>4.9%</td>
</tr>
<tr>
<td>3</td>
<td>10.3 4.6%</td>
<td>55.1 4.5%</td>
<td>2,462 4.3%</td>
<td>11.2%</td>
</tr>
<tr>
<td>4</td>
<td>21.5 4.4%</td>
<td>57.9 5.7%</td>
<td>6,406 4.0%</td>
<td>17.2%</td>
</tr>
<tr>
<td>5</td>
<td>246.3 2.0%</td>
<td>63.4 4.1%</td>
<td>33,935 4.1%</td>
<td>32.8%</td>
</tr>
</tbody>
</table>

P-Value Q1-Q5 0.0000 0.0001 0.0001 0.0000
Table 5: Optimal general contracts

This table describes the optimal general contract (from equation (20)) for six different values of the risk-aversion parameter $\gamma$. Panel A displays the mean and median of the savings, the location of the minimum wage, and the inflection point. Savings are the difference in compensation costs between observed contract and optimal contract expressed as a percentage of the costs of the observed contract, $(\pi_0^d - \pi_0^*)/\pi_0^d$. The location of the minimum is the end-of-period stock price $P_T$ where the agent receives the smallest wage. The inflection point is the end-of-period stock price $P_T$ where the wage scheme turns from being convex to being concave. Both, the location of the minimum and the inflection point are expressed as percentage of the beginning-of-period stock price $P_0$. Panel B 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 $\gamma$ because we exclude all CEO-$\gamma$-combinations for which the observed contract implies positive risk-taking incentives $RTI$ from equation (13).

Panel A: Savings and contract shape

<table>
<thead>
<tr>
<th>Risk Aversion</th>
<th>Obs. Converged</th>
<th>Savings</th>
<th>Location of minimum</th>
<th>Inflection point</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>0.5</td>
<td>537</td>
<td>517</td>
<td>0.23%</td>
<td>0.05%</td>
</tr>
<tr>
<td>1.0</td>
<td>665</td>
<td>632</td>
<td>1.63%</td>
<td>0.51%</td>
</tr>
<tr>
<td>2.0</td>
<td>720</td>
<td>687</td>
<td>8.01%</td>
<td>4.48%</td>
</tr>
<tr>
<td>3.0</td>
<td>735</td>
<td>557</td>
<td>16.44%</td>
<td>12.87%</td>
</tr>
<tr>
<td>5.0</td>
<td>735</td>
<td>274</td>
<td>31.67%</td>
<td>30.18%</td>
</tr>
<tr>
<td>8.0</td>
<td>737</td>
<td>178</td>
<td>47.99%</td>
<td>51.20%</td>
</tr>
</tbody>
</table>

Panel B: Minimum wealth

<table>
<thead>
<tr>
<th>Risk Aversion</th>
<th>Minimum wealth divided by observed non-firm wealth $\min(W_T^*) / W_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>0.5</td>
<td>2.44</td>
</tr>
<tr>
<td>1.0</td>
<td>2.64</td>
</tr>
<tr>
<td>2.0</td>
<td>1.85</td>
</tr>
<tr>
<td>3.0</td>
<td>1.38</td>
</tr>
<tr>
<td>5.0</td>
<td>1.18</td>
</tr>
<tr>
<td>8.0</td>
<td>1.15</td>
</tr>
</tbody>
</table>
Table 6: Optimal monotonic contracts

This table describes the optimal monotonic contract (from equation (22)) for six different values of the risk-aversion parameter $\gamma$. Panel A displays the mean and median of the savings, the location of the kink, and the inflection point. 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 location of the kink is the end-of-period stock price $P_T$ where the wage schedule $W(P_T)$ starts to increase. The inflection point is the end-of-period stock price $P_T$ where the wage scheme turns from being convex to being concave. Both the location of the minimum and the inflection point are expressed as percentage of the beginning-of-period stock price $P_0$. Panel B 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 $\gamma$ because we exclude all CEO-$\gamma$-combinations for which the observed contract implies positive risk-taking incentives $RTI$ from equation (13).

### Panel A: Savings and contract shape

<table>
<thead>
<tr>
<th>Risk Aversion</th>
<th>Obs. Converged</th>
<th>Savings Mean</th>
<th>Savings Median</th>
<th>Location of kink Mean</th>
<th>Location of kink Median</th>
<th>Inflection point Mean</th>
<th>Inflection point Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>537 335</td>
<td>0.20%</td>
<td>0.03%</td>
<td>26.7%</td>
<td>22.5%</td>
<td>205.4%</td>
<td>157.5%</td>
</tr>
<tr>
<td>1.0</td>
<td>665 622</td>
<td>0.87%</td>
<td>0.28%</td>
<td>44.5%</td>
<td>42.8%</td>
<td>121.0%</td>
<td>116.3%</td>
</tr>
<tr>
<td>2.0</td>
<td>720 711</td>
<td>5.35%</td>
<td>3.03%</td>
<td>54.8%</td>
<td>54.4%</td>
<td>85.2%</td>
<td>85.2%</td>
</tr>
<tr>
<td>3.0</td>
<td>735 665</td>
<td>12.92%</td>
<td>9.24%</td>
<td>56.2%</td>
<td>56.9%</td>
<td>74.7%</td>
<td>75.5%</td>
</tr>
<tr>
<td>5.0</td>
<td>735 523</td>
<td>30.03%</td>
<td>29.29%</td>
<td>50.1%</td>
<td>50.1%</td>
<td>60.8%</td>
<td>60.2%</td>
</tr>
<tr>
<td>8.0</td>
<td>737 398</td>
<td>46.08%</td>
<td>47.88%</td>
<td>43.7%</td>
<td>43.5%</td>
<td>52.4%</td>
<td>50.9%</td>
</tr>
</tbody>
</table>

### Panel B: Minimum wealth

<table>
<thead>
<tr>
<th>Risk Aversion</th>
<th>Minimum wealth divided by observed non-firm wealth $\min(W_T)/W_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
</tr>
<tr>
<td>0.5</td>
<td>2.97</td>
</tr>
<tr>
<td>1.0</td>
<td>3.11</td>
</tr>
<tr>
<td>2.0</td>
<td>2.88</td>
</tr>
<tr>
<td>3.0</td>
<td>2.24</td>
</tr>
<tr>
<td>5.0</td>
<td>1.50</td>
</tr>
<tr>
<td>8.0</td>
<td>1.33</td>
</tr>
</tbody>
</table>
**Table 7: Probability that the first-order approach is violated**

This table shows descriptive statistics across CEOs for the maximum violation, $maxViol$. $MaxViol$ is the maximum of the probability of violation from equation (25) calculated across all feasible grid points. For each of the three feasibility sets, the table displays (1) the proportion of CEOs for which $maxViol=1$, (2) the mean of $maxViol$ when it is smaller than one, and (3) the median of $maxViol$. The grid points are given by 50 equally spaced points for $P'$ between 0 and the observed $P^d$ and 50 equally spaced points for $\sigma$ between 0 and $2\sigma$. These grid points are feasible if condition (26) holds for $\delta = 0$ (feasibility set 1), $\delta = 0.05$ (feasibility set 2), and $\delta = 0.1$ (feasibility set 3). All calculations are for $\gamma = 3$.

<table>
<thead>
<tr>
<th>Contract Type</th>
<th>Linear</th>
<th>Monotonic</th>
<th>General</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion with $maxViol=1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feasibility Set 1</td>
<td>94.5%</td>
<td>94.4%</td>
<td>100%</td>
</tr>
<tr>
<td>Feasibility Set 2</td>
<td>27.4%</td>
<td>48.4%</td>
<td>100%</td>
</tr>
<tr>
<td>Feasibility Set 3</td>
<td>8.7%</td>
<td>18.5%</td>
<td>100%</td>
</tr>
<tr>
<td>Mean $maxViol$ if $maxViol&lt;1$</td>
<td></td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Feasibility Set 1</td>
<td>35.0%</td>
<td>18.0%</td>
<td>N/A</td>
</tr>
<tr>
<td>Feasibility Set 2</td>
<td>37.5%</td>
<td>39.5%</td>
<td>N/A</td>
</tr>
<tr>
<td>Feasibility Set 3</td>
<td>31.8%</td>
<td>35.1%</td>
<td>N/A</td>
</tr>
<tr>
<td>Median $maxViol$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feasibility Set 1</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100%</td>
</tr>
<tr>
<td>Feasibility Set 2</td>
<td>49.2%</td>
<td>96.0%</td>
<td>100%</td>
</tr>
<tr>
<td>Feasibility Set 3</td>
<td>29.3%</td>
<td>38.2%</td>
<td>100%</td>
</tr>
</tbody>
</table>