Executive Stock Options when Managers Are Loss-Averse*

Ingolf Dittmann† Ernst Maug‡ Oliver Spalt§

September 4, 2006

Abstract

This paper analyzes optimal executive compensation contracts when managers are loss averse. We show that optimal contracts consist of an upward sloping compensation function and a threshold value for the value of the firm below which the manager is fired and suffers a discrete loss of compensation. We parameterize the model using data on compensation contracts and parameters for preferences suggested by the experimental literature. For a representative CEO, we estimate the optimal contract predicted by the model and discuss its comparative static properties. It turns out that the model's predictions are remarkably accurate and that it can explain the use of stock options as part of an optimal contract. The approximation is about one order of magnitude better than those from reasonable parameterizations of the conventional principal agent model with constant relative risk aversion.

JEL Classifications: G30, M52

Keywords: Stock Options, Executive Compensation, Loss Aversion

---

*We are grateful to the collaborative research center 649 on "Economic Risk" for financial support.

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

‡Corresponding author. Universität Mannheim, D-61381 Mannheim, Germany. Email: maug@bwl.uni-mannheim.de, Tel: +49 621 181 1952.

§Universität Mannheim, D-61381 Mannheim, Germany. Email: spalt@bwl.uni-mannheim.de, Tel: +49 621 181 1973.

1 Introduction

In this paper we analyze executive compensation contracts using a simple contracting model where the manager is loss averse in order to explain salient features of observed compensation contracts. We parameterize this model using standard assumptions and then compare the contracts generated by the model with those actually observed. Our main conclusion is that a standard principal agent-model with loss-averse agents can explain the prevalence of stock options far better than the standard model based on expected utility theory and constant relative risk aversion.

The theoretical literature on executive compensation contracts is based almost exclusively on contracting models where shareholders are risk-neutral and where the manager (agent) is risk averse, which is modelled with a concave utility function in a von Neumann-Morgenstern framework. Some highly stylized models can explain option-type features, but quantitative approaches rely more or less entirely on a standard model with constant relative risk aversion, lognormally distributed stock prices, and effort aversion.[1] However, Dittmann and Maug (2006) show that the standard CRRA-lognormal model cannot explain observed compensation practice. In particular, they find that the optimal contract almost never contains any options, and that options are generally more costly in providing incentives to managers than shares. They also reject the conventional explanation that options provide risk-taking incentives to CEOs.

In this paper we suggest a different approach by assuming that managers' preferences exhibit the features proposed by Kahneman and Tversky (1979) and Tversky and Kahneman (1991, 1992). On the basis of experimental evidence they argue that choices under risk exhibit three features: (i) reference dependence, where agents do not value their final wealth levels, but compare outcomes relative to some benchmark or reference level; (ii) loss aversion, which adds the notion that losses (measured relative to the reference level) loom larger than gains; (iii) diminishing sensitivity, so that individuals become progressively less sensitive to incremental gains and, respectively, incremental losses. For brevity, we will refer to all three features as loss aversion. These assumptions accord with a large body

[1] A model that can explain the use of options is Feltham and Wu (2001) who assume that the effort of the agent affects the risk of the firm, and Oyer (2004), who models options as a device to retain employees when recontracting is expensive. In his model, options do not provide incentives. The applications by Haubrich (1994), Haubrich and Popova (1998), and by Margiotta and Miller (2000) use constant absolute risk aversion when calibrating a principal-agent model. Calibration exercises with CRRA preferences and lognormal distributed stock prices include Lambert, Larcker, and Verrecchia (1991), Hall and Murphy (2000), (2002), Hall and Knox (2002), and Lambert and Larcker (2004).

of experimental literature which shows that the standard expected utility paradigm based on maximizing concave utility functions cannot explain a number of prominent patterns of behavior.[2] However, we do not use the notion of decision weights, so our model does not apply all elements of prospect theory. Given our results, this additional element does not seem to be needed.

The main drawback of expected utility approaches to explaining the prevalent use of stock options in compensation contracts is the fact that risk averse managers gain little utility from payoffs when the value of the firm is high. Whenever firm value is high, managers become wealthier and their marginal utility becomes small. This blunts any instrument for providing incentives that pays off only when firm value is high. Contracts that rely less on rewards for good outcomes ("carrots") and more on penalties for bad outcomes ("sticks") are more beneficial as they provide similar incentives at a lower cost. However, these predictions are at odds with observed compensation practice. By comparison, loss aversion implies that managers are more averse to losses than they are attracted by gains, so they demand a particularly high risk premium for being exposed to losses. Shareholders will therefore offer a contract that pays at least the reference wage most of the time in order to avoid this risk premium. As the marginal utility for payouts that exceed the reference level is much higher in the loss aversion framework, the optimal contract will be convex over some region, so loss aversion is potentially able to explain the use of contracts that provide for the "carrots" we observe in practice.

We develop this argument in two steps. The first step provides a standard analytic derivation of the optimal contract, following traditional approaches in principal agent theory. Here we characterize the optimal contract in a general setup and show that under standard assumptions the optimal contract features two parts: above a certain critical stock price the optimal contract always pays off high enough so that the manager perceives the payoff as a gain. In this region, the contract is continuous and monotonically increasing. However, below this region the contract drops discontinuously to a smaller wage that represents some lower bound on the manager's compensation. Hence, below some cut-off

[2] Experimental support for loss aversion is provided by Thaler (1980), Kahneman and Tversky (1984), Knetsch and Sinden (1984), Knetsch (1989), Dunn (1996) , Camerer, Babcock, Loewenstein, and Thaler (1997). This list is not exhaustive. Recently Rabin (2000) has demonstrated that concave utlity functions cannot account for risk-aversion over small stakes-gambles, a feature readily explained by loss aversion. Myagkov and Plott (1997) show document that the risk-seeking implied by prospect theory diminishes with experience, a result also supported by List (2004). Plott and Zeiler (2005) call into question the general interpretation of gaps between the willingness to pay and the willingness to accept as evidence for loss aversion.

the contract is flat and unrelated to the value of the firm. We suggest that the optimal contract is best interpreted as consisting of an arrangement where the manager is fired if the value of the firm falls below some threshold, and obtains a compensation contract that provides positive rewards relative to her reference wage as long as her employment lasts.

In the second step we parameterize the model using assumptions that are based on data and on the experimental evidence. We then compute the optimal contract for a representative CEO for a range of plausible parameter values and compare the optimal contract with the contract actually observed. We then show that the contract predicted by the model is reasonably close to the observed contract in two ways. Firstly, salient features of the contract (the slope for low stock prices, the slope for high stock prices, and the likelihood of CEO dismissal) can be matched to the data. Secondly, the implied costs of the contract to shareholders are similar to those of the observed contract (within about 5% - 7% of the actual costs), which implies that the model cannot find a much better contract than that observed in practice. The overall approximation is about one order of magnitude better compared to the approximation of a conventional CRRA-lognormal model and surprisingly good for a simple static model with only a small number of parameters.

Many authors apply loss aversion successfully to other questions in finance. Benartzi and Thaler (1995, 1999) develop the notion of myopic loss aversion and use it to explain the equity-premium puzzle. Barberis and Huang (2001) and Barberis, Huang and Santos (2001) apply loss aversion to the explanation of the value premium. Haigh and List (2005) find that CBOT-traders are loss averse and more so than inexperienced students, contradicting the effect List (2004) found earlier for consumers. Coval and Shumway (2005) support the same conclusion in their study of intraday risk-taking of CBOT-traders. Kouwenberg and Ziemba (2004) demonstrate theoretically that hedge-fund managers take more risk if their incentive fees become more substantial, an effect that contrasts the implications of a model based on hyberbolic absolute risk aversion (HARA). Their emirical results tend to support the prospect model. Ljungqvist and Wilhelm (2005) base their measure of issuer satisfaction in initial public offerings on loss aversion. The only application that fails to support loss aversion to the best of our knowledge is Massa and Simonov (2005) in their study of individual investor behavior. Despite the usefulness of loss aversion to analyze risk taking incentives in many areas of finance, the only paper so far that rigorously applies loss aversion to principal-agent theory is de Meza and Webb (2005). However, they explore a different specification and focus on endogenous reference levels.

They do not apply their argument to executive compensation contracts. To the best of our knowledge, ours is the first paper that demonstrates the potential of loss aversion to explain observed compensation practice.

In the following Section 2 we develop the model and discuss the main assumptions. In Section 3 we characterize the optimal contract analytically. We then develop our comparative static analysis by parameterizing the model and calibrating it to observed contracts in Section 4. Section 5 concludes. We gather all proofs and technical results in the appendix.

2 The Model

We consider a standard principal-agent model where shareholders (the principal) make a take-it-or-leave-it offer to a CEO (the agent) who then provides effort that enhances the value of the firm. Shareholders can only observe the stock market value of the firm but not the CEO's effort (hidden action).

Contracts and technology. The contract is negotiated at time t. At the end of the contracting period, T, the value of the firm P_T is commonly observed. P_T depends on the CEO's effort e, which is either high or low, $e \in \{\underline{e}, \overline{e}\}$, and on a state of nature so that P_T is distributed with density $f(P_T|e)$. For notational convenience we write $\Delta e = \overline{e} - \underline{e}$, $\Delta C = C(\overline{e}) - C(\underline{e})$, and $\Delta f(P_T|e) = f(P_T|\overline{e}) - f(P_T|\underline{e})$. The monotone likelihood ratio property (MLRP) holds for f, so $\Delta f(P_T|e)/f(P_T|\overline{e})$ is monotonically increasing in P_T. It is always optimal for shareholders to have the CEO implement the higher level of effort.

Preferences and outside options. Shareholders are assumed to be risk-neutral, whereas the manager is loss-averse (Kahneman and Tversky, 1992). We denote the manager's wage by w. The manager treats her income from the firm separately from other sources of income, and her reference income is w^R.[3] The preferences of the manager are separable in income and effort and can be represented by a function $V(w) - C(e)$, where V represents

[3] The literature refers to this phenomenon as "framing" or "mental accounting."This concept was present already in the earlier papers by Kahneman and Tversky. See Thaler (1999) for a survey of the evidence on mental accounting. The conventional modeling framework in the compensation literature implicitly makes the same assumption, usually for tractability to avoid modeling other sources of uncertainty on the agent's overall wealth.

loss-averse preferences over wage income:

$$V(w) = \begin{cases} U_g\left(w - w^R\right) & if \quad w > w^R \\ -U_l\left(w^R - w\right) & if \quad w \leq w^R \end{cases} . \tag{1}$$

Here U_g represents preferences in the gain space and U_l represents preferences in the loss space. We assume that U_g and U_l are both increasing, positive, and concave in $\left|w - w^R\right|$, so that $-U^l$ is increasing, negative, and convex in w. We also assume $U_g(0) = U_l(0) = 0$. The assignment of the case where $w = w^R$ to the loss region of $V(w)$ is therefore arbitrary. Lastly we impose the standard regularity conditions $\lim_{x\to 0} U_g'(x) = \lim_{x\to 0} U_l'(x) = +\infty$ and $\lim_{x\to +\infty} U_g'(x) = \lim_{x\to +\infty} U_l'(x) = 0$.

Here we assume that the reference point w^R is exogenous in two respects. Firstly, the reference point does not depend on any of the parameters of the contract. Alternative assumptions would relate the reference point to the median or the mean payoff of the contract $w(P_T)$, which would increase the mathematical complexity of the argument substantially.[4] Secondly, the reference point is also independent of the level of effort. This is defensible if the costs of effort are non-pecuniary and the manager separates the costs of effort from the pecuniary wage. However, this is potentially a strong assumption if the costs are pecuniary and the manager frames the problem so that she feels a loss if her payoff does not exceed w^R plus any additional expenses for exerting effort. In the second case, $C(e)$ should simply be added to the reference point w^R. We do not pursue this route here for mathematical tractability.

The manager has some outside employment opportunity that provides her with a utility level $\underline{V}$, so any feasible contract must satisfy the ex ante participation constraint $E[V(w)] - C(e) \geq \underline{V}$. Subsequent to accepting the contract the manager can always terminate the contract and receive a wage $\underline{w}$, hence all feasible contracts must satisfy the ex post participation constraint $w(P_T) \geq \underline{w}$. We assume $\underline{w} \leq w^R$.

[4] De Meza and Webb (2005) focus on this aspect of applying loss aversion to principal-agent theory.

3 Analysis

3.1 The general case

We now characterize the optimal contract $w^* (P_T)$ if shareholders want to implement the higher level of effort $\overline{e}$. Then shareholders' problem can be written as:

$$\min_{w(P_T)\geq \underline{w}} \int w(P_T)\, f(P_T|\overline{e})dP_T \tag{2}$$

$$s.t. \int V(w(P_T))\, f(P_T|\overline{e})dP_T \geq \underline{V} + C(\bar{e}) \quad , \tag{3}$$

$$\int V(w(P_T))\, \Delta f(P_T|e)dP_T \geq \Delta C \ . \tag{4}$$

We solve program (2) to (4) in the usual way by setting up the Lagrangian for this problem and then maximizing it pointwise with respect to w. We first address the problem that the constraints do not necessarily define a convex set as the function $V(w)$ is not concave over the loss region. It turns out that we can circumvent this problem by explicitly extending the space of permissible contracts to lotteries, where the manager obtains a random payoff for a given terminal price P_T.

Lemma 1. ***(Lotteries):*** *(i) Consider any contract that pays off $w(P_T)$ in the interior of the loss space with some positive probability, such that $\underline{w} < w(P_T) < w^R$. Then there always exists an alternative contract that improves on the contract $w(P_T)$ where the manager receives the reservation wage w^R with probability $g(P_T)$ and the minimum wage $\underline{w}$ with the remaining probability $1 - g(P_T)$. (ii) Consider any contract where the manager receives a random wage in the gain space. Then there always exists another contract that improves on this contract where the manager receives some non-random wage $w(P_T) > w^R$.*

The manager has convex preferences in the loss space, so for any given wage w that lies in the interior of the interval $[\underline{w}, w^R]$ we can find a lottery between $w^R > w$ and $\underline{w} < w$ that she prefers to w. Then shareholders can always find a lottery where the manager receives w^R with probability g such that the manager is indifferent between the wage w and this lottery. Since the manager is risk-loving in the loss space, such a lottery would reduce the costs of the contract, so only $\underline{w}$ and w^R can be optimal payoffs in the loss space. The opposite argument holds in the gain space, so we do not need to consider probabilistic payoffs there as part of the optimal contract. The important insight from Lemma 1 is that we can restrict ourselves to contracts that pay off either a non-random wage $w(P_T) \geq w^R$, or that pay off as a lottery between the minimum wage and the reference wage w^R. We can therefore ignore payoffs in the interior of the loss space altogether and only need the randomizing probability $g(P_T)$ in order to characterize contracts in the loss space.

We now want to solve program (2) to (4) by minimizing the corresponding Lagrangian. Denote the Lagrange multiplier on the participation constraint (3) by μ_{PC} and the Lagrange multiplier on the incentive compatibility constraint (4) by μ_{IC}. We can now characterize the optimal contract separately for the gain space and for the loss space. From Lemma 1, write contracts as a combination of a payoff function in the gain space and a lottery over the minimum wage and the reference wage, $\{g(P_T), w_g(P_T)\}$.

Lemma 2. *(i) The participation constraint (3) and the incentive compatibility constraint (4) define a convex set of permissible contracts.*
(ii) Whenever the optimal contract pays off in the gain space it satisfies the condition

$$\frac{1}{U_g'\left(w_g^*(P) - w^R\right)} \geq \mu_{PC} + \mu_{IC}\frac{\Delta f(P\,|e)}{f(P\,|\overline{e})}\ , \tag{5}$$

where (5) holds as an equality for all interior wages $w_g^*(P) > w^R$ and $w_g^*(P) = w^R$ if the inequality is strict. $w_g^*(P_T)$ is monotonically increasing in P_T.
(iii) If the optimal contract pays off in the loss space, then the manager receives $\underline{w}$ for all $P_T \leq P^R$ and she receives w^R for all $P_T \geq P^R$, where P^R is a uniquely defined cutoff value.

We need the first part of Lemma 2 in order to apply a Lagrangian approach to maximization. Part (ii) shows that for the gain space we obtain the familiar Holmström condition (Holmström, 1979). This is intuitive, since the problem in the gain space is not fundamentally different from a standard utility-maximizing framework. Observe that whenever the optimal contract pays off in the gain space, then preferences are represented by a concave function U_g, and the result is therefore similar to that of a conventional principal-agent contract where the agent has a concave utility function. It is important to note that we have not yet proved that the optimal contract ever pays off in the gain space.

For the loss space we already know that the optimal contract either pays off the minimum wage or the reference wage. The additional contribution of Lemma 2 is the result that these lotteries are degenerate and that the optimal contract pays off either $\underline{w}$ or w^R with probability one conditional on the firm value P_T. Again, we have said nothing so far about whether P^R is actually in the loss space. Neither can we exclude the case where $P^R = 0$, so that the contract always pays off w^R in the loss space, nor can we exclude the case where $P^R = \infty$, so that the contract always pays off $\underline{w}$ in the loss space. Also, we have not yet shown that the optimal contract ever pays off in the loss space at all.

We are now in a position to characterize the optimal contract. From Lemma 2 the optimal payoff $w^*(P_T)$ must equal either $\underline{w}$, w^R, or $w_g^*(P_T)$ as defined by (5). We still need to establish when the contract pays off in the loss space and when it pays off in the gain space. We also want to show monotonicity of the optimal contract, so we need to rule out that the optimal payoff equals $w_g^*(P_T)$ for some range of firm values and then $\underline{w}$ for some higher range of firm values.

Proposition 3. ***(Optimal contract):*** *Assume that $U_g'(x)\,x < U_g(x)$ for all $x \geq 0$. Then the optimal contract $w^*(P)$ can be written as:*

$$w^*(P) = \begin{cases} w^R + U_g'^{-1}\left(\left[\mu_{PC} + \mu_{IC}\frac{\Delta f(P|e)}{f(P|\bar{e})}\right]^{-1}\right) & if \quad P_T > \widehat{P} \\ \underline{w} & if \quad P_T \leq \widehat{P} \end{cases}. \tag{6}$$

Proposition 3 provides us with a general characterization of the optimal principal agent contract with a loss-averse manager. The contract is simple. For some region the optimal contract pays off in the gain space, where it is continuous and monotonically increasing as long as f and U_g are smooth functions. However, there is a discontinuity at some point $\widehat{P}$ where the manager's salary jumps discretely from $\underline{w}$ to some value $w^*(P) \geq w^R > \underline{w}$. In the context of executive compensation we can interpret the payoff $\underline{w}$ also as the consequence of firing the manager when she underperforms too much, so $\widehat{P}$ is the cutoff point below which she is fired. Interestingly, the optimal contract combines punishments ("sticks") with rewards ("carrots"), which sets this contract apart from that observed within a utility maximizing framework with constant relative risk aversion.

The result presented above requires the additional regularity condition $U_g'(x)\,x < U_g(x)$. This condition is satisfied by power functions with exponents in the unit interval that are commonly used in the literature and that we use in our parametric analysis below. If this condition is relaxed, then it is not guaranteed that the optimal contract does not fall back into the loss space once it was in the gain space. Moreover, without this regularity condition the optimal contract may have a third region where the payoff jumps discretely from the minimum wage $\underline{w}$ to the reference wage w^R at some stock price $\widetilde{P}_T$, stays flat for some interval, and then increases monotonically for prices above some higher stock price $\hat{P}$. We assume that the restriction $U_g'(x)\,x < U_g(x)$ holds from now on.

3.2 The continuous case

We have analyzed the case of discrete effort extensively here because then we do not need to make more restrictive assumptions apart from the regularity conditions on U_g and U_l. We now extend our analysis to the case where effort is continuous, so $e \in [0, \infty)$. In order to be able to solve this problem analogously to the way we did for the discrete case, we have to apply the first-order approach, i.e., we replace the agent's incentive compatibility constraint (4) with the first order conditions for (4). It is always legitimate to do this if we can ensure that the manager's maximization problem when choosing her effort level is globally concave, so that the first order conditions uniquely identify the maximum of her objective function.[5] In our case, this implies that

$$\frac{\partial^2 E\left(V\left(w\left(P\right)\right)|e\right)}{\partial e^2} = \int V\left(w\left(P_T\right)\right)\frac{\partial^2 f\left(P_T\,|e\right)}{\partial e^2}dP_T - \frac{\partial^2 C\left(e\right)}{\partial e^2} < 0\ . \tag{7}$$

This condition will not hold generally. Firstly, the value function V is convex over some regions. Moreover, the optimal contract $w\left(P_T\right)$ is also not concave from Proposition 3 and may well be convex over the entire gain space, depending on the functional form of U_g'. However, we can ensure that condition (7) holds for some cost functions C and some density functions in two ways. Firstly, equation (7) shows that this condition will be satisfied for sufficiently convex cost functions, so that $\partial^2 C\left(e\right)/\partial e^2$ is bounded from below so that (7) holds. Secondly, if we can rewrite P_T so that $P_T = P_0\left(e\right)\eta_T$, where η_T is a random variable and $P_0\left(e\right)$ is a concave production function, then (7) will also be satisfied if $P_0\left(e\right)$ is sufficiently concave (such that P_0'' is sufficiently small for all effort levels).

Proposition 4. ***(Continuous effort):*** *Assume that shareholders wish to implement some given effort level $\hat{e}$ and that condition (7) is satisfied for all e. Also, assume that $\frac{\partial f(P_T|e)/\partial e}{f(P_T|e)}$ is monotonically increasing in P_T (MLRP) and that $U_g'(x)\,x < U_g(x)$ for all*

[5] The literature on the principal-agent model has identified conditions where this "first-order approach" is valid. See e.g. Jewitt (1988) and Rogerson (1985).

$x \geq 0$. *Then the optimal contract* $w^*(P)$ *can be written as:*

$$w^*(P) = \begin{cases} w^R + U_g'^{-1}\left(\left[\mu_{PC} + \mu_{IC}\frac{\partial f(P|e)/\partial e}{f(P|\hat{e})}\right]^{-1}\right) & if \quad P_T > \widehat{P} \\ \underline{w} & if \quad P_T \leq \widehat{P} \end{cases}, \tag{8}$$

where $\widehat{P}$ *is a uniquely defined cut-off value.*

Proposition 4 is not surprising and shows that the whole argument of the previous subsection goes through with the same implications for the optimal contract, providing that we can assume that condition (7) holds.

4 Comparative static analysis

4.1 Parameterizing the model

We now turn to the comparative static analysis of the optimal contract. There is little we can say analytically about the contract described in Propositions 3 and 4 without further assumptions on the functional forms for the probability law f and for the value function V. We therefore parameterize the model and then provide a numerical comparative static analysis. We use the following value function, which is a special case of (1):[6]

$$V(w) = \begin{cases} \left(w - w^R\right)^\alpha & if \quad w \geq w^R \\ -\lambda\left(w^R - w\right)^\beta & if \quad w < w^R \end{cases}, \text{ where } \alpha, \beta < 1. \tag{9}$$

For this value function the condition $U_g'(x)\,x < U_g(x)$ for all $x \geq 0$ that we assume in Propositions 3 and 4 is always satisfied. Then we assume that the stock price follows a lognormal distribution and specify:[7]

$$P_T(u, e) = P_0(e)\exp\left\{\left(r_f - \frac{\sigma^2}{2}\right)T + u\sqrt{T}\sigma\right\}, \quad u \sim N(0,1), \tag{10}$$

where r_f is the risk-free rate of interest, σ^2 the variance of the returns on the stock, T the time horizon and u a standard normal random variate. This allows us to write the optimal contract for the parametric model.

[6] See Kahneman and Tversky (1979) and Tversky and Kahneman (1992).

[7] Our specification ignores dividends in order to simplify the exposition. We include dividends in the numerical analysis below.

Proposition 5. ***(Parametric model):*** *Assume that the manager's preferences are given by (9) and that the stock price is distributed lognormal as in (10). Assume that shareholders wish to implement the effort level $\hat{e}$ and that condition (7) holds for all effort levels. Then the optimal contract is given by*

$$w^*(P_T) = \begin{cases} w^R + (\gamma_0 + \gamma_1 \ln P_T)^{\frac{1}{1-\alpha}} & if \quad P_T > \widehat{P} \\ \underline{w} & if \quad P_T \leq \widehat{P} \end{cases} , \tag{11}$$

$$\textit{where: } \mu(e) = \ln(P_0(e)) + \left(r_f - \frac{\sigma^2}{2}\right) T , \tag{12}$$

$$\gamma_1 = \alpha \mu_{IC} \frac{P_0'(e)}{P_0(e)\sigma^2 T} , \tag{13}$$

$$\gamma_0 = \alpha \left(\mu_{PC} - \mu_{IC} \frac{P_0'(e)}{P_0(e)} \frac{\mu(e)}{\sigma^2 T}\right) = \alpha\mu_{PC} - \gamma_1 \mu(e) , \tag{14}$$

and $\hat{P}$ is uniquely defined by the condition:

$$\alpha\left(w^R - \underline{w}\right) = \left(\gamma_0 + \gamma_1 \ln \hat{P}\right) \lambda \left(w^R - \underline{w}\right)^\beta + (1-\alpha)\left(\gamma_0 + \gamma_1 \ln \hat{P}\right)^{\frac{1}{1-\alpha}} . \tag{15}$$

The optimal parametric contract has therefore a firing region, where the manager is fired and paid only the minimum wage $\underline{w}$ if the terminal value of the firm falls below some pre-specified price $\hat{P}$ determined by (15). Above that cutoff value the manager obtains a value w_g^* above her reference wage w^R specified by (11). Note that this function is strictly increasing in P_T and is convex as long as $P_T \leq \exp\{1 - \gamma_0/\gamma_1\}$. Above this value the function is concave, so that we cannot conclude that the optimal contract can easily explain convex payoff functions. It is therefore an empirical question whether the contract described in Proposition 5 can describe observed contracts, as the concave region may or may not be empirically relevant.[8]

We can now identify the parameters that we need to determine in order to analyze the optimal contract numerically. Firstly, we have to find appropriate values for the preference parameters α, β and w^R and for the lower bound of the wage $\underline{w}$. For these we rely on the empirical literature and on data for executive compensation contracts. Then we need the parameters that describe the distribution of P_T in (10). These are the return variance σ^2, the maturity of the contract T, the risk-free rate r_f, and the value of the firm P_0. All these

[8] A similar observation applies to the CRRA-lognormal model for parameters of relative risk aversion smaller than one.

can be determined from available data. However, we cannot determine the parameters γ_0, γ_1 and $\hat{P}$ from data directly, as they depend on the Lagrange multipliers μ_{IC} and μ_{PC} and the derivative of the production function $P_0'(e)$. If we would specify a complete model and commit ourselves to parametric forms of the production function $P_0(e)$ and of the cost function $C(e)$, then all our conclusions would depend on the assumed functional forms. We want to avoid this and therefore proceed by investigating another implication of the optimal contract that relies only on the characterization (11). This is a reduced form with three parameters: the two coefficients γ_0 and γ_1 and the cut-off point $\hat{P}$, where $\hat{P}$ is determined from (15), so only γ_0 and γ_1 are free parameters that need to be determined numerically.

Consider a CEO for whom we can completely characterize the observed contract $w^d(P)$, where we use the superscript 'd' in order to refer to 'data.' Our null hypothesis is that $w^d(P_T)$ is an optimal contract, so it can be rationalized as the outcome of an optimization program, where we make the auxiliary assumptions that preferences are parameterized as in (9) and the technology is parameterized as in (10).[9] If $w^d(P_T)$ is indeed optimal, then it should not be possible to find another contract that (i) provides the same incentives as the observed contract, (ii) provides the same utility to the CEO as the observed contract, and (iii) costs less to shareholders compared to the observed contract. We search for such a contract over all admissible parameters $\{\gamma_0, \gamma_1\}$ (recall that $\hat{P}$ is given from (15)) and change notation by writing the wage function as $w(P_T|\gamma_0, \gamma_1)$. We therefore solve the following program numerically:

$$\min_{\gamma_0,\gamma_1} \pi\left(w(P_T|\gamma_0,\gamma_1)\right) \equiv \int w(P_T|\gamma_0,\gamma_1) f(P_T) dP_T \tag{16}$$

$$s.t. \quad \int V\left(w(P_T|\gamma_0,\gamma_1)\right) f(P_T) dP_T \geq \int V(w^d(P_T)) f(P_T) dP_T \quad , \tag{17}$$

$$\int V\left(w(P_T|\gamma_0,\gamma_1)\right) \frac{\partial f(P_T)}{\partial P_0} dP_T \geq \int V(w^d(P_T)) \frac{\partial f(P_T)}{\partial P_0} dP_T \quad . \tag{18}$$

By writing P_T as in (10) and setting $P_0(e)$ equal to the observed value of the firm, we effectively treat the (unknown) effort level of the CEO as given. We can then write the density without reference to the level of effort as $f(P_T)$.

Effectively, we follow Grossman and Hart (1983) and divide the solution to the optimal contracting problem into two stages, where the first stage solves for the optimal contract

[9] The program is specified in (35) to (37) in the appendix.

for a given level of effort and determines the cost of implementing this effort level. The second stage solves for the optimal contract by trading off the costs and benefits of contracts that are optimal at the first stage. We do not consider the second stage and focus only on the first stage by solving program (16) to (18) as it does not depend on knowledge of the cost function $C(e)$ and of the production function $P_0(e)$. This implies also that we cannot analyze the optimal *level* of incentives (pay for performance sensitivity) for a compensation contract, which would invariably depend on this information. However, with our approach we can analyze the optimal *structure* of compensation contracts for any given level of incentives.

Program (16) to (18) generates a new contract $w^*(P_T)$ that is less costly to shareholders and specifies parameters γ_0^*, γ_1^*, and $\hat{P}^*$. Condition (18) ensures that the CEO has at least the same incentives under the new contract as she had under the observed contract, so that the contract found by the program will not result in a reduced level of effort. Similarly, condition (17) ensures that the contract found by the program provides at least the same value to the CEO as the observed contract, so it should also be acceptable to the CEO. We can then compare the observed contract $w^d(P_T)$ to the optimal contract $w^*(P_T)$ from (16) to (18).

4.2 Implementation and data

We develop our numerical analysis for typical CEOs. We first describe the observed contracts for CEOs using the 2004 version of the Compustat ExecuComp database. We denote the number of shares by n_S and the number of options by n_O and normalize both numbers by the total number of shares outstanding. As we consider the contract that was written in the beginning of 2004, we use the stock and option holdings from the end of the 2003 fiscal year, while we obtain the fixed salary ϕ from 2004 data. We include the annual bonus and other compensation items except long term incentive pay in the fixed salary ϕ. Long term incentive pay is not awarded annually and would therefore distort the salary if it is paid. CEOs typically have options granted at different dates with different strike prices and different remaining maturities. We estimate the maturity T and the strike price K of a single representative option that has the same value and the same Black-Scholes option delta as the actual option portfolio. We estimate the actual option portfolio following the procedure proposed by Core and Guay (2002). The volatility σ^2 is available from ExecuComp.

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile	Repres. CEO
Stock	n_S	2.00%	5.35%	0.03%	0.30%	4.61%	0.32%
Options	n_O	1.60%	1.57%	0.26%	1.14%	3.43%	1.19%
Fixed Salary ('000)	ϕ	2,143	2,288	560	1,485	4,093	1,643
Non-firm wealth ('000)	W_0	52,470	369,946	2,297	9,703	64,409	7,603,756
Firm value ('000)	P_0	8,481,084	24,212,154	329,127	1,850,161	16,348,908	1,675,670
Strike price ('000)	K	6,488,595	20,719,655	233,967	1,260,415	11,471,441	1,073,602
Moneyness	K/P_0	71.52%	19.56%	45.44%	72.92%	95.89%	64.07%
Maturity	T	5.14	1.40	0.45	5.11	6.29	4.18
Stock volatility	σ	49.4%	25.1%	27.4%	42.9%	83.5%	54.0%
Dividend rate	d	0.89%	1.27%	0.00%	0.27%	2.59%	0.00%

Table 1: Description of the dataset and of the representative CEO. This table displays mean, median, standard deviation, 10th and 90th percentile and the corresponding values of the contract parameters for the representative CEO.

We want to compare the results for loss aversion with those for expected utility with constant relative risk aversion, by far the most common model used in the compensation literature. For this conventional model, we need an estimate of the CEO's wealth. We estimate the portion of each CEO's wealth that is not tied up in securities of his or her company by tracking the CEO's income from salary, bonus, and other compensation payments, adding the proceeds from sales of securities, and subtracting the costs from exercising options. We refer to this magnitude as non-firm wealth and denote it by W_0.

Table 1 provides descriptive statistics for all variables and the data for the whole sample and for one representative CEO. The representatives CEO is the CEO with the smallest maximum deviation from the median of the variables listed in Table 1 (excluding the dividend rate). The risk-free rate must be matched to the maturity of the CEO's option holdings. We use the six-year government bond rate at the beginning of 2004 (3.39%) which we calculated from data obtained from the Federal Reserve Board's website. There are three further parameters we need to estimate in order to complete our calibration: the minimum wage $\underline{w}$, the probability of obtaining the minimum wage, and the reference point w^R.

Dismissal probability. One feature of the optimal contract is the discrete jump at the point $\hat{P}$ from $\underline{w}$ to some number above w^R. We interpret this jump as firing the CEO if the stock price falls below $\hat{P}$. Dismissal is not an explicit part of the CEO's contract with the firm. Rather, contracts are negotiated for a limited period of time and not extended, or terminated prematurely as the result of negotiations between the board of directors and the CEO. In these cases the governance structure of the company basically provides the legal context, and we include this in our concept of the optimal contract.

We estimate the observed probability of dismissal by calculating the frequency with which CEOs in the ExecuComp database leave the company within a given four-year period, where the recorded reason is 'resigned.' We repeat this for all four-year periods between 1995 and 2004 and obtain an average dismissal probability of 7.4%. Note that this number is inferred from a cross-section and the *ex ante* probabilities may well vary across CEOs. However, we have no reliable way of modeling this heterogeneity here, so we use parameters inferred from the entire sample. We use the estimated dismissal probability in two ways: First, when calculating the incentives provided by the observed contract, i.e. the right-hand side of (18), we assume that the CEO is fired with a probability of 7%, i.e. for all stock prices below the 7% quantile of the price distribution. Second, we use it as a benchmark with which we compare the dismissal probability implied by the optimal contract: $p(\hat{P}) \equiv \int_0^{\hat{P}} f(P_T)\, dP_T$. A good model of efficient contracting should generate realistic dismissal probabilities.

Note that our analysis assumes that dismissal is always performance-related. The literature on CEO dismissals suggests that some dismissals can be related to stock price performance, but this may only be a minor part. Our procedure therefore tends to overestimate the incentives provided from dismissals in observed contracts.[10]

Minimum wage. We do not have a good theory of the minimum wage in the context of our analysis. We reason that the CEO could be hired into another job with a similar compensation to her current job. However, it seems unlikely that she could obtain such a job offer when her previous company significantly underperformed expectations. It is also not plausible that her new employer would compensate her for giving up restricted stock or stock options that are practically worthless if the stock price drops below $\hat{P}$. We therefore use the fixed salary (which includes bonus payments) as a higher bound for our estimate of the minimum wage $\underline{w}$. On the other end of the distribution we regard it as conceivable (although not likely) that the CEO invests some of her own money in the company's stock, which then becomes worthless subsequently. We regard half of her

[10] See Weisbach (1988) and Kaplan (1994) for earlier contributions to this literature and Engel, Hayes, and Wang (2003) and Farrell and Whidbee (2003) for more recent contributions. Brickley (2003) states in the discussion of the last two papers that he is "struck by the limited explanatory power of the various performance measures in the CEO turnover regressions." (p. 232).

current wealth as a maximum investment, so that her total compensation could fall as low as the negative of half of her current wealth. Given the uncertainty with respect to this parameter, we provide some sensitivity analysis and evaluate the optimal contract for a plausible range of $\underline{w}$.

Reference point. Prospect theory does not provide us with clear guidance with respect to the reference point. It seems plausible that the CEO regards her salary and bonus as a "bird in the hand," which is tangible and any reduction in her fixed compensation would then be regarded as a loss. We would therefore regard a reference point below the current fixed salary (including the current bonus) as implausible. Reference compensation will most likely also include some portion of deferred compensation, where deferred compensation is taken from the existing contract. For example, the baseline for a contract negotiated for 2004 is in all likelihood the 2003 contract. Also, with loss aversion (as with risk aversion) the CEO values shares at some value below their risk-neutral market values that apply to diversified investors. We therefore expect the reference wage to lie somewhere between the fixed salary and the total market value of all compensation items, both evaluated for the previous year. As with the minimum wage, we provide sensitivity analysis in order to capture the uncertainty regarding this parameter.

Preference parameters. For the preference parameters α and λ we rely on the experimental literature for guidance. We therefore use $\alpha = \beta = 0.88$ and $\lambda = 2.25$ as our baseline values.[11]

4.3 Calibration results

We compute optimal contracts for the representative CEO whose contract and company is described in Table 1. We compute the optimal contract by solving program (16) to (18) numerically using the parameters above. Figure 1 shows the optimal contract for our baseline parameters.

We also compute the optimal contract with constant relative risk aversion (with relative risk aversion equal to 3). Visual inspection shows that the contract generated by the model with loss aversion (solid line) is similar to the stylized observed contract (dashed

[11] See Tversky and Kahneman (1992). These values have become somwhat of a standard in the literature. For experimental studies on the preference parameters which yield parameter values in a comparable range see Abdellaoui (2000) and Abdellaoui, Vossmann, and Weber (2005).

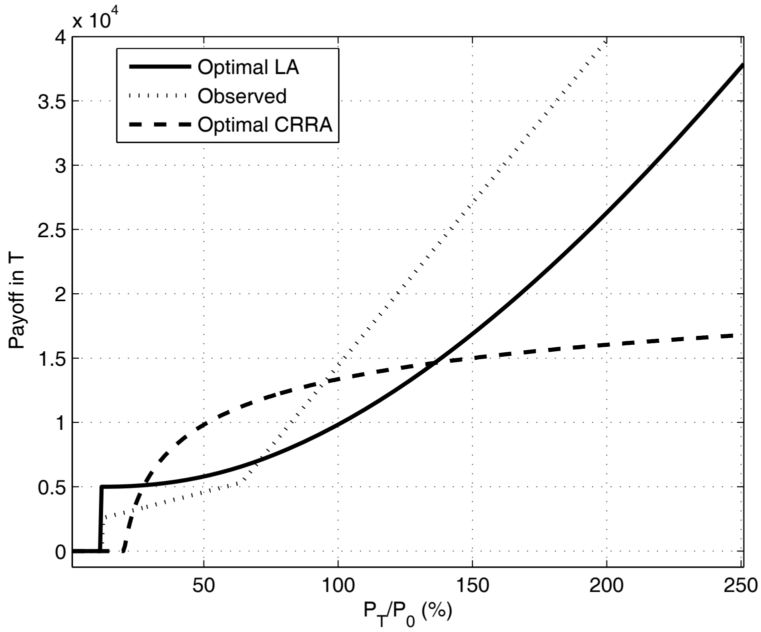

Figure 1: Optimal and observed contracts. The figure shows the optimal contract with loss aversion (solid line), with constant relative risk aversion (CRRA coefficient of 3, dashed line), and the observed contract (dotted line). The parameters for the optimal contract with loss aversion are: $w^R = \$5\text{m}$, $\alpha = \beta = 0.88$, $\lambda = 2.25$. All other parameters are those given in Table 1.

line) in several respects. Both, the observed contract and the optimal contract have a jump (from firing) that occurs at about the same stock price P_T. The optimal contract predicts a probability of dismissal of about 6%. The optimal contract is also convex in a way that is similar to the convexity generated by options in the observed contract, so the concave region here is outside the range that is reached with any reasonable probability. By contrast, the optimal contract with constant relative risk aversion is dramatically different and concave over the entire relevant region.

Clearly, visual inspection is somewhat limited in its scope for identifying the strengths and the weaknesses of the model. We therefore describe the optimal contract predicted by the models with loss aversion and with constant relative risk aversion in more detail and compare them to the observed contract. For this we use the following parameters and statistics:

- Savings represents the savings from recontracting from the observed contract to the optimal contract generated by the program, expressed as a percentage of the observed costs of the contract. We denote savings by S and calculate them as

$$S = \frac{\pi\left(w^{d}\right) - \pi(w^{*}\left(P_T \left|\gamma_0, \gamma_1\right.\right))}{\pi\left(w^{d}\right)} \ , \tag{19}$$

where w^* is the optimal contract and w^d denotes the observed compensation contract.

- The dismissal probability is defined as above, using the cut-off price $\hat{P}$: $p(\hat{P}) \equiv \int_0^{\hat{P}} f\left(P_T\right) dP_T$.

- Δ_{low} refers to the slope between the cut-off point $\hat{P}$ and the observed strike price K and is defined as

$$\Delta_{low} \equiv \int_{\hat{P}}^{K} \frac{\partial w^{*}\left(P_T\right)}{\partial P_T} \frac{f\left(P_T\right)}{F\left(P_0\right) - F\left(\hat{P}\right)} dP_T \ ,$$

  where F is the cumulated density function of f. Δ_{low} therefore measures the slope in the lower region of the contract, which corresponds to the region where observed contracts pay off only as the CEO's options are out of the money. Δ_{low} should therefore be compared to the number of shares in the observed contract.

- Δ_{high} refers to the slope above the observed strike price K and is defined as

$$\Delta_{high} \equiv \int_{K}^{\infty} \frac{\partial w^{*}\left(P_T\right)}{\partial P_T} \frac{f\left(P_T\right)}{1 - F\left(P_0\right)} dP_T \ .$$

  Δ_{high} therefore measures the slope in the upper region of the contract, which corresponds to the region where observed contracts pay off from restricted stock and from stock options.

- We calculate which portion of the CEO's incentives come from the performance sensitivity of her compensation, and which portion of incentives can be attributed to the possibility of firing if the stock price falls below $\hat{P}$. Total incentives are given from (18).

Loss aversion and risk aversion. Table 2 shows the results for the contract with loss aversion from Figure 1. Most parameters are reasonably close to observed values. The savings from recontracting S defined in (19) are 5.06%. Given that the optimal contract is always cheaper than the observed contract by construction, this is a low number. The dismissal probability is 6.28%, which is close to the estimate for the observed contract. For

Model	**Savings**	**Dismissal prob.**	Δ_{low}	Δ_{high}
Observed contract	N/A	7.00%	0.32%	1.51%
Loss aversion	5.06%	6.28%	0.16%	1.05%
CRRA, $\gamma = 0.5$	4.21%	31.64%	2.17%	1.67%
CRRA, $\gamma = 3$	46.12%	16.61%	1.90%	0.27%
CRRA, $\gamma = 5$	60.13%	11.10%	1.39%	0.12%

Table 2: Observed and predicted contracts. This table compares the observed contract of the representative CEO with the optimal contracts predicted by four different models. It shows for all contracts the probability of dismissal, the average slope of the pay function for stock prices below the observed strike price Δ_{low}, and the average slope of the pay function for stock prices above the observed stock price Δ_{high}. In addition, the table displays the savings that – according to the model – can be realized by switching from the observed contract to the respective optimal contract. The "loss aversion" contract is the optimal contract from Proposition 5, where the reference wage w_R has been set to \$5,000,000. The coefficient of loss aversion λ is set to 2.25, and the parameters for the curvature of the value function are $\alpha = \beta = 0.88$. The "CRRA" contract refers to the optimal contract from the conventional model in which the manager exhibits constant relative risk aversion. The results of this model are shown for three different values of the parameter of risk aversion γ. For all models, the minimum payout $\underline{w}$ is assumed to be zero.

the CRRA-model the results depend significantly on the degree of relative risk aversion. If risk aversion is very low, then savings are also low, but then the contract generates a high proportion of the incentives from the threat of dismissal, which is not the case for observed compensation contracts. Such a low level of risk aversion is also implausible as it implies that the CEO would hold a highly levered portfolio and invest several times her wealth in the stock market. For higher and more plausible levels of risk aversion the dismissal probability is more in line with its empirical counterpart, but now savings from recontracting are in the 40% to 60%-range and therefore about one order of magnitude larger than they are for loss aversion. These savings stem from the difference between the observed (convex) contract and the optimal (concave) contract, where the latter saves on the large risk premium CEOs demand for options. For all levels of risk aversion, the optimal contract is concave in the sense that $\Delta_{low} > \Delta_{high}$. It therefore does not predict positive option holdings with the observed strike price. We therefore conclude that loss aversion is able to explain the observed contract better than the standard CRRA model frequently applied in the literature.

w^R	**Savings**	**Dismissal prob.**	Δ_{low}	Δ_{high}	**Incentives dismissal**
3,000	2.73%	0.73%	0.41%	1.21%	0.76%
4,000	3.68%	2.71%	0.25%	1.14%	2.92%
5,000	5.06%	6.28%	0.16%	1.05%	6.79%
6,000	6.37%	10.51%	0.10%	0.95%	11.41%
7,000	7.32%	14.61%	0.07%	0.87%	16.02%
10,000	9.16%	24.98%	0.04%	0.69%	28.42%
13,000	10.35%	32.80%	0.02%	0.57%	38.53%
16,000	11.27%	38.87%	0.01%	0.48%	46.86%
19,000	12.10%	43.75%	0.01%	0.41%	53.86%
22,000	12.87%	47.78%	0.01%	0.35%	59.87%

Table 3: Sensitivity analysis with respect to the reference wage w^R. This table shows salient features of the optimal contract from Proposition 5 for different choices of the reference point w^R. For all contracts, it displays the probability of dismissal, the average slope of the pay function for stock prices below the observed strike price Δ_{low}, and the average slope of the pay function for stock prices above the observed stock price Δ_{high}. In addition, the table displays the savings that – according to the model – can be realized from switching from the observed contract to the respective optimal contract, and the proportion of incentives that are generated by the threat of dismissal. The coefficient of loss aversion λ is set to 2.25, the parameters for the curvature of the value function are $\alpha = \beta = 0.88$, and the minimum wage $\underline{w}$ is zero.

The reference wage. Table 3 summarizes the contract with the parameters specified above for the representative CEO, where the reference wage w^R ranges from the base salary ϕ to about twice the value of the contract if it is valued at the current price P_0 (which includes a higher range than what we regard as plausible). We observe that both slopes, Δ_{low} and Δ_{high} for the optimal contract are uniformly decreasing in the reference wage w^R. The higher the reference wage, the more likely it is that the CEO is fired under the optimal contract, as the price $\hat{P}$ increases in w^R. Hence, the discrete jump in the wage from $\underline{w}$ to w^R does not only become larger as w^R increases, but it also becomes more likely. As a result, a larger fraction of the incentives is provided through firing the CEO, and the contribution of the slope of the compensation contract above $\hat{P}$ becomes less important and the contract becomes therefore flatter. The contract predicted by the model looks similar to the observed contract if the reference wage w^R is about \$4 million to \$6 million, which corresponds to about 20%-30% of the market value of her pay package and is 2.5 times to 4 times her fixed compensation ϕ. In this range, the slopes Δ_{low} and Δ_{high} correspond approximately to her shares and options and the dismissal probability is lower

$\underline{w}$		Savings	Dismissal prob.	Δ_{low}	Δ_{high}	Incentives dismissal
-6,083	= -0.8*W_0	7.63%	4.35%	0.22%	0.98%	11.05%
-4,563	= -0.6*W_0	7.05%	4.71%	0.21%	1.00%	10.14%
-3,042	= -0.4*W_0	6.44%	5.13%	0.19%	1.02%	9.14%
-1,521	= -0.2*W_0	5.78%	5.64%	0.18%	1.03%	8.04%
0		5.06%	6.28%	0.16%	1.05%	6.79%
821	= 0.5*ϕ	4.64%	6.69%	0.14%	1.05%	6.02%
1,642	= 1.0*ϕ	4.18%	7.17%	0.13%	1.06%	5.19%

Table 4: Sensitivity analysis with respect to the minimum wage $\underline{w}$. This table shows salient features of the optimal contract from Proposition 5 for different choices of the minimum wage $\underline{w}$. For all contracts, it displays the probability of dismissal, the average slope of the pay function for stock prices below the observed strike price Δ_{low}, and the average slope of the pay function for stock prices above the observed stock price Δ_{high}. In addition, the table displays the savings that – according to the model – can be realized from switching from the observed contract to the respective optimal contract, and the proportion of incentives that are generated by the threat of dismissal. The reference wage w^R has been set to \$5,000,000, the coefficient of loss aversion λ is set to 2.25, and the parameters for the curvature of the value function are $\alpha = \beta = 0.88$.

than the probability estimated above for the whole sample, but in a similar range. Recall also that our procedure overstates the performance-related probability of CEO dismissals.

It is interesting to note that the probability of dismissal increases in w^R. Hence, a CEO who has a higher reference wage will generally demand and also obtain a higher level of compensation. However, at the same time this higher wage increases the threshold level below which the CEO is fired, and we interpret this threshold level as a performance target. So, higher reference wages are associated with higher performance targets.

We can also compare the observed contract to the contract predicted by the model with respect to the implied costs of the contract. The contract predicted by the model is cheaper than the observed contract by construction: it has the theoretically optimal shape from (11) and satisfies the constraints (18) and (17). We can therefore use the savings as a metric of "closeness" between the theoretical contract and the observed contract from the point of view of shareholders, who consider the costs of the contract. If w^R is in the range described above, then savings are around 5% to 7% of current compensation costs.

λ	**Savings**	**Dismissal prob.**	Δ_{low}	Δ_{high}	**Incentives dismissal**
1.00	2.06%	27.00%	0.67%	1.17%	10.88%
1.25	2.24%	16.01%	0.44%	1.14%	8.08%
1.50	2.81%	10.76%	0.31%	1.12%	7.02%
1.75	3.52%	8.25%	0.24%	1.09%	6.62%
2.00	4.28%	6.96%	0.19%	1.07%	6.59%
2.25	5.06%	6.28%	0.16%	1.05%	6.79%
2.50	5.83%	5.92%	0.13%	1.03%	7.15%
2.75	6.60%	5.76%	0.11%	1.00%	7.65%
3.00	7.35%	5.72%	0.10%	0.98%	8.24%

Table 5: Sensitivity analysis with respect to the coefficient of loss aversion λ. This table shows salient features of the optimal contract from Proposition 5 for different choices of the coefficient of loss aversion λ. For all contracts, it displays the probability of dismissal, the average slope of the pay function for stock prices below the observed strike price Δ_{low}, and the average slope of the pay function for stock prices above the observed stock price Δ_{high}. In addition, the table displays the savings that – according to the model – can be realized from switching from the observed contract to the respective optimal contract, and the proportion of incentives that are generated by the threat of dismissal. The reference wage w^R has been set to \$5,000,000, the parameters for the curvature of the value function are $\alpha = \beta = 0.88$, and the minimum wage $\underline{w}$ is zero.

The minimum wage. Table 4 shows how the minimum wage $\underline{w}$ influences the optimal contract. We vary $\underline{w}$ from minus 80% of the wealth of the CEO to the current fixed salary. Generally, the efficiency loss from suboptimal contracting as captured by the savings decreases as the minimum wage increases. A higher minimum wage makes firing less painful for the CEO and therefore easier from an *ex ante* contracting perspective. The dismissal probability therefore increases in $\underline{w}$. However, there is a countervailing effect: the threat of dismissal becomes also less effective, and this effect dominates, so that overall a smaller proportion of the incentives is contributed by the threat of dismissal. At the same time, the contribution of incentives in the low region ("shares", below K) declines, and incentives in the high region ("options", above K) become progressively more important as $\underline{w}$ increases. However, all of these effects appear somewhat small: in comparison to the reference wage w^R the minimum wage $\underline{w}$ appears less important for our results. We therefore conclude that our analysis is robust to mistakes in assessing this parameter.

Loss aversion (λ). The coefficient of loss aversion, λ, is critical for the understanding of the problem. The uncertainty surrounding this parameter seems low as most papers based on experimental data agree on a relatively narrow range.[12] We tabulate our results in Table 5 and find that for $\lambda = 1$, savings from recontracting are very small. This is not surprising, because we also use $\alpha = 0.88$, so with $\lambda = 1$ the CEO becomes practically risk-neutral and the choice of compensation contract becomes less relevant as all instruments that generate the same performance sensitivity (and therefore satisfy (18)) and the same utility (and therefore satisfy (17)) have almost the same costs. For $\alpha = 1$ and $\lambda = 1$ we should have that savings are precisely equal to zero. We cannot compute this because in this case the program (16) to (18) does not have a unique solution anymore and becomes degenerate and therefore numerically instable. Savings from recontracting increase in λ as the risk premium the manager demands increases, so the costs of inefficient contracting also increase. Hence, we cannot regard the savings from recontracting as a yardstick of the model here, as risk-neutral managers would always be indifferent between different forms of incentive provision and the structure of compensation contracts would become irrelevant. It is therefore appropriate to calibrate λ by relying on experimental data.

Higher loss aversion also changes the balance between incentives from the threat of firing and incentives from performance-sensitive pay towards the threat of firing. For the manager the loss in prospect value from being paid $\underline{w}$ instead of w^R is directly proportional to λ, so the threat of firing becomes larger. Of course, this means that the incentives from the threat of firing increase just as much as the risk premium associated with this prospect. The table shows that the higher risk-premium dominates the increased incentives: The optimal contract features progressively lower dismissal probabilities as the CEO becomes more loss averse.

Risk aversion parameters (α, β). We relate α here to risk aversion, even though, strictly speaking, this is true only in the gain space. Consider the extreme scenario where the contract would pay off only in the gain space. Then the value function would be similar to a utility function with an Arrow-Pratt measure of constant relative risk aversion of $1 - \alpha$, so our choice of $\alpha = 0.88$ would be very close to risk neutrality and a higher value of α corresponds to lower risk aversion. Table 6 tabulates our comparative static analysis in terms of α. Interestingly, savings from recontracting vary only moderately over the relevant range and become much larger only for very low estimates of α. However, the dismissal probability becomes surprisingly large for low values of α. Most estimates

[12] See the papers on experimental measurements of preference parameters in footnote 11 on p. 11 above.

α	**Savings**	**Dismissal prob.**	Δ_{low}	Δ_{high}	**Incentives dismissal**
0.60	15.84%	25.14%	0.89%	1.04%	30.60%
0.65	12.52%	23.89%	0.78%	1.11%	26.76%
0.70	9.72%	22.05%	0.66%	1.15%	23.02%
0.75	7.46%	19.40%	0.52%	1.17%	19.25%
0.80	5.85%	15.62%	0.37%	1.16%	15.21%
0.85	5.04%	10.31%	0.23%	1.11%	10.40%
0.90	5.32%	3.39%	0.12%	0.99%	4.00%
0.92	5.80%	0.92%	0.10%	0.93%	1.27%

Table 6: Sensitivity analysis with respect to the curvature parameter of the value function in the gain space, α. This table shows salient features of the optimal contract from Proposition 5 for different choices of the curvature parameter of the value function in the gain space α. For all contracts, it displays the probability of dismissal, the average slope of the pay function for stock prices below the observed strike price Δ_{low}, and the average slope of the pay function for stock prices above the observed stock price Δ_{high}. In addition, the table displays the savings that – according to the model – can be realized from switching from the observed contract to the respective optimal contract, and the proportion of incentives that are generated by the threat of dismissal. The reference wage w^R has been set to \$5,000,000, the coefficient of loss aversion λ is set to 2.25, the parameter for the curvature of the value function in the loss space is $\beta = 0.88$, and the minimum wage $\underline{w}$ is zero.

in the literature lie in the range from 0.85 to 0.92, and within this range the model seems to be robust to changes in this parameter.

Table 7 provides the same analysis of the parameter β. We set α equal to β in our baseline case, but the two parameters have different interpretations and we therefore conduct the sensitivity analysis for both of them independently. Varying β does not have a dramatic impact on savings and the dismissal probability, but it significantly affects the balance of "shares" (Δ_{low}) and "options" (Δ_{high}). β measures the diminishing sensitivity towards progressively larger losses, and lower $\beta-$values imply that the sensitivity towards larger losses is smaller. Interestingly (and somewhat counterintuitively) a higher sensitivity (lower β) reduces the performance sensitivity in the lower range more than it does in the upper range.

5 Conclusion

We have developed a principal agent model with a loss-averse agent in order to explain observed executive compensation contracts. We develop the optimal contract and show

β	**Savings**	**Dismissal prob.**	Δ_{low}	Δ_{high}	**Incentives dismissal**
0.60	9.23%	8.78%	0.06%	0.92%	11.43%
0.65	8.38%	8.02%	0.07%	0.95%	10.22%
0.70	7.57%	7.41%	0.09%	0.97%	9.19%
0.75	6.81%	6.93%	0.11%	1.00%	8.33%
0.80	6.10%	6.58%	0.12%	1.02%	7.62%
0.85	5.43%	6.35%	0.14%	1.04%	7.06%
0.90	4.82%	6.25%	0.17%	1.05%	6.63%
0.95	4.25%	6.28%	0.19%	1.07%	6.31%

Table 7: Sensitivity analysis with respect to the curvature parameter of the value function in the loss space, β. This table shows salient features of the optimal contract from Proposition 5 for different choices of the curvature parameter of the value function in the loss space β. For all contracts, it displays the probability of dismissal, the average slope of the pay function for stock prices below the observed strike price Δ_{low}, and the average slope of the pay function for stock prices above the observed stock price Δ_{high}. In addition, the table displays the savings that – according to the model – can be realized from switching from the observed contract to the respective optimal contract, and the proportion of incentives that are generated by the threat of dismissal. The reference wage w^R has been set to \$5,000,000, the coefficient of loss aversion λ is set to 2.25, the parameter for the curvature of the value function in the gain space is $\alpha = 0.88$, and the minimum wage $\underline{w}$ is zero.

that it can be characterized by an upward sloping function that is convex over the relevant region for plausible parameterizations and by a firing rule for the manager. The manager suffers a discrete loss of compensation if the stock price falls below a contractually specified threshold.

We parameterize this model in a way that is standard in the literature. For the preference parameters we choose values that emerge from the experimental literature. We assume that stock prices are distributed lognormal with parameters taken from data. For the reference wage of the manager and for the minimum compensation that specifies a lower bound on her contract the theory and the data only suggest ranges, so we perform sensitivity analyses on these. We can calibrate optimal contracts surprisingly well for typical CEOs. The slope and convexity of the contract appears close to those of observed contracts that feature restricted stock, stock options, and fixed salaries. We also calculate the potential savings from switching from the observed (piecewise linear) contracts to the contracts suggested by the model. We regard these savings from recontracting as a metric that measures how closely the model predicts observed contracts and find that for plausible parameter values the savings from recontracting fall below 5% of the actual costs of compensating the CEO, which we regard as a good approximation for an arguably simple static contracting model.

We make a number of assumptions when implementing this model on which empirical evidence is still scarce. Firstly, we assume that CEOs regard fixed salaries and deferred compensation as part of one integral compensation package and that they trade off gains and losses across all compensation items. It seems to be equally plausible that CEOs would regard current cash compensation as separate from deferred compensation and mentally account for it separately. The implications for our analysis from changing this assumption would probably be minor and then our results would apply to the structure of deferred, incentive-related compensation only.

We have used only some of the components of prospect theory by using the value function proposed by Kahneman and Tversky. We have neglected the other component, namely the probability weighting function. From the point of view of prospect theory, this is a major compromise since risk aversion is modeled through the decision weights as well as through the value function. However, at this point an inclusion of the probability weighting function appears analytically intractable as we would have to find conditions that preserve the monotone likelihood ratio property after transforming the decision weights. Finally, we have demonstrated our results only for one representative CEO. We regard our research here as a pilot study and will extend this analysis to a larger sample in future research.

6 Appendix

Proof of Lemma 1:

(i) We first show that it is optimal to replace any contract that pays off in the interior off the loss space by a lottery. Consider the proposed candidate contract $w(P_T)$ that pays off $\underline{w} < w(P) < w^R$ at some price P_T with certainty. Since $U_l\left(w^R - w\right)$ is monotonically decreasing in w, we have $U_l\left(w^R - w^R\right) < U_l\left(w^R - w(P)\right) < U_l\left(w^R - \underline{w}\right)$. Hence, there exists a unique number $g(w)$ for each $w \in \left[\underline{w}, w^R\right]$ such that

$$g(w)\, U_l\left(w^R - w^R\right) + (1 - g(w))\, U_l\left(w^R - \underline{w}\right) = U_l\left(w^R - w(P)\right) \ . \tag{20}$$

This implies that replacing the payoff $w(P)$ with the lottery $\left\{g(w), w^R; 1 - g(w), \underline{w}\right\}$ leaves the participation constraint (3) and the incentive compatibility constraint (4) unchanged. From the concavity of U_l we also have:

$$g(w)\, U_l\left(w^R - w^R\right) + (1 - g(w))\, U_l\left(w^R - \underline{w}\right) \leq U_l\left(w^R - \left(g(w)\, w^R + (1 - g(w))\, \underline{w}\right)\right) \ . \tag{21}$$

Combining equations (20) and (21) yields:

$$U_l\left(w^R - w(P)\right) \leq U_l\left(w^R - \left(g(w)\, w^R + (1 - g(w))\, \underline{w}\right)\right) \ . \tag{22}$$

U_l is increasing in its argument and therefore decreasing in w, therefore $g(w)\, w^R + (1 - g(w))\, \underline{w} \leq w(P)$, so the lottery $\left\{g(w), w^R; 1 - g(w), \underline{w}\right\}$ improves on the original contract $w(P)$. Finally, consider a contract that pays off w with $\underline{w} < w < w^R$ with some probability p less than one. Then we can use the same argument as above, but we replace the random payoff w with the lottery $\left\{g(w)\, p, w^R; (1 - g(w))\, p, \underline{w}\right\}$.

(ii) Suppose the optimal contract pays off in the gain space so that the manager receives wages $w \geq w^R$ with probabilities described by some probability law $H(w\,|P_T)$. We can always define lotteries H' that extend over the gain region and the loss region by redefining the cumulative density function as $dH = dH' / \left(1 - H\left(w^R\right)\right)$, so that $\int_{w^R}^{\infty} dH = 1$. Then from the concavity of U_g we can always find a fixed payment $\hat{w} < E_H(w)$ such that $U_g(\hat{w}) = E_H(U_g(w))$, where E_H is the expectations operator with respect to H. Hence, any lottery in the gain space is dominated by some fixed payoff in the gain space. **Q.E.D.**

Proof of Lemma 2:

Step 1: Legitimacy of Lagrangian approach (statement (i)). Before we prove the results stated in the Lemma we need to show that we can legitimately set up the problem as a Lagrangian. Denote by $I(P_T)$ the indicator function that is equal to one if the contract pays out in the gain space and equal to zero if it pays out in the loss space. With the result from Lemma 1 and the assumption that $U_g\left(w^R - w^R\right) = 0$, we can rewrite the optimization problem (2) - (4) as follows:

$$\min_{w(P_T)\geq w^R} \int \left[I(P_T)w\left(P_T\right) + (1 - I(P_T))(g(P_T))w^R + (1 - g(P_T))\underline{w})\right] f(P_T|\overline{e})dP_T \quad (23)$$

$$s.t. \int \left[I(P_T)U_g(w\left(P_T\right) - w^R) - (1 - I(P_T))(1 - g(P_T)U_l(w^R - \underline{w})\right] f(P_T|\overline{e})dP_T$$

$$\geq \underline{V} + C\left(\bar{e}\right), \quad (24)$$

$$\int \left[I(P_T)U_g(w\left(P_T\right) - w^R) - (1 - I(P_T))(1 - g(P_T))U_l(w^R - \underline{w})\right] \Delta f(P_T|e)dP_T \geq \Delta C\ .$$

In order to ensure quasi-convexity of the contract space, we allow for $I(P_T) \in [0, 1]$, even though a contract with $0 < I(P_T) < 1$ is not economically meaningful. It will turn out that $I(P_T) \notin \{0, 1\}$ will never be optimal. The Lagrange approach is justified if the contract space is quasi-convex, i.e. if the space defined by the two restrictions

$$\int \left[I(P_T)U_g(w\left(P_T\right) - w^R) - (1 - I(P_T))(1 - g(P_T)U_l(w^R - \underline{w})\right] f(P_T|\overline{e})dP_T \geq K_1 \quad (25)$$

$$\int \left[I(P_T)U_g(w\left(P_T\right) - w^R) - (1 - I(P_T))(1 - g(P_T))U_l(w^R - \underline{w})\right] \Delta f(P_T|e)dP_T \geq K_2 \quad (26)$$

is convex for all constants K_1 and K_2. Assume that the two contracts $(I_1(P_T), w_1(P_T), g_1(P_T))$ and $(I_2(P_T), w_2(P_T), g_2(P_T))$ fulfill these two equations. Then we obtain for any $\alpha \in (0, 1)$:

$$\int \left[(\alpha I_1(P_T) + (1-\alpha)I_2(P_T))\, U_g(\alpha w_1(P_T) + (1-\alpha)w_2(P_T) - w^R)\right.$$
$$\left. -(1-\alpha I_1(P_T) - (1-\alpha)I_2(P_T))(1-\alpha g_1(P_T) - (1-\alpha)g_2(P_T))U_l(w^R - \underline{w})\right] f(P_T|\overline{e})dP_T$$
$$\geq \alpha^2 \int \left[I_1(P_T)U_g(w_1(P_T) - w^R) - (1-I_1(P_T))(1-g_1(P_T)U_l(w^R - \underline{w})\right] f(P_T|\overline{e})dP_T$$
$$+ \alpha(1-\alpha) \int \left[I_1(P_T)U_g(w_2(P_T) - w^R) - (1-I_1(P_T))(1-g_2(P_T)U_l(w^R - \underline{w})\right] f(P_T|\overline{e})dP_T$$
$$+ \alpha^2 \int \left[I_2(P_T)U_g(w_1(P_T) - w^R) - (1-I_2(P_T))(1-g_1(P_T)U_l(w^R - \underline{w})\right] f(P_T|\overline{e})dP_T$$
$$+ \alpha(1-\alpha) \int \left[I_2(P_T)U_g(w_2(P_T) - w^R) - (1-I_2(P_T))(1-g_2(P_T)U_l(w^R - \underline{w})\right] f(P_T|\overline{e})dP_T$$
$$\geq K_1$$

Similarly, one can show that the convex combination of the two contracts satisfies (26). For this, one only needs to replace $f(P_T|\overline{e})$ with $\Delta f(P_T|e)$ and K_1 with K_2. Thus the space of contracts defined by the constraints is quasi-convex, so that the Kuhn-Tucker approach can be applied.

Step 2: Optimal contract in the gain space (statement (ii)). The first order conditions for $w(P_T)$ become is:

$$\frac{\partial \mathcal{L}}{\partial w(P_T)} = f(P_T|\overline{e}) - \mu_{PC}U_g'\left(w^* - w^R\right) f(P_T|\overline{e}) - \mu_{IC}U_g'\left(w^* - w^R\right) \Delta f(P_T|e) \tag{27}$$
$$= U_g'\left(w^* - w^R\right) f(P_T|\overline{e}) \left[\frac{1}{U_g'\left(w^* - w^R\right)} - \mu_{PC} - \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}\right] \geq 0\ . \tag{28}$$

Note that $U_g'\left(w^* - w^R\right) f(P_T|\overline{e}) > 0$. Then the condition has to hold as an equality for all $w_g^*(P_T) > w^R$. Otherwise, if $\frac{\partial \mathcal{L}}{\partial w} > 0$ over the entire gain space, then the solution is at the lowest possible value at the constraint $w_g^*(P_T) \geq w^R$ is binding. From MLRP and (5) we can infer directly that the optimal contract is monotonically increasing in the gain space.

Step 3: Optimal contract in the loss space (statement (iii)). The first order condition for the optimal choice of g becomes:

$$\frac{\partial \mathcal{L}}{\partial g(P_T)} = f(P_T|\overline{e})U_l\left(w^R - \underline{w}\right) \left[\frac{w^R - \underline{w}}{U_l\left(w^R - \underline{w}\right)} - \mu_{PC} - \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}\right] = 0\ . \tag{29}$$

We have $f(P_T|\overline{e})U_l\left(w^R - \underline{w}\right) > 0$ by assumption. The only part of the expression in brackets that depends on P_T is $\Delta f(P_T|e)/f(P_T|\overline{e})$, which is increasing in P_T from assuming MLRP, hence there can be at most one cut-off point P^R that satisfies (29) as an equality. For any point $P_T > P^R$ we have $\frac{\partial \mathcal{L}}{\partial g(P_T)} < 0$, so that $\mathcal{L}$ is minimized by increasing g to its upper limit, so $g = 1$. Conversely, for any point $P_T < P^R$ we have $\frac{\partial \mathcal{L}}{\partial g(P_T)} > 0$, so that $\mathcal{L}$ is minimized by reducing g to its lower limit, so $g = 0$. Hence, interior probabilities $0 < g < 1$ are never optimal and the optimal lottery is always degenerate. Then the resulting contract is deterministic with a cut-off value P^R. **Q.E.D.**

Proof of Proposition 3:

For notational ease define $x \equiv w_g(P_T) - w^R$, and $y = w^R - w_l(P_T)$, i.e. $y = w^R - \underline{w}$ if $P < P^R$ and $y = 0$ if $P > P^R$. Then, the Lagrangian becomes

$$\begin{aligned}\mathcal{L} = & \int \left[(1 - I(P_T))\, w_l(P_T) + I(P_T)\, w_g(P_T)\right] f(P_T|\overline{e})dP_T \\ & + \mu_{PC}\left[\underline{V} + C(\overline{e}) + \int \left[(1 - I(P_T))\, U_l(y) - I(P_T)\, U_g(x)\right] f(P_T|\overline{e})dP_T\right] \\ & + \mu_{IC}\left[\Delta C + \int \left[(1 - I(P_T))\, U_l(y) - I(P_T)\, U_g(x)\right] \Delta f(P_T|e)dP_T\right] .\end{aligned} \tag{30}$$

Differentiating with respect to $I(P_T)$ yields

$$\frac{\partial \mathcal{L}}{\partial I(P_T)} = f(P_T|\overline{e})\left[U_l(y) + U_g(x)\right]\left[\underbrace{\frac{x + y}{U_l(y) + U_g(x)}}_{>0} - \mu_{PC} - \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}\right] \tag{31}$$

As $f(P_T|\overline{e})\left[U_l(y) + U_g(x)\right] > 0$, the term in the large brackets determines the sign of equation (31). Now we have to consider two cases:

Case 1: $\mu_{PC} + \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})} < 0$. Since we assume MLRP, this can only be the case for all P_T smaller than some $\widetilde{P}_T$ for which $\mu_{PC} + \mu_{IC}\frac{\Delta f(\widetilde{P}_T|e)}{f(\widetilde{P}_T|\overline{e})} = 0$. In this case we then have from equation (31) that $\frac{\partial \mathcal{L}}{\partial I(P_T)} > 0$. Hence for all $P_T < \widetilde{P}_T$ it is optimal to set $I(P_T)$ to its lowest possible level, zero. But this implies by construction that the contract always pays off in the loss space for all $P_T \in (0, \widetilde{P}_T)$, i.e. $(0, \widetilde{P}_T) \subset \Pi^l$.

Case 2: $\mu_{PC} + \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})} > 0$. In this case, we can define the function $x(P_T)$:

$$\frac{1}{U_g'(x)} = \mu_{PC} + \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}. \tag{32}$$

For all P_T where the contract pays off in the gain space, this is the exactly the condition for the optimal contract as established in Lemma (2). However, it should be noted that equation (32) is defined over all $P_T \in (\widetilde{P}_T, \infty)$ and not just over the gain space, which by Case 1 must be a subset of $(\widetilde{P}_T, \infty)$. Hence at this point we presume nothing about whether the contract actually pays off in the loss space, or in the gain space, for any given $P_T > \widetilde{P}_T$. Now, using (32) in (31) we get

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial I\left(P_T\right)} &= f(P_T|\overline{e})\left[U_l\left(y\right)+U_g\left(x\right)\right]\left[\frac{x+y}{U_l\left(y\right)+U_g\left(x\right)}-\frac{1}{U_g'\left(x\right)}\right] \\
&= \frac{f(P_T|\overline{e})}{U_g'\left(x\right)}\left[U_g'\left(x\right)\left(x+y\right)-U_l\left(y\right)-U_g\left(x\right)\right] \\
&= \frac{f(P_T|\overline{e})}{U_g'\left(x\right)}\cdot z\left(x(P_T),x(P_T)\right),
\end{aligned}$$

where $z\left(x,y\right)\equiv U_g'\left(x\right)\left(x+y\right)-U_l\left(y\right)-U_g\left(x\right)$. Note that y is constant on the intervals $(-\infty, P^R)$ and (P^R, ∞). Hence, $z(x,y)$ is a strictly decreasing function in x because $z'\left(x\right)=U_g''\left(x\right)\left(x+y\right)<0$ as $U_g\left(\cdot\right)$ is concave. As $x(P_T)$ defined by (32) is strictly increasing in P_T, $z(x,y)$ is strictly decreasing in P_T on these two intervals. Consequently, there can be at most two solutions to the first order condition $\frac{\partial\mathcal{L}}{\partial I(P_T)}=0$: one for $y=0$ and one for $y=w^R-\underline{w}$. In the first case, $z(x,y)=0$ is equivalent to $U_g'\left(x\right)x-U_g\left(x\right)$ which is by assumption smaller than zero. Consequently, there is at most one solution to the first order condition that defines a unique value $\widehat{P}$ for which it holds that

i) $\frac{\partial\mathcal{L}}{\partial I(P_T)}>0$, for all $P_T<\widehat{P}$
ii) $\frac{\partial\mathcal{L}}{\partial I(P_T)}<0$, for all $P_T>\widehat{P}$

$\widehat{P}$ is given by $z(x,y)=0$, i.e.:

$$U_g'\left(w^*\left(\widehat{P}\right)-w^R\right)\left(w^*\left(\widehat{P}\right)-\underline{w}\right)-\left[\lambda U_l\left(w^R-\underline{w}\right)+U_g\left(w^*\left(\widehat{P}\right)-w^R\right)\right]=0 \tag{33}$$

Hence, we have established in Case 1 and Case 2, that loss space and gain space are non-empty intervals, $\Pi^l=(0,\widehat{P}_T)$ and $\Pi^g=(\widehat{P}_T,+\infty)$. To establish that the optimal contract cannot feature a region in the loss space where $w_l^*\left(P_T\right)=w^R$, look again at equation (29) from the Proof of Lemma 2, which we state here again for convenience

$$\frac{\partial\mathcal{L}}{\partial g\left(P_T\right)}=f(P_T|\overline{e})U_l\left(w^R-\underline{w}\right)\left[\frac{w^R-\underline{w}}{U_l\left(w^R-\underline{w}\right)}-\mu_{PC}-\mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}\right]. \tag{34}$$

This is zero if the term in the square brackets is zero, which can only be the case for $\mu_{PC} + \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\overline{e})} > 0$. By the same logic as before, we can rewrite this for $P_T > \widetilde{P}_T$, using (32) as

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial g(P_T)} &= f(P_T|\overline{e}) \left[U_g^{'}(x)\left(w^R - \underline{w}\right) - U_l\left(w^R - \underline{w}\right)\right] \\ &\geq f(P_T|\overline{e}) \left[U_g^{'}(x)\, y - U_l(y)\right], \forall P_T > \widetilde{P}_T.\end{aligned}$$

Comparing the term in square brackets in this equation with $z(x)$, and using the assumption that $U_g'(x)\, x < U_g(x)$ for all $x \geq 0$, we have that $z(x)$ is always zero before the jump in the loss space from $\underline{w}$ to w^R occurs, which is just what equation (29) determines. Hence the optimal contract pays off $\underline{w}$ in the loss space for all $P_T < \widehat{P}_T$, and $w_g^*(P_T)$ in the gain space for $P_T > \widehat{P}_T$, where $w_g^*(P_T)$ can be found by solving equation (32) for $w_g(P_T)$. **Q.E.D.**

Proof of Proposition 4:

Shareholders' problem if they wish to minimize the contracting costs for implementing effort level $\hat{e}$ can be written as:

$$\min_{w(P_T) \geq \underline{w}} \int w(P_T)\, f(P_T|\hat{e}) dP_T \tag{35}$$

$$\begin{aligned} s.t. \; & - \int_{\Pi^l} U_l\left(w^R - w(P_T)\right) f(P_T|\hat{e}) dP_T \\ & + \int_{\Pi^g} U_g\left(w(P_T) - w^R\right) f(P_T|\hat{e}) dP_T \geq \underline{V} + C(\hat{e}) \quad , \end{aligned} \tag{36}$$

$$- \int_{\Pi^l} U_l\left(w^R - w(P_T)\right) f_e(P_T|\hat{e}) dP_T + \int_{\Pi^g} U_g\left(w(P_T) - w^R\right) f_e(P_T|\hat{e}) dP_T \geq C' \;, \tag{37}$$

where C' denotes the first derivative of C and f_e denotes the first derivative of f with respect to e. Since optimization of program (35) to (37) is pointwise, the only changes with respect to program (2) to (4) are: replace ΔC with C', which is a constant for a given level of effort in both programs; replace $f(P_T|\overline{e})$ with $f(P_T|\hat{e})$, which is just a density that has the same properties in both programs; replace $\Delta f(P_T|e)$ with $f_e(P_T|\hat{e})$, which also has the same properties in both programs as we assume MLRP in both cases. Hence, the same arguments as in Lemmas 1 and 2 and in Proposition 3 goes through as before. Substituting the parametric form of the value function (9) and the functional form of the optimal contract (11) into condition (33) yields (15). **Q.E.D.**

Proof of Proposition 5:

The general characterization of the optimal contract follows directly from Proposition 5. From (10) and (12) we have $\ln(P_T) = \mu(e) + u\sigma\sqrt{T}$, which is distributed normal with mean $\mu(e)$ and standard deviation $\sigma\sqrt{T}$. We write the density $f(P_T|e)$ of the lognormal distribution as:

$$f(P_T|e) = \frac{1}{P_T\sqrt{2\pi T}\sigma} \exp\left\{-\frac{[\ln P_T - \mu(e)]^2}{2\sigma^2 T}\right\} . \tag{38}$$

Then the likelihood ratio is

$$\frac{\partial f(P_T|e)/\partial e}{f(P_T|e)} = \frac{P_0'(e)}{P_0(e)} \frac{\ln P_T - \mu(e)}{\sigma^2 T} . \tag{39}$$

This allows us to rewrite the equivalent of condition (5) as:

$$\frac{1}{\alpha\left(w_g^*(P_T) - w^R\right)^{\alpha-1}} = \mu_{PC} + \mu_{IC}\frac{P_0'(e)}{P_0(e)} \frac{\ln P_T - \mu(e)}{\sigma^2 T} \tag{40}$$

From this and Proposition 4 equation (11) follows immediately. **Q.E.D.**

References

[1] Abdellaoui, Mohammed, 2000, Parameter-Free Elicitation of Utility and Probability Weighting Functions, Management Science, 46(11), pp. 1497—1512.

[2] Abdellaoui, Mohammed, Frank Vossmann, and Martin Weber, 2005, Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty, Management Science, forthcoming.

[3] Barberis, Nicholas, and Ming Huang, 2004, The Loss Aversion / Narrow Framing Approach to the Equity Premium Puzzle, (May)

[4] Barberis, Nicholas, and Ming Huang, 2001, Mental Accounting, Loss Aversion, and Individual Stock Returns , The Journal of Finance 56, no. 4 (August), pp. 1247-1292.

[5] Barberis, Nicholas, Ming Huang, and Tano Santos 2001, Prospect Theory and Asset Prices, The Quarterly Journal of Economics, CXVI(1), pp. 1-53.

[6] Benartzi, Shlomo, and Richard H. Thaler, 1995, Myopic Loss Aversion and the Equity Premium Puzzle , The Quarterly Journal of Economics 110, no. 1 (February), pp. 73-92

[7] ———, 1999, Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments , Management Science 45, no. 3 (May), pp. 364-381

[8] Brickley, James A., 2003 , Empirical Research on CEO Turnover and Firm-Performance: a Discussion, Journal of Accounting and Economics 36, no. 1-3 (December), pp. 227-233

[9] Camerer; Colin ; Linda Babcock; George Loewenstein, and Richard H. Thaler, 1997, Labor Supply of New York City Cabdrivers: One Day at a Time , The Quarterly Journal of Economics 112, no. 2 (May), pp. 407-441

[10] Core, John E., and Wayne R. Guay, 2002, Estimating the Value of Stock Option Portfolios and Their Sensitivities to Price and Volatility, Journal of Accounting Research 40, no. 3 (June), pp. 613-640

[11] Coval, Joshua D., and Tyler Shumway, 2005, Do Behavioral Biases Affect Prices?, Journal of Finance (forthcoming)

[12] de Meza, David, and David C. Webb, 2005, Incentive Design Under Loss Aversion, (July)

[13] Dittmann, Ingolf, and Ernst. Maug, 2006, Lower Salaries and No Options? On the Optimal Structure of Executive Pay, Journal of Finance, no. forthcoming (July)

[14] Dunn, L., 1996, Loss Aversion and Adaptation in the Labor Market: Empirical Indifference Functions and Labor Supply, Review of Economics and Statistics, 78(3), pp. 441-450.

[15] Engel, Ellen; Rachel M. Hayes, and Xue Wang, 2003, CEO Turnover and Properties of Accounting Information, Journal of Accounting and Economics 36, pp. 197-226

[16] Farrell, Kathleen A., and David A. Whidbee, 2003, Impact of Firm Performance Expectations on CEO Turnover and Replacement Decisions, Journal of Accounting and Economics 36, pp. 165-196

[17] Feltham, Gerald A., and Martin G. H. Wu, 2001, Incentive Efficiency of Stock Versus Options, Review of Accounting Studies 6, no. 1 (March), pp. 7-28

[18] Grossman, Sanford J., and Oliver D. Hart, 1983, An Analysis of the Principal-Agent Problem, Econometrica 51, no. 1 (January), pp. 7-45

[19] Hall, Brian J., and Thomas A. Knox, 2004, Underwater Options and the Dynamics of Executive Pay-to-Performance Sensitivities, Journal of Accounting Research 42, no. 2 (May), pp. 365-412

[20] Haigh, Michael S., and John A. List, 2005, Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis, Journal of Finance 49, no. 1 (February), pp. 523-534

[21] Hall, Brian J., and Kevin J. Murphy, 2000, Optimal Exercise Prices for Executive Stock Options, American Economic Review 90, (May), pp. 209-214

[22] ———, 2002, Stock Options for Undiversified Executives, Journal of Accounting and Economics 33, no. 2 (April), pp. 3-42

[23] Haubrich, Joseph G., 1994, Risk Aversion, Performance Pay, and the Principal-Agent-Problem, Journal of Political Economy 102, no. 2 (April), pp. 258-275

[24] Haubrich, Joseph G., and Ivilina Popova, 1998, Executive Compensation: A Calibration Approach, Economic Theory 12, (December), pp. 561-581

[25] Holmström, Bengt, 1979, Moral Hazard and Observability, Bell Journal of Economics 10, pp. 74-91

[26] Jewitt, Ian, 1988, Justifying the First-Order Approach to Principal-Agent Problems, Econometrica 56, no. 5 (September), pp. 1177-1190

[27] Jost, Peter-J., and Florian C. Wolff, 2005, Subjective Expected Utility Theory and the Value and Incentive Effects of Stock Options for Executives

[28] Kahneman, Daniel, and Amos. Tversky, 1979, Prospect Theory: An Analysis of Decision Under Risk, Econometrica 47, no. 2 (March), pp. 263-292

[29] ———, 1984, Choices, Values, and Frames, American Psychologist, 39(4), pp. 341-350.

[30] Kaplan, Steven N, 1994, Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States, Journal of Political Economy 102, no. 3, pp. 510-545

[31] Knetsch, Jack L., 1989, The Endowment Effect and Evidence of Nonreversible Indifference Curves, The American Economic Review 79, no. 5 (December), pp. 1277-1284

[32] Knetsch, Jack L., and A. J. Sinden, 1984, Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value , The Quarterly Journal of Economics 99, no. 3 (August), pp. 507-521

[33] Kouwenberg, Roy, and William T. ZIemba, 2005, Incentives and Risk Taking in Hedge Funds, Mimeo, University of British Columbia, (July)

[34] Lambert, Richard A., and David F. Larcker, 2004, Stock Options, Restricted Stock, and Incentives, Mimeo, University of Pennsylvania, (April)

[35] Lambert, Richard A.; David F. Larcker, and Robert Verrecchia, 1991, Portfolio Considerations in Valuing Executive Compensation, Journal of Accounting Research 29, no. 1 (Spring), pp. 129-149

[36] List, John A., 2004, Neoclassical Theory Versus Prospect Theory: Evidence From the Marketplace, Econometrica 72, no. 2 (March), pp. 615-625

[37] Ljungqvist, Alexander, and William J. Wilhelm, 2005, Does Prospect Theory Explain IPO Market Behavior?, Journal of Finance 49, no. 4 (August), pp. 1759-1790

[38] Margiotta, M.-M., and R. M. Miller, 2000, Managerial Compensation and the Cost of Moral Hazard, International Economic Review 41, no. 3 (August), pp. 669-719

[39] Massa, Massimo, and Andrei Simonov, 2005, Behavioral Biases and Investment, Mimeo, INSEAD

[40] Myagkov, Mikhail, and Charles R. Plott, 1997, Exchange Economies and Loss Exposure: Experiements Exploring Prospect Theory and Competitive Equilibria in Market Environments, American Economic Review 87, no. 5 (December), pp. 801-828

[41] Plott, Charles R., and Kathryn Zeiler, 2005, the Willingness to Pay-Willingness to Accept Ga, the "Endowment Effect," Subject Misconcedptions, and Experimental Procedures for Elicitng Valuations, American Economic Review 95, no. 3 , pp. 530-545

[42] Rabin, Matthew, 2000, Risk Aversion and Expected-Utility Theory: A Calibration Theorem , Econometrica 68, no. 5 (September), pp. 1281-1291

[43] Rogerson, William P., 1985, The First-Order Approach to Principal-Agent Problems, Econometrica 53, no. 6 (November), pp. 1357-1367

[44] Thaler, Richard H., 1980, Toward a Positive Theory of Consumer Choice, Journal of Economic Behavior and Organization, 1, pp. 39-60.

[45] ———, 1999, Mental Accounting Matters, Journal of Behavioral Decision Making 12, pp. 183-206

[46] Tversky, Amos, and Daniel Kahneman, 1991, Loss Aversion in Riskless Choice: A Reference-Dependent Model, The Quarterly Journal of Economics, 106, no. 4 (November), pp. 1039-1061

[47] ———, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty, 5, pp. 297-323.

[48] Weisbach, Michael S., 1988, Outside Directors and CEO Turnover, Journal of Financial Economics 20, pp. 431-460