Executive Stock Options when Managers Are Loss-Averse*

Ingolf Dittmann[†] Ernst Maug[‡] Oliver Spalt[§]

March 15, 2007

Abstract

This paper analyzes optimal executive compensation contracts when managers are loss averse. We establish the general optimal contract analytically and parameterize the model using data on compensation contracts for 916 CEOs. Parameters for preferences are based on the experimental literature. Overall, the Loss Aversion-model dominates an equivalent Expected Utility-model, especially with respect to its ability to predict options as part of the optimal contract. The Loss Aversion-model performs well in terms of predicting observed compensation contract if the reference wage is assumed to lie not too far above previous year's fixed wage. Our results suggest that loss aversion is a better paradigm for analyzing design features of stock options and for developing preference-based valuation models.

JEL Classifications: G30, M52

Keywords: Stock Options, Executive Compensation, Loss Aversion

*We are grateful to seminar participants at the universities Cologne, Mannheim, and Tilburg and to Axel Börsch-Supan and David De Meza for their feedback. We also thank the collaborative research center 649 on "Economic Risk" in Berlin for financial support. Ingolf Dittmann acknowledges financial support from NWO through a VIDI grant.

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

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

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

Executive Stock Options when Managers Are Loss-Averse

Abstract

This paper analyzes optimal executive compensation contracts when managers are loss averse. We establish the general optimal contract analytically and parameterize the model using data on compensation contracts for 916 CEOs. Parameters for preferences are based on the experimental literature. Overall, the Loss Aversion-model dominates an equivalent Expected Utility-model, especially with respect to its ability to predict options as part of the optimal contract. The Loss Aversion-model performs well in terms of predicting observed compensation contract if the reference wage is assumed to lie not too far above previous year's fixed wage. Our results suggest that loss aversion is a better paradigm for analyzing design features of stock options and for developing preference-based valuation models.

JEL Classifications: G30, M52

Keywords: Stock Options, Executive Compensation, Loss Aversion

1 Introduction

In this paper we explain salient features of observed compensation contracts with a simple contracting model where the manager is loss averse. We parameterize this model using standard assumptions and then compare the contracts generated by the model with those actually observed for a large sample of U.S. CEOs. 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 largely based 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. This raises a concern for the widespread application of this model to the valuation of executive stock options and to the analysis of their design (strike price, indexing, reloading, and repricing).[2]

In this paper we suggest a different approach by assuming that managers' preferences exhibit loss aversion as described 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 incremental losses. For brevity, we will refer to all three features as loss aversion and to the corresponding

[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. Inderst and Müller (2003) explain options as instruments that provide outside shareholders with better liquidation incentives. In Oyer (2004) and Inderst and Müller (2003), 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).

[2] Examples on design features include Hall and Murphy (2000), (2002) on the strike price, Meulbroek (2001) on the indexing of strike prices relative to benchmark variables, and Hemmer, Matsunaga, and Shevlin (1998) and Huddart, Jagannathan, and Saly (1999) on reloading.

principal-agent model as the Loss Aversion-model. These assumptions accord with a large body 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.[3] We do not use the notion of decision weights (another feature of prospect theory), 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 in 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 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. The Loss Aversion-model suggests contracts that combine positive option holdings with positive fixed salaries, whereas the Expected Utility-model predicts contracts with much higher stock holdings combined with zero or even negative salaries and option holdings.

We develop this argument in two steps. The first step provides a standard analytic derivation of the optimal contract. We show that under standard assumptions the optimal contract features two parts: above a certain critical stock price the optimal contract always pays off the reference wage of the CEO plus a performance-related part that is represented by an increasing and (mostly) convex function of the stock price. Below this critical stock price compensation falls discontinuously to some lower bound.

The comparative abilities of the Loss Aversion-model Expected Utility-model to explain observed compensation practice is an empirical question. In the second step of our analysis we therefore parameterize both models using assumptions that are based on data and on prior research,

[3] Experimental support for loss aversion is provided by Thaler (1980), Kahneman and Tversky (1984), Knetsch and Sinden (1984), Knetsch (1989), Dunn (1996), and 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. There are also some papers that take a more critical stance. Myagkov and Plott (1997) document that 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.

especially the experimental evidence. Then we calibrate the models using data on 916 CEOs for whom we have complete data. We represent their contracts as consisting of base salaries, stock, and stock options. Then we compute the optimal contract for each CEO for the Loss Aversion-model and for the Expected Utility-model for a range of plausible parameterizations and assess how well each model predicts the observed contract.

It turns out that the performance of the Loss Aversion-model depends critically on the parameterization, especially on the assumed reference wage. If the reference wage is not far above the current base salary (which in our stylized representation also includes most bonus components), then this model predicts observed contracts well. In particular, it can rationalize the use of stock options. If the reference wage is higher and close to total value of the contract, including all options and restricted stock at market values, then the Loss Aversion-model performs poorly. The Expected Utility-model always performs poorly and never predicts options and positive base salaries. Overall, we find that the Loss Aversion-model predicts observed contracts better than the Expected Utility model.

We also calculate the general nonlinear contracts for each CEO in our sample, i.e. we drop the restriction that the contract is piecewise-linear. Above some threshold level, these contracts are mostly convex, and at the threshold level they feature a discontinuous drop to the lowest feasible wage, which is reminiscent of a dismissal. For plausible parameterizations of the Loss Aversion-model we estimate that shareholders would save an additional 0.3% to 3.8% of current compensation costs if they would replace the optimal piecewise linear contract with the optimal nonlinear contract. We conclude that the governance costs of incentive provision through CEO dismissals (with big drops in compensation, i.e. without severance pay) rather than through high-powered wage functions is probably not worth the costs for most companies.

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) study the incentives and investment decisions of hedge-fund managers, and 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 (2007). However, they do not apply their argument to executive compensation contracts and explore a different specification from ours. To the best of our knowledge, ours is the first paper that explores empirically 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. Section 4 develops our empirical methodology in detail. Section 5 analyzes contracts that consist of fixed salaries, stock, and options. Section 6 extends this analysis to general nonlinear contracts. Section 7 concludes. All proofs and derivations are deferred to 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 a wage function $w(P_T)$ that specifies the wage of the manager for a given realization of the company value P_T at time T . Contract negotiations take place at time 0. At the end of the contracting period, T, the value of the firm P_T is commonly observed and the wage is paid according to $w(P_T)$. P_T depends on the CEO's effort e and the state of nature.

The agent's effort e is either high or low, $e \in \{\underline{e}, \overline{e}\}$ so that P_T is distributed with density $f(P_T|e)$. (Later we will also allow for continuous effort.) For notational convenience we write $\Delta e = \overline{e} - \underline{e}$, $\Delta C = C(\overline{e}) - C(\underline{e})$, and $\Delta f(P_T|e) = f(P_T|\overline{e}) - f(P_T|\underline{e})$. We require the monotone likelihood ratio property (MLRP) to hold for f, so $\Delta f(P_T|e)/f(P_T|\overline{e})$ is monotonically increasing in P_T.

Preferences and outside options. Throughout we assume that shareholders are risk-neutral. The manager's preferences are separable in income and effort and can be represented by

$$V\left(w\left(P_T\right)\right)-C\left(e\right), \tag{1}$$

where $C\left(e\right)$ is an increasing and convex cost function, and where we assume preferences over wage income, $w\left(P_T\right)$, of the form[4]

$$V\left(w\left(P_T\right)\right)=\begin{cases}\left(w\left(P_T\right)-w^R\right)^{\alpha} & if \quad w\left(P_T\right)\geq w^R \\ -\lambda\left(w^R-w\left(P_T\right)\right)^{\beta} & if \quad w\left(P_T\right)< w^R\end{cases}, \text{ where } \alpha,\beta<1 \text{ and } \lambda\geq 1. \tag{2}$$

Here, w^R denotes the reference wage. If the payoff of the contract at time T exceeds the reference wage, then the manager codes this as a gain, whereas a payoff lower than w^R is coded as a loss. We will refer to the range of the wage above w^R as the *gain space* and to the range below w^R as the *loss space*. There are three aspects that set this specification apart from standard von Neumann-Morgenstern concave utility specifications. First, the parameter $\lambda\geq 1$ gives a higher weight to payoffs below the reference wage. This reflects the observation from psychology that losses loom larger than gains of comparable size.[5] Formally, this introduces a kink in the value function at w^R and thus locally infinite risk-aversion. This characteristic is also called 'first-order risk aversion' (Segal and Spivak, 1990). Second, the manager treats her income from the firm separately from other sources of income, a phenomenon that is often referred to as "framing" or "mental accounting." See Thaler (1999) for a survey of the evidence on mental accounting. Third, while $V\left(w\left(P_T\right)\right)$ is concave over gains, it is convex over losses. Throughout the remainder of this paper, we will refer to a CEO with preferences of the form (2) as *loss averse* and to the corresponding principal agent-model as the Loss Aversion-model or, for brevity, as the LA-model. We will often compare this model to the Expected Utility model with constant relative risk-aversion, that has

[4]This preference specification was originally proposed by Tversky and Kahneman (1992). It has been introduced into the finance literature by Benartzi and Thaler (1995) and was used by Shumway (1997), Langer and Weber (2001), Berkelaar, Kouwenberg, and Post (2004), and Barberis and Huang (2005).

[5]Rabin (2000) calls loss aversion "the most firmly established feature of risk preferences." For experimental evidence see Tversky and Kahneman (1991) and the references therein as well as McNeil, Pauker, Sox and Tversky (1982), Knetsch and Sinden (1984), Kahneman, Knetsch and Thaler (1986), Tversky and Kahneman (1986), Samuelson and Zeckhauser (1988), Knetsch (1989), Loewenstein and Adler (1995), Post et al. (2007). For evidence from the finance literature see the papers mentioned in the Introduction.

become standard in the literature on executive compensation. The preferences for this model are:

$$V^{EU}\left(w\left(P_T\right)\right) = \frac{\left(W_0 + w\left(P_T\right)\right)^{1-\gamma}}{1-\gamma}\ , \tag{3}$$

where W_0 denotes wealth and γ represents the coefficient of relative risk aversion. We will refer to this model as the Expected Utility-model, or, for brevity, the EU-model. This nomenclature uses general labels to refer to specific, though commonly used parameterizations of each model. Our theoretical analysis focuses on the LA-model only as the EU-model has been analyzed in many places in the literature (see footnote 1 in the Introduction).

We assume that the reference point w^R is exogenous in two respects. First, 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 $w\left(P_T\right)$, which would increase the mathematical complexity of the argument substantially. De Meza and Webb (2007) focus on this aspect of applying loss aversion to principal-agent theory. Second, the reference point is also independent of the level of effort. This is defensible if the cost of effort is non-pecuniary and if 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, $C\left(e\right)$ 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 $E\left[V\left(w\left(P_T\right)\right)\right] - C\left(e\right) \geq \underline{V}$. We assume that there exists a lower bound $\underline{w}$ on the wage function proposed by the shareholders such that $\underline{w} \leq w\left(P_T\right)$ for all P_T and $\underline{w} < w^R$. Such a lower bound arises naturally with limited liability. In the most extreme case, the manager could be required to invest all her wealth in the securities of the firm, but even then her total payoff cannot fall below $-W_0$ in any state of the world, in which case she would lose all her initial wealth.

3 Analysis

3.1 Discrete effort

We characterize the optimal contract $w^*\left(P_T\right)$ under the assumption that effort e is either high or low, $e \in \{\underline{e}, \overline{e}\}$, and that shareholders want to implement the higher level of effort $\overline{e}$. Following the standard principal agent approach as in Holmström (1979), shareholders' problem can then be written as:

$$\min_{w(P_T)\geq \underline{w}} \int w\left(P_T\right) f(P_T|\bar{e})dP_T \tag{4}$$

$$s.t. \int V\left(w\left(P_T\right)\right) f(P_T|\bar{e})dP_T \geq \underline{V} + C\left(\bar{e}\right) \quad , \tag{5}$$

$$\int V\left(w\left(P_T\right)\right) \Delta f(P_T|e)dP_T \geq \Delta C \ . \tag{6}$$

We derive the solution to this problem in Appendix A. The convexity of the agent's preferences over losses turns out to be a major complication. The contract space is not (quasi) convex, so that we can only establish necessary but not sufficient conditions for an optimum. Over the gain space, the optimal solution is characterized by setting up the Lagrangian for this problem and then maximizing it pointwise with respect to $w\left(P_T\right)$. We denote the Lagrange multiplier on the participation constraint (5) by μ_{PC} and the Lagrange multiplier on the incentive compatibility constraint (6) by μ_{IC}.

Proposition 1. ***(Optimal contract):*** *Given the preference structure in (1) and assuming MLRP the optimal contract $w^*\left(P_T\right)$ for the principal agent problem in (4) to (6), is given by:*

$$w^*\left(P_T\right) = \begin{cases} w^R + \left[\alpha\left(\mu_{PC} + \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\bar{e})}\right)\right]^{\frac{1}{1-\alpha}} & if \quad P_T > \widehat{P} \\ \underline{w} & if \quad P_T \leq \widehat{P} \end{cases} , \tag{7}$$

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

Proposition 1 provides us with a general characterization of the optimal contract with a loss-averse manager. For some region $P_T > \widehat{P}$ the optimal contract is continuous, monotonically increasing and pays off only in the gain space. Moreover, the function is convex for $\alpha < 1$ unless the likelihood ratio is concave in P_T and its concavity is sufficiently strong. For $P_T \leq \widehat{P}$ the optimal contract pays off the lowest possible wage $\underline{w}$. The contract also features a discontinuity at $\widehat{P}$ where the manager's wage jumps discretely from $\underline{w}$ to some value $w^*\left(P_T\right) \geq w^R > \underline{w}$.[6] Interestingly, the optimal LA-contract (7) combines punishments ("sticks") with rewards ("carrots"), which sets

[6]De Meza and Webb (2006) find a similar discontinuity in a principal agent model with loss aversion. In their specification, however, the payoff jumps from $\underline{w}$ to w^R and is flat at w^R before it possibly increases continuously. A flat payout at the reference wage w^R occurs if the slope of the line that connects $(0, \underline{w})$ and $(\widehat{P}, w^R)$ is steeper than the slope of the utility function entering the gain space. With the Kahneman and Tversky (1992) value function (2), this cannot occur because the slope entering the gain space is infinite, so that the agent prefers a fair gamble over $\underline{w}$ and $w^R + \varepsilon$ to w^R.

this contract apart from optimal EU-contract that implies only punishments, but no significant rewards.

Equation (7) shows that for the gain space where $P_T > \widehat{P}$ we obtain a result very similar to the familiar Holmström condition (Holmström, 1979) for optimal contracts in the standard concave utility model. This is intuitive, since the problem in the gain space, where preferences are concave, is not fundamentally different from a standard utility-maximizing framework. However, in the loss space the optimal contract features a jump at the reference wage and pays the lower bound $\underline{w}$ for all $P_T \leq \widehat{P}$. While the proof in Appendix A is lengthy, the intuition is straightforward: Since the preferences of the manager are convex over the loss space, any payoff $\underline{w} \leq w(P_T) \leq w^R$ can be replaced by a lottery that pays off $\underline{w}$ with some probability p and $\overline{w} \geq w^R$ with the complementary probability $1 - p$ so that both the incentive constraint and the participation constraint are still satisfied, but expected compensation costs are reduced. We then show that these lotteries are always degenerate for the optimal contract and that the optimal contract pays off $\underline{w}$ for all $P_T \in (0, \widehat{P}]$ and $\overline{w} \geq w^R$ for all $P_T \in (\widehat{P}, \infty)$, where $\widehat{P}$ is a uniquely defined cut-off value. In the final step, we derive the functional form (7) of the optimal contract.

3.2 Continuous effort

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 (6) (more precisely, its analogue for continuous effort) with the first order conditions for (6). 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.[7] In our case, this requires that

$$\frac{\partial^2 E\left(V\left(w\left(P\right)\right)|e\right)}{\partial e^2} = \int V\left(w\left(P_T\right)\right)\frac{\partial^2 f\left(P_T|e\right)}{\partial e^2} dP_T - \frac{\partial^2 C\left(e\right)}{\partial e^2} < 0 \ . \tag{8}$$

This condition will not hold generally. In our setting, one issue is the convexity of the function $V(P_T)$ over the loss space. Moreover, the optimal contract $w(P_T)$ may be convex over some regions of the gain space. However, we can ensure that condition (8) holds for some cost functions C and some density functions in two ways. Firstly, equation (8) shows that this condition will be satisfied

[7] 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).

for sufficiently convex cost functions, so that $\partial^2 C(e)/\partial e^2$ is bounded from below such that (8) holds. Secondly, if the production function $P_T(e)$ is sufficiently concave (such that $\partial^2 P_T(e)/\partial e^2$ is sufficiently small for all effort levels), then (8) will also be satisfied. In the remainder of this paper we will assume that equation (8) holds. The following proposition shows that under this assumption the whole argument of the previous subsection goes through with the same implications for the optimal contract.

Proposition 2. ***(Continuous effort):*** *Assume that the agent's effort is continuous, $e \in [0, \infty)$ and condition (8) holds for each effort level. Then, the results from Proposition 1 continue to hold when the discrete likelihood ratio $\Delta f(P_T|e)/f(P_T|\bar{e})$ is replaced by the continuous ratio $f'(P_T|e)/f(P_T|e)$.*

4 Implementation and Data

4.1 Implementation

The general Loss Aversion-contract. In our empirical implementation, we assume that the stock price follows a lognormal distribution and specify:[8]

$$P_T(u, e) = P_0(e) \exp\left\{\left(r_f - \frac{\sigma^2}{2}\right)T + u\sqrt{T}\sigma\right\} \ , \ \ u \sim N(0, 1) \ , \tag{9}$$

where r_f is the risk-free rate of interest, σ^2 the variance of the returns on the stock, T the time horizon, u is a standard normal random variate and $P_0(e)$ is a strictly increasing and concave function. The expected present value of $P_T(u, e)$ under the risk-neutral density is equal to $P_0 = E[P_T \exp\{-r_f T\}]$.[9] Our assumptions on $P_0(e)$ imply that exerting more effort by the CEO *ceteris paribus* leads to a higher probability that the end-of-period share price will be higher, and that the marginal productivity of effort is decreasing. Note that in any rational expectations equilibrium, P_0 is equal to the market value of equity at the effort level e^* chosen by the manager under the given contract, so $P_0(e^*)$ is equal to the observed market capitalization.

We show in Appendix B that the optimal contract $w^*(P_T)$ for the problem in (4) to (6), can be written as:

[8] This specification ignores dividends for simplicity of exposition. We include dividends in our numerical analysis.

[9] Here and in the following all expectations are taken with respect to the probability distribution of $u \sim N(0, 1)$. Instead of writing $P_T(u, e)$ and $w(P_T(u, e))$ as functions of u we submerge reference to u for ease of exposition.

$$w^*(P_T) = \begin{cases} w^R + (\gamma_0 + \gamma_1 \ln P_T)^{\frac{1}{1-\alpha}} & if \quad P_T > \widehat{P} \\ \underline{w} & if \quad P_T \le \widehat{P} \end{cases}, \tag{10}$$

where γ_0 and γ_1 depend on the two Lagrange multipliers and the (unknown) production function $P_0(e^*)$. $\widehat{P}$ is uniquely defined by the condition:

$$\alpha\left(w^R - \underline{w}\right) = \left(\gamma_0 + \gamma_1 \ln \widehat{P}\right) \lambda \left(w^R - \underline{w}\right)^\beta + (1-\alpha)\left(\gamma_0 + \gamma_1 \ln \widehat{P}\right)^{\frac{1}{1-\alpha}}. \tag{11}$$

Hence, we can represent the nonlinear LA-contract by the coefficients γ_0 and γ_1 and write it as $\mathcal{C}^{LA} = \{\gamma_0, \gamma_1\}$. This specification implies that the contract predicted by the model is strictly increasing in P_T and that it is convex as long as $P_T \le \exp\{\alpha/(1-\alpha) - \gamma_0/\gamma_1\}$. Above this value $w^*(P_T)$ is concave. It is therefore an empirical question whether the contract described in Equation (10) can explain option contracts, because the concave region may or may not be empirically relevant.

Next we identify the parameters that we need to determine in order to analyze the optimal contract numerically. We find appropriate values for the preference parameters α, β, λ, and w^R and for the lower bound of the wage $\underline{w}$ from the experimental literature and from data for executive compensation contracts. The parameters that describe the distribution of P_T in (9) include 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. These can also be determined from available data.

The parameterized model given by equations (10) and (11) contains only two parameters that we cannot determine: γ_0 and γ_1. They depend on the production function $P_0(e)$ and on the cost function $C(e)$, which we both do not model explicitly. We determine these two parameters numerically.

Finding optimal contracts. Our null hypothesis is that the observed contract $w^d(P_T)$ is an optimal contract, so it can be rationalized as the outcome of an optimization program, where we assume that preferences are parameterized as in (1) and that the technology is parameterized as in (9). (The program is specified in Equations (46) to (48) in the appendix.) 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 therefore determine the contract parameters by solving the following program numerically:

$$\min_{\mathcal{C}} \ \pi\left(w(P_T\,|\mathcal{C}\,)\right) \equiv \int w(P_T\,|\mathcal{C}\,)f(P_T)dP_T \tag{12}$$

$$s.t. \ \int V\left(w(P_T\,|\mathcal{C}\,)\right) f(P_T)dP_T \geq \int V(w^d\left(P_T\right))f(P_T)dP_T \ \ , \tag{13}$$

$$\int V\left(w(P_T\,|\mathcal{C}\,)\right) \frac{\partial f(P_T)}{\partial P_0}dP_T \geq \int V(w^d\left(P_T\right))\frac{\partial f(P_T)}{\partial P_0}dP_T \ \ . \tag{14}$$

This program uses a slightly more general notation as we write the wage function as $w(P_T\,|\mathcal{C}\,)$, where $\mathcal{C}$ can refer to different types of contract. For the time being, we only consider $\mathcal{C} = \mathcal{C}^{LA}$. By writing P_T as in (9) and setting $P_0\left(e\right)$ equal to the observed value of the firm, we 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 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 focus only on the first stage by solving program (12) to (14) as it does not depend on knowledge of the cost function $C\left(e\right)$ or of the production function $P_0\left(e\right)$. We therefore do not consider the second stage. 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 (12) to (14) generates a new contract $w^*\left(P_T\right)$ that is less costly to shareholders and specifies the parameters of the optimal contract. Condition (14) 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 (assuming validity of condition (8). Similarly, condition (13) 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 $w^d\left(P_T\right)$ to the optimal contract $w^*\left(P_T\right)$ from (12) to (14).

Piecewise linear contracts Observed contracts consist of salaries, bonus payments, and holdings of corporate securities in addition to many other provisions and perquisites. We simplify the observed contracts by assuming that they only consist of a fixed salary ϕ^d that is paid at time 0, n_S^d shares and n_O^d options, where the total number of shares the company has outstanding is normalized to one. We use the superscript 'd' in order to refer to 'data.' Hence, we write

$$w^d(P_T) = \phi^d e^{r_f T} + n_S^d P_T + n_O^d \max(P_T - K, 0) \ , \tag{15}$$

where K is the strike price of the option. We abstract from other details of observed contracts and consolidate each CEOs portfolio of options into one representative option. The main reason is that different option grants have different maturities and therefore cannot be modelled within the standard one-period principal agent model.

Comparing the non-linear optimal contract $w^*(P_T)$ from the loss aversion model $\mathcal{C}^{LA}$ with the piecewise linear observed contract $w^d(P_T)$ might appear unfair. We therefore also calculate optimal contracts that are restricted to be piecewise linear. For this, we replace the general wage function (10) with the piecewise linear wage function that can be generated by fixed salary ϕ, shares n_S, and options n_O. We will refer to this contract as the piecewise linear LA-contract and denote it by $\mathcal{C}_{Lin}^{LA} = \{\phi^{LA}, n_S^{LA}, n_O^{LA}\}$. The strike price K and the maturity T of the option grant are set equal to the strike price and maturity of the representative option that is estimated form the data.

We also compare the loss aversion model with the expected utility model. We therefore also calculate optimal piecewise linear EU-contracts, which we denote by $\mathcal{C}_{Lin}^{EU} = \{\phi^{EU}, n_S^{EU}, n_O^{EU}\}$. This is the solution to program (12) to (14) with $\mathcal{C} = \mathcal{C}_{Lin}^{EU}$, where the wage function is again linear as in (15) and preferences are given by (3).

Proposition 1 provides only necessary but not sufficient conditions. We therefore solve the optimization problem for different start values in order to find the global optimum. For none of the CEOs in our sample and none of the parameter constellations considered did we find any indication that there is more than one local optimum.

4.2 Data

We identify all CEOs in the ExecuComp database who are CEO at least from January 2004 to December 2005. We also delete all CEOs who where executives in more than one company in either 2004 or 2005 and separately estimate CEOs' contracts in 2005 and in 2004. The 2004 contracts are only needed to construct the reference wage for 2005. We set P_0 equal to the market capitalization at the end of 2004 and take the dividend rate d, the stock price volatility σ^2, and the proportion of shares owned by the CEO n_S from the 2004 data, while the fixed salary ϕ is calculated from 2005 data.[10]

We estimate the option portfolio held by the CEO from 2004 data using the procedure proposed by Core and Guay (2002). We then map this option portfolio into one representative option by first setting the number of options n_O equal to the sum of the options in the option portfolio. Then we determine the strike price K and the maturity T of the representative option such that n_O representative options have the same market value and the same Black-Scholes option delta at the estimated option portfolio. We take into account the fact that most CEOs exercise their stock options before maturity by multiplying the maturity of the individual options in the estimated portfolio by 0.7 before calculating the representative option (see also Huddart and Lang, 1996, and Carpenter, 1998). The maturity T determines the contracting period and the risk-free rate r_f is the U.S. government bond rate from January 2005 with maturity closest to T. After deleting 4 CEOs with stock volatility exceeding 250%, our dataset contains 916 CEOs.

For the EU-model we also 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 from historical data for a subsample of 496 CEOs who have a history of at least five years (as executive of any firm) in the database. We cumulate the CEO's income from salary, bonus, and other compensation payments, add the proceeds from sales of securities, and subtract the costs from exercising options. For this subsample, the median ratio of non-firm wealth to the risk-neutral value of the CEO's pay package (including fixed salary, stock and options) is 0.34. We therefore estimate each CEO's non-firm wealth W_0 by calculating the risk-neutral value of the CEO's pay package and then set W_0 equal to 34% of this value. This procedure sacrifices some accuracy for the breadth of the sample, since the requirements for estimating wealth directly would lead to the loss of more than half of our sample.

[Insert Table 1 here]

Table 1 provides descriptive statistics for all variables in our dataset. The median CEO receives a fixed salary of $1.6m, owns 0.3% of the firm's equity and has options on another 1% of the firm's equity. The median firm value is $2.1bn and the median moneyness K/P_0 is 0.7, so most options are clearly in the money. The median maturity is 4.5 years. The distribution of the contract parameters is highly skewed, so their means are substantially larger than their medians. There are two further parameters we need to estimate in order to complete our calibration: the minimum wage $\underline{w}$, and the reference wage w^R.

[10] ϕ is the sum of the following four ExecuComp data types: Salary, Bonus, Other Annual, and All Other Total. We do not include LTIP (long-term incentive pay), as these are typically not awarded annually.

Minimum wage. We do not have a good theory of the minimum wage in the context of our analysis. Clearly limited liability implies that the minimum wage cannot drop below $-W_0$. In this case, the CEO would be required to invest all her wealth into securities of her own firm and could lose her entire wealth. While this appears rather unrealistic, such a contract is legally clearly feasible. There is also a lot of anecdotal evidence that newly hired executives are asked to invest some of their private wealth into their new company. In our base case, we set the minimum wage $\underline{w}$ therefore equal to $-W_0$. We argue that we should not exclude contracts with negative payouts just because we rarely observe them. Instead, a good model should *generate* contracts with non-negative payouts. Nevertheless, we also repeat our analysis with the minimum wage set equal to zero.

Reference point. Prospect theory does not provide us with clear guidance with respect to the reference point. The reference wage is the wage below which the CEO regards the payments she receives from the company as a loss. We therefore study alternative values for the reference wage. We use a range that is based on the notion that the reference wage reflects expectations the CEO forms based on her previous year's salary. For this reason we look at the previous year's (i.e., 2004) contract of each CEO. It seems natural that the CEO regards a total compensation (fixed and variable) below the fixed salary of the previous year as a loss and we use this as a lower bound. In addition, she may also build in some part of her deferred compensation into her reference wage. Most likely, she will evaluate her securities at a substantial discount relative to their value for a well-diversified investor. This discount depends on her loss aversion and her framing of the wage-setting process. We therefore regard the value of her previous contract based on the current stock price and the number of shares and options she inherited from the previous period as a (rather implausible) upper bound for the reference wage. We denote the value of her deferred compensation in 2005 based on the number of shares and options she held in 2004 by MV and write:

$$w_{2005}^{R}(\theta) = \phi_{2004} + \theta \cdot MV(n_{2004}^{S}, n_{2004}^{O}, P_{2005}) \; , \tag{16}$$

The parameter θ is an index of the discount the CEO applies to her deferred compensation. If $\theta = 0$, then the reference wage for 2005 equals her base salary for 2004. If $\theta = 1$, then the reference wage equals the market value of her total compensation in the previous year, valued at current market prices and without a discount for risk. We will look at a grid of alternative values for θ.

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]

5 Contracts with restricted stock and options

5.1 Loss Aversion vs. Expected Utility

We now describe piecewise linear contracts predicted by the LA-model and compare them to the contracts predicted by the standard EU-model. For our base case we assume that option awards n_O and base salaries ϕ can be negative and are restricted only by the number of shares the manager owns ($n_O > -n_S \exp(dT)$) and, respectively, by the manager's wealth ($\phi > -W_0$). For each CEO, we compare the observed contract $\mathcal{C}^d = \{\phi^d, n_S^d, n_O^d\}$ with the optimal piecewise linear contract $\mathcal{C}^M = \{\phi^M, n_S^M, n_O^M\}$, where the superscript M denotes the contracts predicted by model $M \in \{EU, LA\}$. Minimization of program (12) to (14) is subject to the additional constraints $n_O > -n_S \exp(dT)$ and $\phi > -W_0$.

We use two methods to compare contracts predicted by the models to observed contracts. The first method relies on the following metric D_{Lin} that measures the average distance between optimal contracts and observed contracts:

$$D_{Lin} = \frac{1}{N}\sum_{i=1}^{N}\left[\underbrace{\left(\frac{\phi^* - \phi^d}{\sigma_\phi}\right)}_{error(\phi)}^2 + \underbrace{\left(\frac{n_S^* - n_S^d}{\sigma_S}\right)}_{error(n_S)}^2 + \underbrace{\left(\frac{n_O^* - n_O^d}{\sigma_O}\right)}_{error(n_O)}^2\right]^{1/2}, \qquad (17)$$

$$where: \sigma_S = \frac{1}{N}\sum_{i=1}^{N}\left(n_S^{d,i} - \bar{n}_S^d\right)^2, \quad \sigma_O = \frac{1}{N}\sum_{i=1}^{N}\left(n_O^{d,i} - \bar{n}_O^d\right)^2,$$

$$\sigma_\phi = \frac{1}{N}\sum_{i=1}^{N}\left(\phi^{d,i} - \overline{\phi}^d\right)^2.$$

Here summation is over all $N = 916$ CEOs in the sample and arithmetic means over all CEOs are denoted by a bar. This metric measures the distance between the observed contracts and the model contracts and gives more weight to those parameters that have lower cross-sectional dispersion. A similar approach was used in Carpenter (1998) and Bettis, Bizjak, and Lemmon (2005). Our second method simply compares the salient features of observed contracts, specifically, the fact that almost

[11] See Tversky and Kahneman (1992). These values have become somewhat of a standard in the literature e.g. Benartzi and Thaler (1995), Shumway (1997), Langer and Weber (2001), Berkelaar, Kouwenberg and Post (2004), Barberis and Huang (2005). For experimental studies on the preference parameters which yield parameter values in a comparable range see Abdellaoui (2000) and Abdellaoui, Vossmann, and Weber (2005).

all contract feature positive base salaries and positive option holdings, and then asks to what extent the models can replicate these salient features.

[Insert Table 2 here]

Table 2 Panel A summarizes the results for the LA-Model for seven different levels of the reference wage as parametrized by θ (see (16)). Panel B shows the results for the EU-model for six values of the coefficient of relative risk-aversion γ. For each model we show the means and medians of the contract parameters and the scaled mean deviations of the contract parameters of the models from their observed counterparts (referred to as errors in Equation (17)).

Both models tend to underpredict base salaries and options while overpredicting stock. A notable exception is the LA-model for intermediate values of w^R ($\theta = 0.2$ and 0.4), where predicted salaries are on average higher than observed salaries. The results for the EU-model are less sensitive to the parametrization. Here base salaries and option holdings are generally substantially below those for the LA-model and also below those of the observed contracts.

The results for shareholdings mirror those of option holdings: lower option holdings are always matched by higher holdings of the firm's stock. This follows from the incentive compatibility constraint (6) for both models, which ensures that overall incentives correspond to those of the observed contract. Then lower option holdings imply higher shareholdings. Whenever the model replaces some options with stock, then, in the EU-model, the contract becomes more valuable to the CEO since options are worth less than the corresponding number of shares that generate the same incentives. Hence, replacing options with stock implies that base salaries decrease in the EU-model. In the LA-model, this is not generally true. Here, the manager might attach a higher value to incentives provided by options than to incentives provided by stock, because options potentially insulate her from losses.

Our results for the distance metric D_{Lin} in Table 2 depend strongly on the parameterization of each model, and do so more than on the model type (EU or LA) itself. For the EU-model, the lowest distances between observed and model contracts occur for the highest levels of risk aversion γ, because high risk aversion reduces the subjective value of stock and options, so that fixed salaries do not decrease so much as options are replaced by stock. On the other hand, the EU-model also becomes (slightly) more accurate as the risk-aversion γ decreases and converges to zero. This reflects the fact that any observed contract is optimal (i.e. cost minimizing) if the agent is risk-neutral ($\gamma = 0$), because subjective values are then identical to market values. We do not consider values of γ below 0.1 in Table 2 as they as they lead to implausible implications and are never used in the literature on executive compensation.[12] High levels of risk aversion lead to even more concave contracts with negative option holdings (or zero option holdings if negative holdings are not permitted).

The LA-model generates consistently better forecasts for low reference wages than for high reference wages. The reason is that options can limit losses only if the reference wage w^R is lower than the (comparatively high) wage that is associated with option pay. If the reference wage is higher, then any stock price P_T below the option's strike price K will lead to a wage w_T that is coded as a loss; then a contract with a lower wage, less options and more stock might be more desirable. Consequently, options are used more often (and salaries are higher) if the reference wage is low.

This argument does not hold as the reference wage decreases further and further, because eventually any feasible contract will only pay out in the gain space. As the manager is slightly risk-averse in the gain space, the optimal contract would then contain no options and only stock (assuming that this is feasible) just as in the EU-model with a low value of γ. This is the reason why the LA-model predicts the largest option holdings for $\theta = 0.1$, where it is also most accurate.

A comparison of the distance measure D_{Lin} between the two models shows that the LA-model with low reference wages w^R predicts contracts that are closer to observed contracts than any parameterization of the EU-model in Table 2 and parameterizations with a reference wage where $\theta \geq 0.4$ seem to generate poor predictions for the model parameters. By comparison, the EU-model generates poor predictions of the compensation parameters throughout and generates the lowest value for the distance metric only for high values of γ. Both models predict more concave rather than convex contracts for higher levels of the reference wage, respectively, higher levels of risk aversion. Hence, if we evaluate the models in terms of their ability to predict convex contracts that can rationalize the use of options we would prefer parameterizations with lower reference wages, respectively, lower levels of risk aversion.

[12] The literature on executive compensation has often discussed values for γ in the range between 2 and 3. Hall and Murphy (2000) use these values that seem to go back to Lambert, Larcker, and Verrecchia (1991). Lambert and Larcker (2004) more recently proposed a value as low as 0.5. A useful point of reference here is the portfolio behavior of the CEO, since very low levels of risk aversion (below 1) imply that CEOs have implausibly highly leveraged investments in the stock market. Ingersoll (2002) develops a parameterization of the EU-model that is sufficiently similar to ours but includes investments in the stock market. Using his equation (8) and assuming a risk premium on the stock market as low as 4% and a standard deviation of the market return of 20% gives an investment in the stock market (including exposure to the stock market through holding securities in his own firm) equal to $1/\gamma$. E.g., $\gamma = 0.1$, the lowest value considered in Table 2, would imply that the CEO invests ten times her wealth in the stock market. We do not wish to take such a restrictive stance in order not to bias our analysis in favor of the LA-model and therefore allow for levels of risk aversion as low as 0.1, even though we regard such values as implausible.

As risk preferences and behavior of CEOs vary widely as the model parameters change, it does not seem warranted to compare *all* parametrization of the LA-model with *all* parameterizations of the EU-model. Instead, it seems more sensible to compare the two models based on *comparable* parameterizations, i.e. we need an additional variable that we hold constant across models. We propose to compare parameterizations that imply the same overall attitude to risk. More specifically, we compare parameterizations that generate the same valuation of the observed contract by the same CEO. We define the certainty equivalent of model M from: $E\left(V^M\left(w^d\left(P_T\right)\right)\right) = CE^M$. We fix θ to establish the reference wage of each CEO and then define an equivalent γ from

$$CE^{LA}(\theta) \equiv CE^{EU}(\gamma_e) \ . \tag{18}$$

We refer to the value of γ_e that satisfies (18) as the equivalent degree of relative risk aversion, because it holds the certainty equivalent constant. A straightforward implication of this step is that we also hold the risk premium with respect to the observed contract constant for both models. For each CEO and for each θ we calculate the equivalent γ_e and the optimal EU-contract with $\gamma = \gamma_e$. Table 3 compares the two models.

[Insert Table 3 here]

Table 3, Panel A reports the mean and the median difference $D_{Lin}^{EU} - D_{Lin}^{LA}$ of the distance metric D_{Lin} between the two models. We test whether this difference is significantly different from zero with the standard t-test, the Wilcoxon signed rank test, and the sign test. The distribution of D_{Lin} is skewed, so we sometimes obtain conflicting results for the means and for the medians. The equivalent γ_e's are generally very low and below the range we regard as plausible (see footnote 12). We note also that they are non-monotonic in θ: Larger reference wages move more and more probability mass into the loss space, so that the risk aversion at the reference wage becomes less important.

For low and medium size reference wages ($\theta < 0.5$) the LA-model dominates the EU-model, for $\theta \leq 0.2$ even for more than 90% of all CEOs in our sample. For $\theta = 0.5$ and $\theta = 0.6$ the two models are equally successful if judged by the sign of the difference in accuracy, and for larger values of θ the EU-model performs better than the LA-model, but the mean difference becomes significant only for very high reference wages ($\theta \geq 0.9$). The EU-model dominates the LA-model in most cases for $\theta > 0.6$. However, when it fails, then its failure is often more extreme than that of the LA-model, which gives rise to a more skewed distribution of D_{Lin} and hence insignificant t-statistics.

Table 3, Panel B compares the models in terms of the overall features of the contract parameters. We report the proportion of predicted contracts that feature positive option holdings or positive salaries for both models. The LA-model is more successful in explaining positive options and positive salaries for all parameter choices, whereas the proportion of CEOs for which the EU-model predicts positive option holdings *and* positive salaries is virtually zero. The model exchanges options for more stock and lower salaries until either the restriction on salaries ($\phi > -W_0$) or the restriction on option holdings ($n_O > -n_S \exp(dT)$) bind. Therefore this model can *never* explain positive option holdings *and* positive salaries simultaneously, while 99.8% of the CEOs in our sample have such a contract (two CEOs have zero fixed salaries and zero option holdings). In contrast, the LA-model can explain contracts with both, options and positive salaries for up to 72% of the CEOs in our sample.

The discussion of Tables 2 and 3 leads to the conclusion that both models perform poorly if we assume high levels of the reference wage for the LA-model, and high levels of risk aversion for the EU-model. Then both models predict concave contracts and generate large values of the distance metric D_{Lin}. The performance of the LA-model depends more strongly on the assumed reference wage, whereas the performance of the EU-model is more consistent, even though it also depends on assumed parameter values. Figure 1 illustrates this for the case of a typical CEO. The figure shows the optimal LA-contract, the optimal EU-contract and the observed contract for the same CEO for $\theta = 0.1$ and for $\theta = 0.8$. For $\theta = 0.1$, the LA-contract and the observed contract are virtually indistinguishable with a value of $D^{LA}_{Lin} = 0.70$ for the distance metric. By contrast, the corresponding EU-contract is concave and differs substantially from the observed contract, which is reflected in a much higher value of the distance metric of $D^{EU}_{Lin} = 49.57$. Clearly, the LA-model performs well with this parameterization even though it overpredicts the fixed salary, whereas the EU-model performs poorly. For $\theta = 0.8$, both models perform poorly, but the deterioration is much stronger for the LA-model than it is for the EU-model, which now compares favorably in terms of the distance metric because it predicts a less negative base salary, even though it predicts even lower option holdings.

[Insert Table 4 here]

Table 4 shows a similar comparison of the two models as Table 3 if we impose the stricter constraints that salary and option holdings cannot become negative, i.e. $\phi \geq 0$ and $n_O \geq 0$. Table

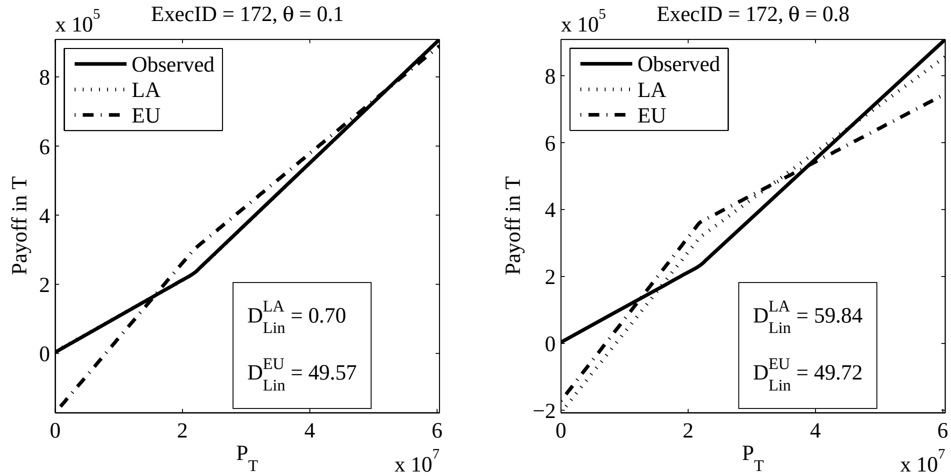

Figure 1: The figure shows the observed contract, the LA-contract and the EU-contract for the CEO with ExecID # 172 for $\theta = 0.1$ and $\theta = 0.8$. Observed contract parameters are $\phi^d = 2.92$m, $n_S^d = 1.04\%$, and $n_O^d = 0.70\%$. For $\theta = 0.1$ (left panel) the EU-model predicts $\phi = -141.92$m, $n_S = 2.15\%$, and $n_O = -0.64\%$, whereas the LA-model predicts $\phi = 4.460$m, $n_S = 1.02\%$, and $n_O = 0.72\%$. For $\theta = 0.8$ (right panel) the EU-model predicts $\phi = -141.92$m, $n_S = 2.42\%$, and $n_O = -1.45\%$, whereas the LA-model predicts $\phi = -171.5$m, $n_S = 2.38\%$, and $n_O = -1.00\%$.

4 Panel B reveals that the EU-model is still not able to generate positive salaries *and* positive option holdings, because one of the two new constraints always binds: Either option holdings are equal to zero, then salary is positive, or the predicted salary is zero and option holdings are positive. With the stricter restrictions on ϕ and n_O, the EU-model is therefore comparatively more successful at predicting *either* positive option holdings *or* positive salaries. However, it can never predict both. The median values in Table 4 show that for most CEOs the LA-model still dominates the EU-model for almost all values of the reference wage w^R. Nonetheless, the accuracy of the EU-model increases and is higher than the accuracy of the LA-model on average, as shown by the mean values in Panel A of Table 4.

We conclude that the EU-model is not able to generate positive salaries or positive option holdings - except if we impose this as a restriction in the maximization problem. In contrast, the LA-model predicts positive option holdings for 89% of all CEOs and positive salaries for 76% of all CEOs if $\theta = 0.1$ (see Table 3 Panel A).

5.2 Agents and Owners: The Limitations of the LA-Model

The accuracy of the LA-model is comparatively high for low reference wages w^R, but it deteriorates markedly as the reference wage increases (see Table 2 Panel A). In order to understand this effect, we regress the distance between the predicted LA-contract and the observed contract, D_{Lin}^{LA}, on the observed stock holdings, the observed option holdings, and an intercept. Table 5 shows the results.

[Insert Table 5 here]

The Table exhibits that the accuracy of the LA-model is significantly positively correlated with observed stock holdings and not significantly correlated with observed option holdings. So the LA-model seems to be particularly bad for those CEOs who own a large proportion of shares in their company. We therefore split the sample into a subsample with the 97 owner-executives who own 5% or more of the shares of their firm and a subsample with the remaining 819 CEOs who own less than 5%. Table 6 displays the parameters of the optimal contracts predicted by the LA-model for the two subsamples; it is a breakdown of the results shown in Table 2 Panel A.

[Insert Table 6 here]

Comparing the results in Table 6 with those for the LA-model in Table 2 shows that for CEOs with an ownership of less than 5% of their companies the LA-model performs more consistently and is more precise overall. By comparison, the right part of Table 6 demonstrates that the LA-model is very imprecise for owner-executives and the discrepancy between observed and predicted contracts are large.

The fact that the LA-model cannot generate the contracts of owner-managers is not necessarily an argument against this model. Like any other simple principal-agent model, it aims at describing the contract of a salaried agent who is employed by outside shareholders and who needs to be incentivized to provide adequate effort. If the CEO owns a large block of shares herself, her objectives and the bargaining process between principal and agent are likely to be much different from the situation modelled in a simple principal-agent model. On the other hand, our model works very well for non-owner CEOs for whom the principal-agent paradigm appears plausible, or about 90% of our sample.

6 General Non-Linear Loss-Aversion Contracts

Our analysis in the previous section relies on a stylized representation of the contracts. Our theoretical analysis above shows however that the optimal contract is highly non-linear. In this section we describe and analyze the optimal nonlinear contracts generated by the loss-aversion model in order to gain a better understanding of the advantages and disadvantages of this model.

One feature of the optimal nonlinear contract in the LA-model is the discrete jump at the point $\hat{P}$ from $\underline{w}$ to some number above w^R. This jump is caused by the fact that the LA-contract pays out above the reference wage w^R as long as possible: if it pays out below this threshold it therefore pays the lowest possible wage $\underline{w}$. This drop can be interpreted as a dismissal of the manager, and we will also use the word "dismissal" in the tables for brevity. In practice however, dismissals do not always generate a discontinuity in the payoff function, for example when managers receive sufficient severance pay to compensate them for their loss of compensation and it is not possible to calibrate this discontinuity to the data. In order to describe the optimal non-linear contract, we define the dismissal probability p of the optimal model contract as $p(\hat{P}) \equiv \int_0^{\hat{P}} f(P_T)\, dP_T$.

The conceptual difficulty in comparing general nonlinear contracts to the data lies in the fact that contracts like (10) cannot be implemented using a few standard securities like shares and options. In principle they could be approximated with a sufficiently large number of options with different strike prices, providing that the contract is convex. However, the general nonlinear loss-aversion contract (10) also has a range where it is concave, and the concave part can be approximated with options only if we allow for negative option holdings. Another limitation arises from the fact that the observed contracts described in Table 1 above reduce the much more complex contracts observed in practice to a stylized representation in terms of base salaries, stock, and one option grant.

We address these issues by developing some heuristics that allow us to compare the model contracts to the observed contracts. In particular, we look at the average slopes of the nonlinear contract. We define:

$$\Delta_{Low} \equiv \int_0^K \frac{\partial w^*(P_T)}{\partial P_T} \frac{f(P_T)}{F(K)} dP_T \ , \tag{19}$$

$$\Delta_{High} \equiv \int_K^\infty \frac{\partial w^*(P_T)}{\partial P_T} \frac{f(P_T)}{1 - F(K)} dP_T \ . \tag{20}$$

Here Δ_{Low} is the average slope in the region below the strike price of the option, which can be compared to the number of shares n_S. Δ_{High} is the average slope in the region above the strike price and can be compared to shares and options combined. In addition, we are also interested in the convexity and the concavity of the optimal contracts. From (10) we can determine the inflection point P_I of each contract, so that the contract is convex for all terminal stock prices below P_I. We use the probability that the predicted contract pays off in the convex range, $\Pr(w^*(P_T) \leq P_I)$ as another descriptive statistic.[13]

[Insert Table 7 here]

Table 7 reports the average slopes Δ_{Low} and Δ_{High}, the dismissal probability, and the quantile of the inflection point for different parameterizations. We also report the percentage of those CEOs where the nonlinear contract accommodates positive option holdings, defined by the condition that $\Delta_{High} > \Delta_{Low}$, which also measures convexity. We can see that the contracts predicted by the LA-model are mostly convex by both measures of convexity. The slope in the upper range, Δ_{High} is almost always higher than the slope in the lower range, Δ_{Low}. Similarly, almost all of the probability mass for this contract lies to the left of the inflection point, rendering the concave part of the contract irrelevant.

The dismissal probabilities are unrealistically high for the LA-model once the reference point becomes sufficiently high ($\theta-$values above 0.4). This aspect underlines our earlier assessment that high parameterizations with high reference wages lead to poor performance of the LA-model. As the reference wage increases, the threat of dismissals becomes more important. Intuitively, CEOs with a higher reference wage demand a higher compensation, and they receive it in the sense that their compensation while they are employed is larger. However, then incentives are provided to a lesser extent through the slope of the wage function (note how Δ_{Low} declines as w^R increases) and to a larger extent through the threat of dismissals.

[Insert Table 8 here]

The non-linearity of the general LA-contract raises the question how costly it is to restrict the contract shape to being piecewise linear, i.e. implementable by fixed salary, stock and one option grant. In Table 8, we therefore compare the optimal non-linear contract (10) with the optimal piecewise linear contract (**??**). For both contracts, the table shows the average slopes Δ_{Low} and Δ_{High} and the distance metric D_{NonLin}:[14]

[13] There are some CEOs where $P_I \leq \hat{P}$, so the LA-contract has a slope of zero up to the discontinuity and then becomes concave. For these CEOs we calculate $\Pr\left(w^*(P_T) \leq \hat{P}\right)$.

[14] Note that for the piecewise linear contract, $\Delta_{Low} = n_S$ and $\Delta_{High} = n_S + n_O$.

$$D_{NonLin} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(\frac{\Delta_{Low}^{*,i} - \Delta_{Low}^{d,i}}{\sigma_{Low}}\right)^2 + \left(\frac{\Delta_{High}^{*,i} - \Delta_{High}^{d,i}}{\sigma_{High}}\right)^2} \tag{21}$$

$$where: \sigma_{Low}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\Delta_{Low}^{d,i} - \bar{\Delta}_{Low}^{d}\right)^2 \quad , \quad \sigma_{High}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\Delta_{High}^{d,i} - \bar{\Delta}_{High}^{d}\right)^2 \ .$$

Here, $\Delta_{Low}^{d,i}$ and $\Delta_{High}^{d,i}$ represent the slopes of the observed contract corresponding to (19) and (20) and $\bar{\Delta}_{Low}^{d}$ and $\bar{\Delta}_{Low}^{d}$ refer to their sample averages. Note that D_{NonLin} does not contain any fixed salary, which is not a meaningful concept in the context of general nonlinear contracts. In addition, Table 8 shows how much shareholders could save (as a proportion of total observed compensation) if they could recontract and replace the observed contract with the contract predicted by the models. These savings from recontracting are defined as

$$Savings = \frac{E\left(w^d\left(P_T\right)\right) - E\left(w^*\left(P_T\right)\right)}{E\left(w^d\left(P_T\right)\right)} \ , \tag{22}$$

or, in words, the percentage reduction in the costs of the optimal predicted contract compared to those of the observed contract. These savings are effectively what is maximized when our algorithm searches for the optimal contract.

We can see that the accuracy (i.e. the average $-D_{NonLin}$) of the general contract is always higher than the accuracy of the piecewise linear contract. For low reference wages the difference is small, but it increases as the reference wage increases. By construction, the savings relative to the status-quo of the optimal general contract are higher than the savings of the piecewise linear contract.

The savings are not substantial for either version of the contract. This is important, because it shows that even where the distance between the observed contracts and the predicted contracts appears large in terms of the metrics developed above, the savings are insubstantial, particularly for the piecewise linear contract. Hence, replacing the observed contract with the model contract would generate negligible benefits for shareholders. Indirectly, this fact lends support to the model.

Implementing a better approximation of the nonlinear contract would not only require a broad portfolio of options with a range of strike prices (which most CEOs have), but it would also require to implement the sharp drop in compensation at the threshold $\hat{P}$. We do not observe such drops in actual contracts. A possible reason is that such a contract would need tight monitoring of the CEO in order to enforce dismissals and the corresponding fall in wages. Another caveat is that *ex post* the board might have additional information and might know that the bad performance is not caused by the CEO's low effort. If it is not possible to contract on this information, a wage function that features potential dismissals might not be credible. The difference in savings between the piecewise linear contract and the general nonlinear contract would have to be related to the costs of writing and implementing such a contract. These additional savings are 0.3% for $\theta = 0$. (Subtract the 0.18% for the linear contract from the 0.5% for the nonlinear contract in Table 8.) Based on our earlier discussion, the LA-model is a candidate for explaining observed compensation practice only if we assume relatively low reference wages, and then the additional savings from implementing the more complex contract with dismissals are low ($\leq 3.8\%$ for $\theta \leq 0.4$). We conclude that incentive provision through CEO dismissals with big drops in compensation rather than high-powered wage functions is probably not worth the associated costs for most companies.

6.1 Robustness Checks

In this section, we investigate to what extent our results are sensitive to parameter assumptions. We have based our discussion on the estimates of α, β, and λ on the experimental literature, and this may well be inappropriate for the study of CEOs.

[Insert Table 9 here]

Table 9 reports the results of a comparative static analysis in terms of the preference parameters where the reference wage w^R is set to last year's fixed salary (i.e. $\theta = 0$). We report the same parameters as in the previous table. The results for the piecewise linear model are hardly affected by changes in β and λ, whereas higher values for α are associated with lower values for D_{NonLin} and also substantially lower savings. As α increases, CEOs become increasingly more risk-neutral, and we observed before that this trivially improves the fit of the model. Overall, it seems safe to conclude that none of our qualitative conclusions is affected by our particular choice of model parameters.

7 Conclusion

We develop a principal agent model with a loss-averse agent in order to explain observed executive compensation contracts. We derive the optimal contract and show 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. We parameterize this model in a way that is standard in the literature and calibrate it to observed contracts.

We find that the Loss Aversion-model performs better in several respects in comparison to the Expected Utility-model:

- For reasonable parameter levels (low reference wages), the Loss Aversion-model predicts observed contracts more closely compared to the Expected Utility-model based on our tests.
- For other parameterizations with high reference wages (respectively, higher levels of risk aversion), both models predict concave contracts with negative option holdings.
- The Loss Aversion-model can explain the prevalence of stock options in observed compensation contracts. It generates interior solutions for option holdings and positive base salaries for realistic parameterizations, whereas the Expected Utility-model does not.
- The Expected Utility-model comes close to the Loss Aversion-model only if we impose constraints on base salaries (which this model cannot generate endogenously).

Our results are of particular importance to the substantial literature on the design and the valuation of executive stock options that relies so far on variants of the Expected Utility-model (see footnote 1 in the introduction). We argue that for these applications the Loss Aversion-model is more useful than the Expected Utility-model. Our analysis also gives some guidance regarding relevant ranges of the reference wage: Predicted contracts most closely resemble observed contracts for relatively low reference wages that are set close to the previous fixed salary (including bonus payments).

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. We have not investigated this alternative specification as it would not allow us to compare the Loss Aversion-model to the Expected Utility model on an equal footing. We conjecture that the implications for our analysis would be minor and then our results would apply to the structure of deferred, incentive-related compensation only.

A Appendix

Proof of Proposition 1: We prove the proposition in three steps. In the first step, Lemma 1 shows that the contract never pays out in the interior of the loss space. So whenever the agent realizes a loss, it will be the largest possible loss $\underline{w}$. Lemma 2 then shows that the optimal contract pays out $\underline{w}$ for all realized stock prices below some threshold. If the stock price exceeds this threshold, the contract always pays out wages that are perceived as gains by the agent. Lemma 2 greatly reduces the set of contracts from which we have to find the optimal contract. In the third step, we write down the Lagrangian for the simplified problem and derive the solutions to the first-order condition.

For Lemma 2, we extend the set of permissible contracts to contracts that pay out lotteries. The agent is risk-seeking over losses, so lotteries might be part of the optimal contract. Lemma 2 shows, however, that the optimal contract does not contain lotteries.

Lemma 1. ***(Lotteries):*** *Consider a contract $w(P_T)$ that, for some realized stock price P_T, pays off w' in the interior of the loss space with some positive probability, such that $\underline{w} < w' < w^R$. Then there always exists an alternative contract that improves on the contract $w(P_T)$ where the manager receives instead of w' the reference wage w^R with probability g and the minimum wage $\underline{w}$ with the remaining probability $1-g$.*

Proof of Lemma 1: We first consider a contract that pays off in the interior of the loss space with certainty. Consider the proposed candidate contract $w(P_T)$ that pays off $\underline{w} < w(P_T) < w^R$ at some price P_T with certainty. Since $\lambda\left(w^R - w(P_T)\right)^\beta$ is monotonically decreasing in $w(P_T)$, we have $\lambda\left(w^R - w^R\right)^\beta < \lambda\left(w^R - w(P_T)\right)^\beta < \lambda\left(w^R - \underline{w}\right)^\beta$. Hence, there exists a unique number $g(P_T)$ for each $w(P_T) \in \left[\underline{w}, w^R\right]$ such that

$$g(P_T)\,\lambda\left(w^R - w^R\right)^\beta + \left(1 - g(P_T)\right)\lambda\left(w^R - \underline{w}\right)^\beta = \lambda\left(w^R - w(P_T)\right)^\beta \ . \tag{23}$$

This implies that replacing the payoff $w(P_T)$ with the lottery $\left\{g(P_T), w^R; 1-g(P_T), \underline{w}\right\}$ leaves the participation constraint (5) and the incentive compatibility constraint (6) unchanged. From the concavity of U_l we also have:

$$g(P_T)\,\lambda\left(w^R - w^R\right)^\beta + \left(1 - g(P_T)\right)\lambda\left(w^R - w(P_T)\right)^\beta \le \lambda\left(w^R - \left(g(P_T)\,w^R + \left(1-g(P_T)\right)\underline{w}\right)\right)^\beta \ . \tag{24}$$

Combining equations (23) and (24) yields:

$$\lambda\left(w^R - w(P_T)\right)^\beta \le \lambda\left(w^R - \left(g(P_T)\,w^R + \left(1-g(P_T)\right)\underline{w}\right)\right)^\beta \ . \tag{25}$$

U_l is increasing in its argument and therefore decreasing in $w(P_T)$, therefore $g(P_T)\,w^R + (1-g(P_T))\,\underline{w} \leq w(P_T)$, so the lottery $\left\{g(P_T), w^R; 1-g(P_T), \underline{w}\right\}$ improves on the original contract $w(P_T)$. 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 $\left\{g(P_T)\,p, w^R; (1-g(P_T))\,p, \underline{w}\right\}$. $\square$

Note that due to the concavity of the agent's preferences over gains, lotteries among payouts in the gain space are never optimal.

Lemma 2. ***(Shape of the loss space):*** *There exists a uniquely defined cut-off value $\widehat{P}$ such that the optimal contract $w^*(P_T)$ pays out in the loss space for all $P_T \leq \widehat{P}$ and in the gain space for all $P_T > \widehat{P}$. When the contract pays out in the loss space, it always pays the minimum feasible wage: $w^*(P_T | P_T \leq \widehat{P}) = \underline{w}$.*

Proof of Lemma 2: According to Lemma 1, we can represent the optimal contract by three functions: $\widetilde{w}(P_T) = (g(P_T), \overline{w}(P_T), \underline{w}(P_T))$, where $\overline{w}(P_T) \geq w^R$ and $\underline{w}(P_T) = \underline{w}$ are non-random wage functions and $g(P_T) \in [0,1]$ is the probability that $\overline{w}(P_T)$ is paid. With probability $1 - g(P_T)$ the wage $\underline{w}(P_T)$ is paid.

We prove Lemma 2 by contradiction. If there is no cut-off value that separates the loss space from the gain space, then there exists a unique point $\widetilde{P} \in [0, \infty)$ such that the probability that the contract pays out in the gain space below $\widetilde{P}$ is positive and equal to the probability that the contract pays out in the loss space above $\widetilde{P}$. More formally:

$$\textstyle\int_0^{\widetilde{P}} g(P_T) f(P_T|\overline{e}) dP_T = \int_{\widetilde{P}}^{\infty} (1 - g(P_T)) f(P_T|\overline{e}) dP_T =: s > 0. \tag{26}$$

$\widetilde{P}$ exists because the distribution of P_T is continuous. We then construct an alternative contract, where we exchange the "wrong" gains to the left of $\widetilde{P}$ with the "wrong" losses to the right of $\widetilde{P}$. More precisely, we replace the gains below $\widetilde{P}$ by the lowest possible loss $\underline{w}$, and all losses above $\widetilde{P}$ by a constant payout in the gain space that is chosen such that the costs of the two contracts to the firm are identical. This constant payout is equal to the expected payout across the "removed" gains below $\widetilde{P}$. We then show that this alternative contract strictly relaxes the participation constraint and the incentive compatibility constraint. This implies that the agent is better off with the new contract and has stronger incentives to exert high effort. This alternative contract is obviously not optimal, but its existence shows that the initial contract cannot be optimal.

Consider the alternative contract $\widetilde{w}'(P_T) = (g'(P_T), \overline{w}'(P_T), \underline{w}'(P_T))$ which is defined as follows:

$$g'(P_T) = g(P_T) \tag{27}$$

$$\overline{w}'(P_T) = \begin{cases} \underline{w}, & \text{if } P_T \leq \widetilde{P} \\ \overline{w}(P_T), & \text{if } P_T > \widetilde{P} \end{cases} \tag{28}$$

$$\underline{w}'(P_T) = \begin{cases} \underline{w}(P_T) = \underline{w}, & \text{if } P_T \leq \widetilde{P} \\ \frac{1}{s}\int_0^{\widetilde{P}} g(P_T)\overline{w}(P_T) f(P_T|\overline{e}) dP_T \geq w^R, & \text{if } P_T > \widetilde{P} \end{cases} \tag{29}$$

By construction, the costs of $\widetilde{w}(P_T)$ and $\widetilde{w}'(P_T)$ are identical for the principal. In the remaining part of the proof, we show that the new contract $\widetilde{w}'(P_T)$ relaxes both, the participation constraint and the incentive compatibility constraint. Therefore, the initially considered contract $\widetilde{w}(P_T)$ cannot be optimal. Note that the $\widetilde{w}'(P_T)$ is also not optimal as it pays a lottery in the gain space where the agent's preferences are concave. So the contract can further be improved by replacing these lotteries pointwise with sure payoffs. Note that this does not interfere with the argument in the proof, as this is a pointwise change in the contract, whereas the proof is concerned with a shift of payouts between states of the world.

Participation Constraint: We need to show that the following difference is positive:

$$\begin{aligned} &\int \left[g'(P_T)V(\overline{w}'(P_T)) + (1 - g'(P_T))V(\underline{w}'(P_T))\right] f(P_T|\overline{e}) dP_T \\ &- \int \left[g(P_T)V(\overline{w}(P_T)) + (1 - g(P_T))V(\underline{w}(P_T))\right] f(P_T|\overline{e}) dP_T \end{aligned} \tag{30}$$

Substituting in the definitions (27) to (29) and rearranging gives:

$$\begin{aligned} &\int_0^{\widetilde{P}} g(P_T) \left[V(\underline{w}) - V(\overline{w}(P_T))\right] f(P_T|\overline{e}) dP_T \\ &+ \int_{\widetilde{P}}^{\infty} (1 - g(P_T)) \left[V\left(\underline{w}'(P_T)\right) - V(\underline{w})\right] f(P_T|\overline{e}) dP_T \end{aligned} \tag{31}$$

With the definition of the agent's preferences (2) and further rearranging we obtain:

$$\begin{aligned} &\int_{\widetilde{P}}^{\infty} (1 - g(P_T)) \left[\left(\underline{w}'(P_T) - w^R\right)^{\alpha} + \lambda\left(w^R - \underline{w}\right)^{\beta}\right] f(P_T|\overline{e}) dP_T \\ &- \int_0^{\widetilde{P}} g(P_T) \left[\left(\overline{w}(P_T) - w^R\right)^{\alpha} + \lambda\left(w^R - \underline{w}\right)^{\beta}\right] f(P_T|\overline{e}) dP_T \end{aligned} \tag{32}$$

Note that $\underline{w}'(P_T)$ is constant and does not depend on P_T. With the definitions of $\widetilde{P}$ and s in equation (26) we get the following simplification:

$$s\left(\underline{w}'(P_T)-w^R\right)^{\alpha}-\int_0^{\widetilde{P}}g(P_T)\left(\overline{w}(P_T)-w^R\right)^{\alpha}f(P_T|\overline{e})dP_T \tag{33}$$

Substitution in the definition of $\underline{w}'(P_T)$ from equation (29) and recognizing that $\frac{1}{s}g(P_T)f(P_T|\overline{e})$ is a density function on $[0,\widetilde{P}]$ gives

$$\begin{aligned}&s\left(\frac{1}{s}\int_0^{\widetilde{P}}g(P_T)\left(\overline{w}(P_T)-w^R\right)f(P_T|\overline{e})dP_T\right)^{\alpha}\\&-\int_0^{\widetilde{P}}g(P_T)\left(\overline{w}(P_T)-w^R\right)^{\alpha}f(P_T|\overline{e})dP_T\end{aligned} \tag{34}$$

If we divide this expression by s and move the factor $1/s$ into the integrands, the integrands become expectations because $\frac{1}{s}g(P_T)f(P_T|\overline{e})$ is a density function on $[0,\widetilde{P}]$. From Jensen's inequality and the strict concavity of the agent's preferences in the gain space, it follows that (34) and therefore (30) is strictly positive.

Incentive Compatibility Constraint: When the contract $\widetilde{w}(P_T)$ is replaced by our candidate contract $\widetilde{w}'(P_T)$, the agent gains for some realized stock prices above $\widetilde{P}$ and loses for some realized stock prices below $\widetilde{P}$. In expectation, the utility gains are higher than the utility losses, which is just a restatement of our result that expression (30) is strictly positive. We assume that the likelihood ratio $\Delta f(P_T|e)/f(P_T|\overline{e})$ is monotonous. So if we multiply the integrands in (30) with the likelihood ratio, gains are multiplied by bigger numbers than losses. Consequently, the new expression is also strictly positive:

$$\begin{aligned}&\int\left[g'(P_T)V(\overline{w}'(P_T))+(1-g'(P_T))V(\underline{w}'(P_T))\right]\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}f(P_T|\overline{e})dP_T\\&-\int\left[g(P_T)V(\overline{w}(P_T))+(1-g(P_T))V(\underline{w}(P_T))\right]\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}f(P_T|\overline{e})dP_T>0\end{aligned} \tag{35}$$

Hence, switching from the initial contract $\widetilde{w}(P_T)$ to the alternative contract $\widetilde{w}'(P_T)$ also relaxes the incentive compatibility constraint. □

Lemma 2 allows us to rewrite the principal's program (4) to (6) as follows:

$$\min_{\widehat{P},w(P_T)\geq w^R} \int_{\widehat{P}}^{\infty} w\left(P_T\right) f(P_T|\overline{e}) dP_T + \underline{w} F(\widehat{P}|\overline{e}) \tag{36}$$

$$s.t. \int_{\widehat{P}}^{\infty} V\left(w\left(P_T\right)\right) f(P_T|\overline{e}) dP_T + V(\underline{w}) F(\widehat{P}|\overline{e}) \geq \underline{V} + C\left(\bar{e}\right) \quad , \tag{37}$$

$$\int_{\widehat{P}}^{\infty} V\left(w\left(P_T\right)\right) \Delta f(P_T|e) dP_T V(\underline{w}) \left[F(\widehat{P}|\overline{e}) - F(\widehat{P}|\underline{e})\right] \geq \Delta C \ . \tag{38}$$

The contract space that is defined by the constraints is not quasi convex, because the lower bound of the integral is a parameter of the problem and because $w(P_T)$ is not defined for $P_T < \widehat{P}$. Therefore, the Lagrangian approach only yields necessary conditions for an optimum. We cannot show sufficiency.

The derivative of the Lagrangian function with respect to $w\left(P_T\right)$ is:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial w(P_T)} &= f(P_T|\overline{e}) - \mu_{PC} \cdot \alpha \left(w\left(P_T\right) - w^R\right)^{\alpha-1} f(P_T|\overline{e}) - \mu_{IC} \cdot \alpha \left(w\left(P_T\right) - w^R\right)^{\alpha-1} \Delta f(P_T|e) \\ &= \alpha \left(w\left(P_T\right) - w^R\right)^{\alpha-1} f(P_T|\overline{e}) \left[\frac{1}{\alpha}\left(w\left(P_T\right) - w^R\right)^{1-\alpha} - \mu_{PC} - \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}\right] \ . \end{aligned} \tag{39}$$

Setting this equal to zero and solving for $w\left(P_T\right)$ yields the expression for $P_T > \widehat{P}$ in (7):

$$w\left(P_T\right) = w^R + \left[\alpha\left(\mu_{PC} + \mu_{IC} \frac{\Delta f\left(P_T\left|e\right.\right)}{f\left(P_T\left|\overline{e}\right.\right)}\right)\right]^{\frac{1}{1-\alpha}} \tag{40}$$

The derivative of the Lagrangian with respect to $\widehat{P}$ is:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \widehat{P}} &= \left(\underline{w} - w(\widehat{P})\right) f(\widehat{P}|\overline{e}) + \mu_{PC}\left(w(\widehat{P}) - w^R)^{\alpha} + \lambda\left(w^R - \underline{w}\right)^{\beta}\right) f(\widehat{P}|\overline{e}) \\ &\quad + \mu_{IC}\left(w(\widehat{P}) - w^R)^{\alpha} + \lambda\left(w^R - \underline{w}\right)^{\beta}\right) \Delta f(\widehat{P}|\overline{e}) && (41)\\ &= -\left(w(\widehat{P}) - w^R)^{\alpha} + \lambda\left(w^R - \underline{w}\right)^{\beta}\right) f(\widehat{P}|\overline{e}) && (42)\\ &\quad \left[\frac{\left(w(\widehat{P}) - \underline{w}\right)}{\left(w(\widehat{P}) - w^R)^{\alpha} + \lambda\left(w^R - \underline{w}\right)^{\beta}\right)} - \mu_{PC} - \mu_{IC}\frac{\Delta f(\widehat{P}|e)}{f(\widehat{P}|\overline{e})}\right] \ . && (43)\end{aligned}$$

This derivative is zero if the term in the brackets is zero. Substituting in equation (40) for $P_T = \widehat{P}$ yields:

$$\frac{\left(w(\widehat{P}) - \underline{w}\right)}{\left(w(\widehat{P}) - w^R)^{\alpha} + \lambda\left(w^R - \underline{w}\right)^{\beta}\right)} - \frac{1}{\alpha}\left(w(\widehat{P}) - w^R\right)^{1-\alpha} = 0 \tag{44}$$

$$\Leftrightarrow \alpha\left(w(\widehat{P})-w^{R}\right)^{a-1}\left(w(\widehat{P})-\underline{w}\right)-\lambda\left(w^{R}-\underline{w}\right)^{\beta}-\left(w(\widehat{P})-w^{R}\right)^{a}=0 \tag{45}$$

The left hand side of equation (45) is strictly decreasing in $w(\widehat{P})$. This can be shown by taking the first derivative of the LHS of (45) with respect to $w(\widehat{P})$. As $w(P)$ is strictly increasing in P from (40), the left hand side of equation (45) is strictly decreasing in $\widehat{P}$. Therefore, there can never be more than one solution to equation (45). ■

Proof of Proposition 2:

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

$$\min_{w(P_T)\geq\underline{w}} \int w\left(P_T\right) f(P_T|\hat{e})dP_T \tag{46}$$

$$s.t. \quad -\int\left(1-I(P_T)\right)U_l\left(w^R-w\left(P_T\right)\right)f(P_T|\hat{e})dP_T \tag{47}$$

$$+\int I(P_T)U_g\left(w\left(P_T\right)-w^R\right)f(P_T|\hat{e})dP_T \geq \underline{V}+C\left(\hat{e}\right) \quad ,$$

$$-\int\left(1-I(P_T)\right)U_l\left(w^R-w\left(P_T\right)\right)f_e(P_T|\hat{e})dP_T \tag{48}$$

$$+\int I(P_T)U_g\left(w\left(P_T\right)-w^R\right)f_e(P_T|\hat{e})dP_T \geq C' \ ,$$

where $I(P_T)$ is one if the contract pays off in the gain space and zero otherwise, C' denotes the first derivative of C and f_e denotes the first derivative of f with respect to e. Since optimization of program (46) to (48) is pointwise, the only changes with respect to program (4) to (6) 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 **??** and in Proposition 1 goes through as before. ■

B The optimal contract when P_T is lognormal and effort is continuous

From our parametric form of P_T in equation (9), we have that $\ln\left(P_T\right)$ is distributed normal with mean $\mu\left(e\right)=\ln\left(P_0\left(e\right)\right)+\left(r_f-\frac{\sigma^2}{2}\right)T$ and standard deviation $\sigma\sqrt{T}$. The density $f\left(P_T\left|e\right.\right)$ of the lognormal distribution is then:

$$f\left(P_T\left|e\right.\right)=\frac{1}{P_T\sqrt{2\pi T}\sigma}\exp\left\{-\frac{\left[\ln P_T-\mu\left(e\right)\right]^2}{2\sigma^2T}\right\}\ , \tag{49}$$

and the likelihood ratio is

$$\frac{\partial f\left(P_T\left|e\right.\right)/\partial e}{f\left(P_T\left|e\right.\right)}=\frac{P_0'\left(e\right)}{P_0\left(e\right)}\frac{\ln P_T-\mu\left(e\right)}{\sigma^2T}\ . \tag{50}$$

Using the continuous effort analogue of the optimal contract as given in equation (7), and defining

$$\gamma_1=\alpha\mu_{IC}\frac{P_0'\left(e\right)}{P_0\left(e\right)\sigma^2T}\ , \tag{51}$$

$$\gamma_0=\alpha\left(\mu_{PC}-\mu_{IC}\frac{P_0'\left(e\right)}{P_0\left(e\right)}\frac{\mu\left(e\right)}{\sigma^2T}\right)=\alpha\mu_{PC}-\gamma_1\mu\left(e\right)\ , \tag{52}$$

allows us to write:

$$\alpha\left(\mu_{PC}+\mu_{IC}\frac{P_0'\left(e\right)}{P_0\left(e\right)}\frac{\ln P_T-\mu\left(e\right)}{\sigma^2T}\right)=\gamma_0+\gamma_1\ln P_T\ . \tag{53}$$

From this, equation (10) follows immediately.

The optimal cut-off point was derived in the proof of Proposition 1 and is implicitly defined, according to equation (**??**), by

$$U_g'\left(w^*\left(\widehat{P}\right)-w^R\right)\left(w^*\left(\widehat{P}\right)-\underline{w}\right)-\left[\lambda U_l\left(w^R-\underline{w}\right)+U_g\left(w^*\left(\widehat{P}\right)-w^R\right)\right]=0\ . \tag{54}$$

Substituting U_g and U_l by their definitions in equations (**??**) and (**??**) yields equation (11).

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

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

[5] Barberis, Nicholas, and Ming Huang, 2005, Stocks and Lotteries: The Implications of Probability Weighting for Security Prices, Working Paper

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

[7] 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

[8] ———, 1999, Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments , Management Science 45, no. 3 (May), pp. 364-381

[9] Berkelaar, Arjan; Roy Kouwenberg, and Thierry Post, 2004, Optimal Portfolio Choice Under Loss Aversion, The Review of Economics and Statistics 86, no. 4 (November), pp. 973-987

[10] Bettis, J. Carr; John M. Bizjak, and Michael L. Lemmon, 2005, The Cost of Employee Stock Options, Journal of Financial Economics 76, pp. 455-470

[11] 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

[12] 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

[13] Carpenter, Jennifer N., 1998, The Exercise and Valuation of Executive Stock Options, Journal of Financial Economics 48, no. 2 , pp. 127-158

[14] 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

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

[16] de Meza, David, and David C. Webb, 2006, Incentive Design Under Loss Aversion, Journal of the European Economic Association, forthcoming

[17] Dittmann, Ingolf, and Ernst. Maug, 2006, Lower Salaries and No Options? On the Optimal Structure of Executive Pay, Journal of Finance, no. 62, pp. 303-343.

[18] 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.

[19] 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

[20] 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

[21] 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

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

[23] 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

[24] 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

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

[26] ———, 2002, Stock Options for Undiversified Executives, Journal of Accounting and Economics 33, no. 2 (April), pp. 3-42

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

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

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

[30] Huddart, Steven J.; Ravi Jagannathan, and P. Jane Saly, 1999, Valuing the Reload Features of Executive Stock Options, NBER Working Paper, no. 7020 (March)

[31] Inderst, Roman, and Holger M. Müller, 2003, A Theory of Broad-Based Option Pay, Mimeo, London School of Economics, (November)

[32] Ingersoll, Jonathan E. Jr., 2002, The Subjective and Objective Evaluation of Incentive Stock Options, Yale ICF Working Paper, no. 02-07 (February)

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

[34] Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, 1986, Fairness as a Constraint on Profit Seeking: Entitlements in the Market, The American Economic Review, 76(4), 728—741.

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

[36] ———, 1984, Choices, Values, and Frames, American Psychologist, 39(4), pp. 341-350.

[37] 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

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

[39] 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

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

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

[42] 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

[43] Langer, T., and Martin Weber, 2001, Prospect Theory, Mental Accounting, and Differences in Aggregated and Segregated Evaluation of Lottery Portfolios, Management Science, 47(5), 716—733.

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

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

[46] Loewenstein, G., and D. Adler, 1995, A Bias in the Prediction of Tastes, The Economic Journal, 105(431), 929—937.

[47] 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

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

[49] McNeil, B., S. G. Pauker, H. J. Sox, and Amos Tversky, 1982, On the elicitation of preferences for alternative therapies, New England Journal of Medicine, 306(21), 1259—62.

[50] Meulbroek, Lisa K., 2001, The Efficiency of Equity-Linked Compensation: Understanding the Full Cost of Awarding Executive Stock Options, Financial Management 30, no. 2 (Summer), pp. 5-30

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

[52] Oyer, Paul, 2004, Why Do Firms Use Incentives That Have No Incentive Effects?, Journal of Finance 59, no. 4 (August), pp. 1619-1650

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

[54] Post, Thierry, Marijn van den Assem, Guido Baltussen, and Richard Thaler, 2007, Deal or No Deal? Decision making under risk in a large-payoff game show, Working paper, Erasmus University Rotterdam

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

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

[57] Samuelson, W., and Richard Zeckhauser, 1988, Status quo bias in decision making, Journal of Risk and Uncertainty, 1, 7—59.

[58] Segal, U., and A. Spivak, 1990, First Order versus Second Order Risk Aversion, Journal of Economic Theory 51, 111–125.

[59] Shumway, T., 1997, Explaining returns with loss aversion. Working paper, University of Michigan.

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

[61] ———, 1999, Mental Accounting Matters, Journal of Behavioral Decision Making 12, pp. 183-206

[62] Tversky, Amos, and Daniel Kahneman, 1986, Rational Choice and the Framing of Decisions, The Journal of Business, 59(4), 251—278.

[63] ———, 1991, Loss Aversion in Riskless Choice: A Reference-Dependent Model, The Quarterly Journal of Economics, 106, no. 4 (November), pp. 1039-1061

[64] ———, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty, 5, pp. 297-323.

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

Table 1: Description of the dataset

This table displays mean, standard deviation, and the 10%, 50% and 90% quantiles of ten variables for our sample of 916 CEOs.

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Stock	n_S	2.08%	5.41%	0.03%	0.33%	5.60%
Options	n_O	1.41%	1.56%	0.14%	0.99%	3.20%
Fixed Salary ('000)	ϕ	2,332	2,896	576	1,560	4,313
Non-firm Wealth	W_0	55,954	512,599	1,715	9,629	63,016
Firm Value	P_0	9,540,284	29,294,103	395,336	2,127,836	17,761,268
Strike Price	K	7,280,536	25,166,019	242,728	1,369,911	12,486,310
Moneyness	K/P_0	69.33%	21.10%	39.60%	70.03%	99.21%
Maturity	T	4.65	1.34	3.44	4.50	6.28
Stock Volatility	σ^2	43.98%	22.73%	22.80%	36.85%	77.90%
Dividend Rate	d	1.21%	2.37%	0.00%	0.61%	3.28%

Table 2: Optimal piecewise linear contracts

This table describes the optimal piecewise linear contract for the unrestricted models where options and salary can become negative ($n_O \geq -n_S\exp(r_fT)$, $\phi \geq -W_0$). It shows mean and median of the three contract parameters base salary ϕ^*, stock holdings n_S^* and option holdings n_O^* together with the mean of the errors $error(\phi)=(\phi^* - \phi^d)/\sigma_\phi$, $error(n_S)=(n^*_S - n^d_S)/\sigma_S$, and $error(n_O)=(n^*_N - n^d_N)/\sigma_N$. The table also shows mean and median of the distance metric D_{Lin} and the average probability of a loss, i.e., $\text{Prob}(w^*(P_T) < w^R)$. Panel A displays the results for the Loss Aversion Model for seven different reference wages parameterized by θ. Panel B shows the results for the Expected Utility Model for six levels of the risk aversion parameter γ. The last row in Panel A shows the corresponding values of the observed contract.

Panel A: Loss Aversion Model

Theta	Obs.	Avg. Prob. of Loss	Salary (ϕ)			Stock (n_S)			Options (n_O)			D_Lin	
			Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median
0.0	914	4.98%	-6,056	228	-2.8977	0.0289	0.0059	0.1503	0.0047	0.0056	-0.6045	3.3520	0.7762
0.1	892	13.64%	2,156	1,288	-0.0630	0.0239	0.0056	0.0743	0.0101	0.0081	-0.2579	2.8843	0.7067
0.2	892	18.53%	2,991	1,198	0.2272	0.0280	0.0072	0.1343	0.0069	0.0054	-0.4457	6.2958	1.5477
0.4	898	28.21%	6,515	-2,782	1.4541	0.0409	0.0128	0.3880	-0.0084	0.0004	-1.4445	15.3007	3.9673
0.6	901	37.24%	-6,821	-7,753	-3.1500	0.0573	0.0183	0.6801	-0.0288	-0.0041	-2.7490	20.5270	5.2819
0.8	904	45.97%	-27,260	-10,105	-10.2202	0.0667	0.0206	0.8462	-0.0417	-0.0073	-3.5874	21.6123	5.8107
1.0	902	53.34%	-59,922	-10,769	-21.4904	0.0707	0.0214	0.9184	-0.0482	-0.0088	-4.0031	23.5704	5.9233
Data	916	N/A	2,332	1,560	N/A	0.0208	0.0033	N/A	0.0141	0.0099	N/A	N/A	N/A

Panel B: Expected Utililty Model

Gamma	Obs.	Salary (ϕ)			Stock (n_S)			Options (n_O)			D_Lin	
		Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median
0.1	916	-44,923	-9,198	-16.3183	0.0619	0.0198	0.7595	-0.0349	-0.0053	-3.1463	17.7886	4.9812
0.2	916	-49,054	-9,481	-17.7447	0.0652	0.0204	0.8196	-0.0401	-0.0063	-3.4775	19.4093	5.0799
0.5	916	-54,101	-9,238	-19.4877	0.0673	0.0212	0.8584	-0.0473	-0.0088	-3.9432	21.4226	5.3102
1.0	915	-51,207	-8,522	-18.4793	0.0642	0.0218	0.8016	-0.0505	-0.0126	-4.1454	20.5301	5.1504
3.0	914	-35,482	-4,801	-13.0479	0.0510	0.0151	0.5562	-0.0493	-0.0143	-4.0675	15.2574	4.0991
6.0	903	-14,391	-767	-5.7448	0.0373	0.0096	0.2988	-0.0389	-0.0099	-3.3985	7.9523	2.5680

Table 3: Comparison of Loss Aversion-model with matched Expected Utility-model

This table compares the optimal Loss Aversion contract with the equivalent optimal Expected Utility contract where each CEO's risk aversion parameter γ is chosen such that both models predict the same certainty equivalent for the observed contract. Contracts are piecewise linear, and options and salary can become negative ($n_O \geq -n_S \exp(r_f T)$, $\phi \geq -W_0$). Panel A shows the average equivalent γ, mean and median of the difference between the distance metric D_{Lin} between the EU-model and the LA-model, and the frequency that this difference is positive. The last three columns show the p-values of three tests about the difference between the two models: the Wilcoxon signed rank test and the sign test both for zero median differences, and the t-test for zero average differences. Panel B shows the frequencies that optimal option holdings are positive, that the optimal salary is positive, and that both (options and salary) are positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

θ	Obs.	Average equivalent γ	$D_{Lin}^{EU} - D_{Lin}^{LA}$			P-values $D_{Lin}^{EU} - D_{Lin}^{LA} = 0$		
			Percent > 0	Mean	Median	T-test	Wilcoxon	Sign-Test
0.0	914	0.1781	98.91%	15.4800	4.0723	0.0000	0.0000	0.0000
0.1	892	0.2362	98.65%	16.2388	4.0915	0.0001	0.0000	0.0000
0.2	892	0.3272	91.70%	13.2197	2.8816	0.0020	0.0000	0.0000
0.3	896	0.4321	79.35%	7.8511	1.7597	0.0019	0.0000	0.0000
0.4	898	0.5565	63.70%	4.8478	0.6597	0.0040	0.0000	0.0000
0.5	899	0.6832	49.72%	4.0860	-0.0055	0.0441	0.0030	0.8939
0.6	901	0.7946	41.51%	0.8017	-0.1930	0.6617	0.0016	0.0000
0.7	900	0.8814	35.44%	-3.0737	-0.2832	0.3242	0.0000	0.0000
0.8	900	0.9291	30.78%	-0.2822	-0.3628	0.8055	0.0000	0.0000
0.9	896	0.9189	27.68%	-2.2416	-0.4482	0.0005	0.0000	0.0000
1.0	898	0.8581	23.16%	-2.0834	-0.5007	0.1488	0.0000	0.0000

Panel B: Positive option holdings and poisitve salaries

θ	Percent with positive option holdings		Percent with positive fixed salary		Percent with positive options and salary	
	EU	LA	EU	LA	EU	LA
0.0	28.12%	80.53%	1.86%	57.77%	0.33%	50.77%
0.1	27.13%	89.13%	1.68%	75.56%	0.11%	71.64%
0.2	26.23%	78.03%	1.91%	62.89%	0.34%	60.20%
0.3	24.67%	63.62%	1.56%	49.11%	0.00%	46.09%
0.4	23.39%	52.23%	2.00%	35.41%	0.56%	33.30%
0.5	21.25%	43.83%	1.89%	24.69%	0.22%	22.58%
0.6	18.42%	35.74%	2.00%	16.76%	0.11%	14.65%
0.7	16.44%	30.56%	2.11%	11.89%	0.00%	9.78%
0.8	15.67%	26.67%	2.44%	8.78%	0.11%	6.00%
0.9	16.18%	25.78%	2.34%	6.58%	0.00%	4.69%
1.0	17.48%	24.05%	2.45%	4.90%	0.11%	3.23%

Table 4: Comparison of restricted models with positive salaries and positive option holdings

This table compares the optimal Loss Aversion contract with the equivalent optimal Expected Utility contract where each CEO's risk aversion parameter γ is chosen such that both models predict the same certainty equivalent for the observed contract. Contracts are piecewise linear and subject to the constraint that option holdings and salaries must be non-negative ($n_O \geq 0$, $\phi \geq 0$). Panel A shows the average equivalent γ, mean and median of the difference between the distance metric D_{Lin} between the EU-model and the LA-model, and the frequency that this difference is positive. The last three columns show the p-values of three tests about the difference between the two models: the Wilcoxon signed rank test and the sign test both for zero median differences, and the t-test for zero average differences. Panel B shows the frequencies that optimal option holdings are positive, that the optimal salary is positive, and that both (options and salary) are positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

θ	Obs.	Average equivalent γ	$D_{Lin}^{EU} - D_{Lin}^{LA}$			P-values $D_{Lin}^{EU} - D_{Lin}^{LA} = 0$		
			Percent > 0	Mean	Median	T-test	Wilcoxon	Sign-Test
0.0	912	0.1781	57.13%	0.1589	0.0060	0.0000	0.0000	0.0000
0.1	891	0.2359	54.32%	-0.7680	0.0068	0.0125	0.0000	0.0109
0.2	887	0.3281	44.98%	-2.7407	-0.0001	0.0019	0.0020	0.0031
0.3	895	0.4342	50.84%	-6.6810	0.0000	0.0088	0.0006	0.6398
0.4	903	0.5566	61.35%	-7.5460	0.0004	0.0293	0.0406	0.0000
0.5	902	0.6831	70.40%	-6.9902	0.0010	0.1367	0.0000	0.0000
0.6	898	0.7923	75.50%	-8.1499	0.0021	0.2244	0.0000	0.0000
0.7	901	0.8786	78.47%	-9.6937	0.0031	0.2703	0.0000	0.0000
0.8	902	0.9265	79.05%	-4.7573	0.0032	0.2815	0.0000	0.0000
0.9	903	0.9171	79.62%	-0.0929	0.0031	0.1212	0.0000	0.0000
1.0	900	0.8531	79.00%	-0.0895	0.0026	0.2029	0.0000	0.0000

Panel B: Positive option holdings and poisitve salaries

θ	Percent with positive option holdings		Percent with positive fixed salary		Percent with positive options and salary	
	EU	LA	EU	LA	EU	LA
0.0	82.13%	87.28%	17.65%	69.52%	0.11%	57.02%
0.1	81.37%	92.70%	18.29%	82.27%	0.00%	75.08%
0.2	79.71%	93.91%	19.73%	72.38%	0.23%	66.29%
0.3	79.22%	91.84%	19.66%	61.01%	0.00%	52.85%
0.4	78.29%	89.70%	20.71%	51.50%	0.44%	41.20%
0.5	77.61%	88.91%	21.29%	42.46%	0.00%	31.49%
0.6	77.28%	86.64%	21.94%	37.08%	0.11%	23.83%
0.7	77.36%	85.90%	21.98%	31.74%	0.11%	17.76%
0.8	77.27%	84.15%	22.17%	32.71%	0.11%	17.07%
0.9	77.41%	83.72%	21.93%	31.01%	0.00%	14.84%
1.0	77.78%	83.11%	21.89%	31.11%	0.11%	14.33%

Table 5: Relation between accuracy and contract parameters

This table shows the results of an OLS regression of the distance metric D_{Lin}^{LA} on an intercept, the observed stockholdings n_S^d and the observed option holdings n_O^d. Contracts are piecewise linear, and options and salary can become negative ($n_O \geq -n_S\exp(r_f T)$, $\phi \geq -W_0$). Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

θ	Obs.	Intercept	Stock (n_S^d)		Options (n_O^d)	
		Estimate	Estimate	P-Value	Estimate	P-Value
0.0	915	1.624	145.104	0.000	-91.598	0.141
0.1	915	1.729	64.880	0.000	-9.651	0.692
0.2	911	3.666	104.513	0.000	33.195	0.606
0.3	909	6.955	369.108	0.000	-176.969	0.245
0.4	906	7.747	561.679	0.000	-256.259	0.237
0.5	904	7.322	690.191	0.000	-320.740	0.283
0.6	902	7.937	984.962	0.000	-538.151	0.187
0.7	902	9.803	1212.630	0.000	-763.371	0.159
0.8	896	11.108	831.841	0.000	-494.806	0.109
0.9	902	12.024	432.531	0.000	-196.682	0.231
1.0	890	13.285	890.354	0.000	-602.084	0.056

Table 6: Linear Loss Aversion contracts for subsamples formed on stock ownership

This table describes the optimal piecewise linear Loss Aversion contract for two subsamples. The left part of the table displays results for CEOs who own less than 5% of their firm's equity, while the right part displays the corresponding results for the remaining CEOs in our sample. Contracts are piecewise linear, and options and salary can become negative ($n_O \geq -n_S\exp(r_f T)$, $\phi \geq -W_0$). For both subsamples, the table shows the mean of the three contract parameters base salary ϕ^*, stock holdings n_S^* and option holdings n_O^* together with the mean of the distance metric D_{Lin}. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

θ	Non-owner executives					Owner executives				
	Obs.	ϕ	n_S	n_O	D_{Lin}^{LA}	Obs.	ϕ	n_S	n_O	D_{Lin}^{LA}
0.0	817	-1,378	0.0117	0.0075	1.6466	97	-45,456	0.1736	-0.0192	17.7164
0.1	801	1,065	0.0115	0.0079	2.1316	91	11,758	0.1337	0.0293	9.5101
0.2	797	-1,708	0.0164	0.0022	4.6481	95	42,407	0.1260	0.0459	20.1193
0.3	801	-5,315	0.0218	-0.0041	7.5513	96	114,429	0.1386	0.0397	50.9975
0.4	808	-9,634	0.0267	-0.0100	8.6561	90	151,495	0.1682	0.0063	74.9539
0.5	805	-15,089	0.0304	-0.0147	8.8282	94	115,894	0.2237	-0.0648	85.3169
0.6	807	-19,731	0.0329	-0.0178	9.3494	94	104,016	0.2666	-0.1227	116.4881
0.7	807	-23,164	0.0345	-0.0199	10.3571	96	104,480	0.3014	-0.1693	143.7020
0.8	809	-25,429	0.0352	-0.0210	11.0672	95	-42,856	0.3341	-0.2176	111.4122
0.9	807	-27,304	0.0360	-0.0221	11.6666	92	-192,107	0.3506	-0.2488	74.4822
1.0	807	-28,782	0.0358	-0.0220	12.1631	95	-324,447	0.3671	-0.2704	120.4726
Data	819	2,356	0.0062	0.0137	N/A	97	2,126	0.1441	0.0169	N/A

Table 7: Optimal nonlinear Loss Aversion contracts

This table describes the optimal Loss Aversion contract for the general non-linear contract. The table shows the average slope of the wage function below the observed strike price Δ_{Low}, the average slope of the wage function above the observed strike price Δ_{High}, and the frequency with which $\Delta_{High} > \Delta_{Low}$. In addition, the table shows (1) the average dismissal probability which is the probability with which the contract pays the minimum wage $\underline{w}$, (2) the incentives from dismissals that are generated by the drop to the minimum wage $\underline{w}$, and (3) the mean inflection quantile, which is the quantile at which the curvature of the optimal wage function changes from convex to concave. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

θ	Obs.	Mean Δ_{Low}	Mean Δ_{High}	Percent $\Delta_{High} > \Delta_{Low}$	Mean Dismissal Probability	Incentives from Dismissals	Mean Inflection Quantile
0.0	871	2.07%	2.66%	90.01%	0.85%	1.40%	85.15%
0.1	874	1.74%	2.81%	96.34%	1.87%	3.51%	92.95%
0.2	876	1.28%	2.69%	98.06%	3.45%	7.24%	96.84%
0.3	883	0.94%	2.58%	98.19%	5.42%	12.06%	97.87%
0.4	882	0.81%	2.44%	98.75%	7.64%	17.54%	98.42%
0.5	882	0.66%	2.35%	99.09%	10.07%	23.94%	98.74%
0.6	891	0.50%	2.27%	99.21%	12.63%	31.19%	98.95%
0.7	887	0.39%	2.12%	99.32%	15.24%	38.29%	99.10%
0.8	878	0.35%	1.94%	99.43%	17.76%	45.00%	99.10%
0.9	875	0.31%	1.85%	99.54%	20.17%	51.47%	99.15%
1.0	857	0.24%	1.52%	99.65%	22.54%	57.18%	99.35%

Table 8: Comparison of linear and nonlinear Loss Aversion models

This table compares the optimal piecewise linear Loss Aversion contract with the optimal nonlinear Loss Aversion contract. For piecewise linear contracts, options and salary can become negative ($n_O \geq -n_S\exp(r_fT)$, $\phi \geq -W_0$), while the minimum wage equals minus the CEO's wealth ($\underline{w} = -W_0$) for nonlinear contracts. For both models, the table shows the average slope of the wage function below the observed strike price, n_S and Δ_{Low}, respectively, the average slope of the wage function above the observed strike price, $n_S + n_O$ and Δ_{High}, respectively, and the average distance metric D_{NonLin}. In addition, the table shows the savings $[E(w^d(P_T)) - E(w^*(P_T))] / E(w^d(P_T))$ the models predict from switching from the observed contract to the optimal contract. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

θ	Obs.	Linear Option Contract				General Nonlinear contract			
		Mean n_S	Mean $n_S + n_O$	Mean Savings	Mean D_{NonLin}	Mean Δ_{Low}	Mean Δ_{High}	Mean Savings	Mean D_{NonLin}
0.0	869	0.0199	0.0261	0.0018	0.2039	0.02065	0.02659	0.0050	0.2022
0.1	854	0.0198	0.0291	0.0049	0.2523	0.01719	0.02791	0.0154	0.1620
0.2	854	0.0247	0.0296	0.0114	0.4652	0.01264	0.02680	0.0319	0.2170
0.3	866	0.0322	0.0302	0.0189	0.5994	0.00948	0.02583	0.0487	0.2812
0.4	870	0.0385	0.0288	0.0257	0.7117	0.00806	0.02398	0.0634	0.3623
0.5	872	0.0480	0.0280	0.0334	0.7764	0.00653	0.02301	0.0783	0.4124
0.6	884	0.0550	0.0266	0.0406	0.8824	0.00499	0.02229	0.0929	0.4704
0.7	884	0.0604	0.0258	0.0468	0.9166	0.00393	0.02128	0.1062	0.5028
0.8	876	0.0652	0.0235	0.0523	0.9509	0.00351	0.01934	0.1185	0.5249
0.9	870	0.0664	0.0224	0.0572	0.9974	0.00299	0.01776	0.1306	0.5568
1.0	855	0.0683	0.0218	0.0615	0.9963	0.00233	0.01519	0.1421	0.5800

Table 9: Comparative statics for the parameters of the value function

This table describes the optimal piecewise linear Loss Aversion contract and the optimal nonlinear Loss Aversion contract for different values of the parameters α, β, and λ of the value function. The reference wage w^R is set equal to last year's fixed salary, i.e. $\theta = 0$. Panel A shows the results for the parameter λ, Panel B for α, and Panel C for β. For piecewise linear contracts, options and salary can become negative ($n_O \geq -n_S \exp(r_f T)$, $\phi \geq -W_0$), while the minimum wage equals minus the CEO's wealth ($\underline{w} = -W_0$) for nonlinear contracts. For both models, the table shows the average slope of the wage function below the observed strike price, n_S and Δ_{Low}, respectively, the average slope of the wage function above the observed strike price, $n_S + n_O$ and Δ_{High}, respectively, and the average distance metric D_{NonLin}. In addition, the table shows the savings $[E(w^d(P_T)) - E(w^*(P_T))] / E(w^d(P_T))$ the models predict from switching from the observed contract to the optimal contract. Some observations are lost because of numerical problems.

Panel A: Loss aversion parameter λ

		Linear Option Contract				General Nonlinear contract			
λ	Obs.	Mean n_S	Mean $n_S + n_O$	Mean Savings	Mean D_{NonLin}	Mean Δ_{Low}	Mean Δ_{High}	Mean Savings	Mean D_{NonLin}
1.00	809	0.0311	0.0195	0.0073	0.6834	0.0177	0.0213	0.0142	0.2463
1.50	907	0.0302	0.0318	0.0021	0.2094	0.0293	0.0336	0.0057	0.1744
2.00	886	0.0242	0.0301	0.0015	0.1485	0.0256	0.0310	0.0046	0.1527
2.25	869	0.0199	0.0261	0.0018	0.2039	0.0206	0.0266	0.0050	0.2022
2.50	859	0.0190	0.0256	0.0017	0.1994	0.0198	0.0260	0.0049	0.1961
3.00	863	0.0195	0.0264	0.0017	0.2032	0.0197	0.0267	0.0050	0.1930
4.00	870	0.0207	0.0278	0.0018	0.1752	0.0213	0.0282	0.0052	0.1680

Panel B: Gain space curvature α

		Linear Option Contract				General Nonlinear contract			
α	Obs.	Mean n_S	Mean $n_S + n_O$	Mean Savings	Mean D_{NonLin}	Mean Δ_{Low}	Mean Δ_{High}	Mean Savings	Mean D_{NonLin}
0.60	697	0.0532	0.0250	0.0327	0.5574	0.0477	0.0297	0.0450	0.4072
0.70	893	0.0464	0.0261	0.0273	0.5023	0.0414	0.0304	0.0397	0.3620
0.80	912	0.0361	0.0304	0.0082	0.2974	0.0336	0.0337	0.0155	0.2179
0.88	869	0.0199	0.0261	0.0018	0.2039	0.0206	0.0266	0.0050	0.2022
0.95	528	0.0177	0.0281	0.0005	0.0886	0.0190	0.0266	0.0025	0.1617

Panel C: Loss space curvature β

		Linear Option Contract				General Nonlinear contract			
β	Obs.	Mean n_S	Mean $n_S + n_O$	Mean Savings	Mean D_{NonLin}	Mean Δ_{Low}	Mean Δ_{High}	Mean Savings	Mean D_{NonLin}
0.60	721	0.0180	0.0261	0.0006	0.1093	0.0227	0.0264	0.0019	0.2152
0.70	886	0.0252	0.0309	0.0015	0.1461	0.0272	0.0320	0.0051	0.1742
0.80	894	0.0264	0.0318	0.0017	0.1663	0.0273	0.0328	0.0053	0.1619
0.88	869	0.0199	0.0261	0.0018	0.2039	0.0206	0.0266	0.0050	0.2022
0.95	866	0.0202	0.0251	0.0019	0.2691	0.0194	0.0260	0.0049	0.2348