

# Executive Stock Options when Managers Are Loss-Averse\*

Ingolf Dittmann<sup>†</sup>      Ernst Maug<sup>‡</sup>      Oliver Spalt<sup>§</sup>

November 30, 2006

## 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 and we compare stylized contracts consisting of fixed salaries, stock, and options generated by the model to observed contracts. 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. 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 for their feedback. We also thank the collaborative research center 649 on "Economic Risk" in Berlin for financial support.

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

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

<sup>§</sup>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 and we compare stylized contracts consisting of fixed salaries, stock, and options generated by the model to observed contracts. 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. 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 analyze executive compensation contracts using a simple contracting model where the manager is loss averse in order to explain salient features of observed compensation contracts. We parameterize this model using standard assumptions and then compare the contracts generated by the model with those actually observed 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.<sup>1</sup> 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, in particular for the widespread application of this model to the analysis of the design (strike price, indexing, reloading, and repricing) and the valuation of executive stock options.<sup>2</sup>

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

---

<sup>1</sup>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).

<sup>2</sup>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.

gains and incremental losses. For brevity, we will refer to all three features as loss aversion and to the corresponding 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.<sup>3</sup> However, we do not use the notion of decision weights, so our model does not apply all elements of prospect theory. Given our results, this additional element does not seem to be needed.

The main drawback of expected utility-approaches to explaining the prevalent use of stock options in compensation contracts is the fact that risk averse managers gain little utility from payoffs when the value of the firm is high. Whenever firm value is high, managers become wealthier and their marginal utility becomes small. This blunts any instrument for providing incentives that pays off only when firm value is high. Contracts that rely less on rewards for good outcomes ("carrots") and more on penalties for bad outcomes ("sticks") are more beneficial as they provide similar incentives at a lower cost. However, these predictions are at odds with observed compensation practice. By comparison, loss aversion implies that managers are more averse to losses than they are attracted by gains, so they demand a 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 offer a higher base salary together with options and thus mirror observed compensation practice much more accurately in this dimension. These are therefore more attractive compared to contracts that use restricted stock together with lower base salaries, as suggested by the Expected Utility-model, thus exposing a larger fraction of the CEO's wealth to risk.

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

---

<sup>3</sup>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 utility functions cannot account for risk-aversion over small stakes-gambles, a feature readily explained by loss aversion. There are also some papers who take a more critical stance. Myagkov and Plott (1997) show document that the risk-seeking implied by prospect theory diminishes with experience, a result also supported by List (2004). Plott and Zeiler (2005) call into question the general interpretation of gaps between the willingness to pay and the willingness to accept as evidence for loss aversion.

and (mostly) convex function of the stock price. Below this critical stock price compensation falls discontinuously to some lower bound. We suggest that the optimal contract is best interpreted as consisting of an arrangement where the manager is fired if the value of the firm falls below some threshold, and where she obtains a compensation contract that provides stock price-related rewards above her reference wage as long as her employment lasts.

Both, the Loss Aversion-model and the Expected Utility-model have different strengths and weaknesses. Their ability to explain observed compensation practice is therefore 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, 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.

We first observe that the performance of each model depends critically on the parameterization, and then compare both models on the basis of parameterizations that imply the same risk premia (i.e., the same certainty equivalents) for the observed contract. Overall, we find that the Loss Aversion-model predicts observed contracts better than the Expected Utility model. Generally, the Expected Utility model predicts only few options, whereas the Loss Aversion-model generates options as part of the optimal contract. About 95% of the CEOs in our sample have options, and for these the Loss Aversion-model is mostly superior. However, the performance of the Loss Aversion-model is poor for those CEOs who have no options but large shareholdings of their companies. We suggest that CEOs who are also large blockholders (with stakes above 5% in our analysis) should be regarded more as owners of their companies rather than as salaried agents. Applying the principal-agent approach to these CEOs is therefore problematic.

We then analyze the general nonlinear contracts without restricting the functional form to contracts that can be implemented with securities. The contracts predicted by the Loss Aversion-model are convex and could be implemented with stock and a portfolio of options, whereas the contracts predicted by the Expected Utility-model are mostly concave and could be implemented with securities only if we allow for short positions in options.

The nonlinear contracts invariably involve approximation errors with respect to the observed, piecewise linear contracts consisting of stock and options and we develop a methodology

that addresses this issue. The general nonlinear contract features dismissals of the CEO as an integral part of the contract and we analyze what the benefits to shareholders would be from replacing a simple contract consisting of a fixed salary, stock, and one option grant with this more complex contract that would arguably also require adjustments of the governance structure. We estimate that shareholders would save between 0.1% and 7.3% of current compensation from the more complex contract, where the higher figures only obtain for extreme parameterizations. We conclude that incentive provision through CEO dismissals rather than 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) demonstrate theoretically that hedge-fund managers take more risk if their incentive fees become more substantial, an effect that contrasts the implications of a model based on hyperbolic absolute risk aversion (HARA). Their empirical results tend to support the prospect model. Ljungqvist and Wilhelm (2005) base their measure of issuer satisfaction in initial public offerings on loss aversion. The only application that fails to support loss aversion to the best of our knowledge is Massa and Simonov (2005) in their study of individual investor behavior. Despite the usefulness of loss aversion to analyze risk taking incentives in many areas of finance, the only paper so far that rigorously applies loss aversion to principal-agent theory is de Meza and Webb (2005). However, they 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 the potential of loss aversion to explain observed compensation practice and demonstrates this empirically.

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 analyzes the general nonlinear contracts and provides some comparative static analysis. 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  $w(P_T)$  which 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,  $e \in \{\underline{e}, \bar{e}\}$  so that  $P_T$  is distributed with density  $f(P_T|e)$ . Later we will also allow for continuous effort  $e \in [0, \infty)$ . For notational convenience we write  $\Delta e = \bar{e} - \underline{e}$ ,  $\Delta C = C(\bar{e}) - C(\underline{e})$ , and  $\Delta f(P_T|e) = f(P_T|\bar{e}) - f(P_T|\underline{e})$ . We require the monotone likelihood ratio property (MLRP) to hold for  $f$ , so  $\Delta f(P_T|e)/f(P_T|\bar{e})$  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

$$V(w(P_T)) - C(e), \quad (1)$$

where  $C(e)$  is an increasing and convex cost function, and where we assume preferences over wage income,  $w(P_T)$ , of the form<sup>4</sup>

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

Here,  $w^R$  denotes the reference wage. If the payoff of the contract at time  $T$  exceeds the reference wage, 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

---

<sup>4</sup>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).

below  $w^R$  as the *loss space*. There are three ingredients which set this specification apart from standard von Neumann-Morgenstern concave utility specifications. First, while both gains and losses enter the utility function, 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.<sup>5</sup> Formally, this introduces a kink in the value function at  $w^R$  and thus locally infinite risk-aversion.<sup>6</sup> Second, the manager treats her income from the firm separately from other sources of income.<sup>7</sup> Third, while  $V(w(P_T))$  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 become standard in the literature on executive compensation. The preferences for this model are:

$$U(W_0 + w(P_T)) = \frac{(W_0 + w(P_T))^{1-\gamma}}{1-\gamma}, \quad (3)$$

where  $W_0$  denotes wealth and  $\gamma$  represents the coefficient of relative risk aversion. We will refer to this model as the Expected Utility-model, or, for brevity, as the EU-model. We alert the reader to the fact that our 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.

We assume that the reference point  $w^R$  is exogenous in two respects. Firstly, the reference point does not depend on any of the parameters of the contract. Alternative assumptions would relate the reference point to the median or the mean payoff of the contract  $w(P_T)$ , which would increase the mathematical complexity of the argument substantially.<sup>8</sup> Secondly, the reference point is also independent of the level of effort. This is defensible if the cost of effort is non-pecuniary and the manager separates the costs of effort from the pecuniary wage. However,

---

<sup>5</sup>Rabin (2000) calls loss aversion "the most firmly established feature of risk preferences". For evidence from psychology see for example 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). For evidence from the financial literature see the papers mentioned in the introduction.

<sup>6</sup>This is also called 'first-order risk aversion' (Segal and Spivak, 1990).

<sup>7</sup>The literature refers to this phenomenon as "framing" or "mental accounting." This concept was present already in the earlier papers by Kahneman and Tversky. See Thaler (1999) for a survey of the evidence on mental accounting. The conventional modeling framework in the compensation literature implicitly makes the same assumption, usually for tractability to avoid modeling other sources of uncertainty on the agent's overall wealth.

<sup>8</sup>De Meza and Webb (2005) focus on this aspect of applying loss aversion to principal-agent theory.



this is potentially a strong assumption if the costs are pecuniary and the manager frames the problem so that she feels a loss if her payoff does not exceed  $w^R$  plus any additional expenses for exerting effort. In the second case,  $C(e)$  should simply be added to the reference point  $w^R$ . We do not pursue this route here for mathematical tractability.

The manager has some outside employment opportunity that provides her with a utility level  $\underline{V}$ , so any feasible contract must satisfy the ex ante participation constraint  $E[V(w(P_T))] - C(e) \geq \underline{V}$ . We assume that there exists a lower bound  $\underline{w}$  on the wage function proposed by the shareholders such that  $\underline{w} \leq w(P_T)$  for all  $P_T$  and  $\underline{w} < w^R$ . Such a lower bound arises naturally with limited liability: e.g., the manager could be required to invest some of her wealth in the securities of the firm, which would put all her wealth at risk, 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 wealth.

### 3 Analysis

#### 3.1 Discrete effort

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

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

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

$$\int V(w(P_T)) \Delta f(P_T|e) dP_T \geq \Delta C \quad . \quad (6)$$

The convexity of the agent's preferences over losses poses a major complication to solving this problem (see Appendix A). Over the gain space, however, the optimal solution is characterized by setting up the Lagrangian for this problem and then maximizing it pointwise with respect to  $w(P_T)$ . Denote the Lagrange multiplier on the participation constraint (5) by  $\mu_{PC}$  and the Lagrange multiplier on the incentive compatibility constraint (6) by  $\mu_{IC}$ .

**Proposition 1. (Optimal contract):** *Given the preference structure in (1) and assuming*

MLRP the optimal contract  $w^*(P_T)$  for the principal agent problem in (4) to (6), is given by:

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

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

Proposition 1 provides us with a general characterization of the optimal contract with a loss-averse manager. The contract is simple. For some region  $P_T > \hat{P}$  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  $P_T \leq \hat{P}$  the optimal contract pays off the lowest possible wage  $\underline{w}$ . The contract also features a discontinuity at  $\hat{P}$  where the manager's wage jumps discretely from  $\underline{w}$  to some value  $w^*(P_T) \geq w^R > \underline{w}$ .<sup>9</sup> Interestingly, the optimal LA-contract (7) combines punishments ("sticks") with rewards ("carrots"), which sets this contract apart from EU-model.

Looking at equation (7) in more detail shows that for the gain space, where  $P_T > \hat{P}$ , we obtain a result which is 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  $P_T \leq \hat{P}$ . 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 probability  $p$  and  $w^R$  with probability  $(1-p)$  so that both the incentive constraint and the participation constraint, but expected compensation costs are reduced. We then show in the next step that these lotteries are always degenerate for the optimal contract and that the optimal contract pays off  $\underline{w}$  for all  $P_T \in (0, \tilde{P}]$  and  $w^R$  for all  $P_T \in (\tilde{P}, \infty)$ , where  $\tilde{P}$  is a uniquely defined cut-off value. In a final step, it can be proven that  $\tilde{P} > \hat{P}$ , i.e. that the jump in the loss space never occurs because the contract jumps directly from  $\underline{w}$  to the gain space instead.

---

<sup>9</sup>In the context of executive compensation we could interpret the payoff  $\underline{w}$  also as the consequence of firing the manager when she underperforms too much, so  $\hat{P}$  is the cutoff point below which she is fired. Contrary to observed contractual practice, one would have to assume then, however, that the pay package to which she is legally entitled can be taken away from her at a later stage.

### 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.<sup>10</sup> In our case, this requires that

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

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 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 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  $f'(P_T|e)/f(P_T|e)$ .*

---

<sup>10</sup>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).

## 4 Implementation and Data

### 4.1 Implementation

**The Loss Aversion-contract.** In our empirical implementation, we assume that the stock price follows a lognormal distribution and specify:<sup>11</sup>

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

where  $r_f$  is the risk-free rate of interest,  $\sigma^2$  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  $P_0 = E[P_T \exp\{-r_f T\}]$ .<sup>12</sup> 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:

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

where  $\gamma_0$  and  $\gamma_1$  depend on the two Lagrange multipliers and the production function  $P_0(e^*)$ .  $\hat{P}$  is uniquely defined by the condition:

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

Hence, we can represent the nonlinear LA-contract by the coefficients  $\gamma_0$  and  $\gamma_1$  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  $P_T \leq \exp\{\alpha/(1-\alpha) - \gamma_0/\gamma_1\}$ . Above this value  $w^*(P_T)$  is concave. It is therefore an empirical question whether the contract described in

<sup>11</sup>Our specification ignores dividends in order to simplify the exposition. We include dividends in the numerical analysis below.

<sup>12</sup>Here and in the following all expectations are taken with respect to the probability distribution of  $u \sim N(0, 1)$ . Instead of writing  $P_0 = E[P_T(u, e) \exp\{-r_f T\}]$  and  $w(P_T)$  as functions of  $u$  we submerge reference to  $u$  for ease of exposition.

equation 10 can describe observed contracts as the concave region may or may not be empirically relevant.

We can now identify the parameters that we need to determine in order to analyze the optimal contract numerically. First, we have to find appropriate values for the preference parameters  $\alpha$ ,  $\beta$ ,  $\lambda$ , and  $w^R$  and for the lower bound of the wage  $\underline{w}$ . For these we rely on the experimental literature and on data for executive compensation contracts. Then we need the parameters that describe the distribution of  $P_T$  in (9). These are the return variance  $\sigma^2$ , the maturity of the contract  $T$ , the risk-free rate  $r_f$ , and the value of the firm  $P_0$ . All these need to be determined from available data.

The parameterized model given by equations (10) and (11) contains only two parameters that we cannot determine:  $\gamma_0$  and  $\gamma_1$ . They depend on the production function  $P_0(e)$  and on the cost function  $C(e)$ . We can determine these numerically.

**The Expected Utility-contract.** We obtain the optimal nonlinear contract for the EU-model from (see Dittmann and Maug (2006)) and represent it in our notation as:

$$w^{EU}(P_T) = \begin{cases} (\delta_0 + \delta_1 \ln P_T)^{1/\gamma} - W_0 \exp(r_f T) & \text{if } P_T \geq \bar{P} \\ -\underline{w} & \text{if } P_T < \bar{P} \end{cases}, \quad (12)$$

$$\text{where : } \ln \bar{P} = ((\underline{w} + W_0 \exp(r_f T))^\gamma - \delta_0) / \delta_1 \quad (13)$$

Hence, we can represent the nonlinear EU-contract by the coefficients  $\delta_0$  and  $\delta_1$  and write it as  $\mathcal{C}^{EU} = \{\delta_0, \delta_1\}$ . The contract described by (12) is also convex of some region and then concave past the inflection point.

**Contracts with Stock and Options.** Observed contracts consist of salaries, bonus payments, and holdings of corporate securities in addition to many other provisions and perquisites. We represent these contracts as consisting of a fixed salary  $\phi$  that is paid at time 0,  $n_S$  shares and  $n_O$  options, where the total number of shares the company has outstanding is normalized to one. We will refer to these contracts as piecewise linear contracts and denote them by  $\mathcal{C}_{Lin}^{LA} = \{\phi^{LA}, n_S^{LA}, n_O^{LA}\}$  and  $\mathcal{C}_{Lin}^{EU} = \{\phi^{EU}, n_S^{EU}, n_O^{EU}\}$  for the LA-model and the EU-model, respectively. Here the subscript ‘Lin’ serves to distinguish the piecewise linear contracts from the nonlinear contracts described above.

**Finding optimal contracts.** We search for optimal contracts over all admissible parameters and generalize our notation by writing the wage function as  $w(P_T | \mathcal{C})$ , where  $\mathcal{C}$  can refer to either the LA-model or the EU-model and to the nonlinear contracts as well as the piecewise linear contracts.

Now consider a CEO for whom we can completely characterize the observed contract  $w^d(P)$ , where we use the superscript ‘ $d$ ’ in order to refer to ‘data.’ We represent observed contracts always as piecewise linear contracts  $\{\phi^d, n_S^d, n_O^d\}$ . Then:

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

where  $K$  is the strike price of the option. Our null hypothesis is that  $w^d(P_T)$  is an optimal contract, so it can be rationalized as the outcome of an optimization program, where we assume that preferences are parameterized as in (1) (for the LA-model) or as in (3) (for the EU-model) and that the technology is parameterized as in (9).<sup>13</sup> 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 solve the following program numerically:

$$\min_{\mathcal{C}} \pi(w(P_T | \mathcal{C})) \equiv \int w(P_T | \mathcal{C}) f(P_T) dP_T \quad (14)$$

$$s.t. \int V(w(P_T | \mathcal{C})) f(P_T) dP_T \geq \int V(w^d(P_T)) f(P_T) dP_T , \quad (15)$$

$$\int V(w(P_T | \mathcal{C})) \frac{\partial f(P_T)}{\partial P_0} dP_T \geq \int V(w^d(P_T)) \frac{\partial f(P_T)}{\partial P_0} dP_T . \quad (16)$$

By writing  $P_T$  as in (9) and setting  $P_0(e)$  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 do not consider the second stage and focus only on the first stage by

---

<sup>13</sup>The program is specified in (43) to (45) in the appendix.

solving program (14) to (16) as it does not depend on knowledge of the cost function  $C(e)$  or of the production function  $P_0(e)$ . This implies also that we cannot analyze the optimal *level* of incentives (pay for performance sensitivity) for a compensation contract, which would invariably depend on this information. However, with our approach we can analyze the optimal *structure* of compensation contracts for any given level of incentives.

Program (14) to (16) generates a new contract  $w^*(P_T)$  that is less costly to shareholders and specifies the parameters of the optimal contract. Condition (16) 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.<sup>14</sup> Similarly, condition (15) ensures that the contract found by the program provides at least the same value to the CEO as the observed contract, so it should also be acceptable to the CEO. We can then compare the observed contract  $w^d(P_T)$  to the optimal contract  $w^*(P_T)$  from (14) to (16).

## 4.2 Data

We identify all CEOs in the ExecuComp database who are CEO at least from January 2004 to December 2005. We restrict ourselves to CEOs in order to avoid multiple observations from one firm that are likely to be correlated. We also delete all CEOs who were executives in more than one company in either 2004 or 2005. We estimate the CEOs' contracts in 2005 and, separately, in 2004 as described shortly. We only analyze the 2005 contracts empirically. 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  $\sigma^2$ , and the proportion of shares owned by the CEO  $n_S$  from the 2004 data, while the fixed salary  $\phi$  is calculated from 2005 data<sup>15</sup>.

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

<sup>14</sup>This is subject to the assumed validity of condition (8).

<sup>15</sup> $\phi$  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.

in the estimated portfolio by 0.7 before calculating the representative option (see 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 moneyiness  $K/P_0$  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 point  $w^R$ .

**Minimum wage.** We do not have a good theory of the minimum wage in the context of our analysis. We reason that the CEO could be hired into another job with a similar compensation to her current job. However, it seems unlikely that she could obtain such a job offer when her previous company significantly underperformed expectations. It is also not plausible that her new employer would compensate her for giving up restricted stock or stock options that are practically worthless if the stock price drops below  $\hat{P}$ . Moreover, the minimum wage assumed for solving the optimization program (14) to (16) should not reflect the lower bound on the support of the observed wage distribution. A good model should be able to generate this lower bound



from the parameters and assumptions of the model and not from this additional constraint, which only reflects additional institutional restrictions that are not reflected in the model itself. For most part of our analysis we will use zero as a natural lower bound on the minimum wage, but we will perform some robustness checks where we allow the minimum wage to become negative.

**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 has formed 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 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:

$$w_{2005}^R(\theta) = \phi_{2004} + \theta \cdot MV(n_{2004}^S, n_{2004}^O, P_{2005}) , \quad (17)$$

The parameter  $\theta$  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  $\theta$ .

**Preference parameters.** For the preference parameters  $\alpha$  and  $\lambda$  we rely on the experimental literature for guidance. We therefore use  $\alpha = \beta = 0.88$  and  $\lambda = 2.25$  as our baseline values.<sup>16</sup>

---

<sup>16</sup>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

## 5 The baseline case: Contracts with restricted stock and options

The general nonlinear contract (10) can be implemented only by using a continuum of securities and can therefore not be directly compared to observed contracts that are piecewise linear. We therefore base our first analysis on a discussion of piecewise linear contracts and consider general non-linear contracts later. For our base case we also assume that option awards  $n_O$  and base salaries  $\phi$  cannot be negative. This seems not only realistic, but we will show later that this also biases our testing procedure against the LA-model, and since our objective is to show how well this model performs relative to the EU-model we want to make the case for the EU-model as strong as possible.

We therefore want to compare the observed contract  $\mathcal{C}^d = \{\phi^d, n_S^d, n_O^d\}$  with the optimal piecewise linear contracts  $\mathcal{C}^M = \{\phi^M, n_S^M, n_O^M\}$ , where the superscript  $M$  denotes the contracts predicted by model  $M \in \{EU, LA\}$  for each CEO. Minimization of program (14) to (16) is subject to the additional constraints  $n_O \geq 0$  and  $\phi \geq 0$ . We use the following distance metric  $D_{Lin}$  in order to compare optimal contracts to observed contracts:

$$D_{Lin} = \frac{1}{N} \sum_{i=1}^N \sqrt{\underbrace{\left(\frac{\phi^* - \phi^d}{\sigma_\phi}\right)^2}_{error(\phi)} + \underbrace{\left(\frac{n_S^* - n_S^d}{\sigma_S}\right)^2}_{error(n_S)} + \underbrace{\left(\frac{n_O^* - n_O^d}{\sigma_O}\right)^2}_{error(n_O)}}, \quad (18)$$

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

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

Here summation is over all  $N = 916$  CEOs in the sample and arithmetic averages 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 (2003).

[Insert Table 2 here]

Table 2 Panel A summarizes the results for the LA-Model for six different levels of the reference wage as parametrized by  $\theta$  (see (17)). Panel B shows the results for the EU-model in a comparable range see Abdellaoui (2000) and Abdellaoui, Vossman, and Weber (2005).

for five values of the coefficient of relative risk-aversion  $\gamma$ . For each model we show the means of the contract parameters and the means and the medians of the errors  $error(\phi)$ ,  $error(n_S)$ , and  $error(n_O)$ . These are defined as the scaled deviations of the observed value and the model value and form the components of  $D_{Lin}$  as indicated in (18).

The EU-model consistently underpredicts base salaries and options while overpredicting stock. For the LA-model, we observe a similar pattern in mean errors for very low and very high  $\theta$ , but the opposite pattern for intermediate values of  $\theta$ . However, the distributions are highly skewed, particularly for the LA-model, which tends to underpredict the base salary of the median CEO and it overpredicts base salaries on average for all values of  $\theta$  between 0.2 and 0.8. Similarly, the LA-model predicts lower option holdings for the median CEO, but higher option holdings on average for intermediate  $\theta$ -values. By comparison, 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 base salaries and option holdings: lower option holdings are always matched by higher holdings of the firm's stock. This is a necessity of both models, since the incentive compatibility constraint (6) ensures that overall incentives correspond to those of the observed contract, so lower option holdings imply higher shareholdings. However, whenever the model replaces some options with stock, then 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.

Our results for the distance metric  $D_{Lin}$  in Table 2 depend strongly on the parameterization of each model, in fact, more than on the model type itself. The EU-model generates consistently better forecasts as risk aversion  $\gamma$  decreases, whereas the accuracy of the LA-model is non-monotonic: the lowest distances between observed and model contracts occur for the lowest levels and the highest levels of the reference wage ( $\theta = 0$  and  $\theta = 1$ ) while the highest distances occur for intermediate levels of  $w^R$  ( $\theta$  between 0.4 and 0.6).

For both models this is a reflection of the fact that the model contracts are always closer to the optimal contracts if the CEO becomes more risk-neutral. To see this, consider parameterizations where the CEOs become completely risk-neutral. For the EU-model this would be the case for  $\gamma = 0$ . For the LA-model this would require either  $w^R = 0$  in order to eliminate the loss space, together with choosing  $\alpha = 1$ , or setting  $w^R$  at a very high value so as to eliminate the gain space, together with  $\beta = 1$ .

The intuition is that the optimization in (4) effectively minimizes the risk premium that shareholders have to pay the CEO for bearing uninsurable risk. Hence, if we eliminate loss aversion (respectively, risk aversion), then we also eliminate this risk premium and all contracts that generate the same pay for performance-sensitivity as the observed contract are potentially optimal: the difference in costs between options and stock can always be balanced by an appropriate adjustment of the base salary (with the provision that this cannot be negative). Hence, with risk neutrality the observed contract is always an optimal contract and  $D_{Lin} = 0$  is always (and trivially) feasible.

With this insight we can now develop an intuition for the non-monotonicity of  $D_{Lin}$  in the reference wage. For the LA-model the importance of loss aversion depends critically on the reference wage  $w^R$ . Table 2 reports the median probability that the LA contract pays off in the loss space,  $\Pr(w(P_T) \leq w^R)$ . For brevity we refer to this as the probability of loss. If this probability is zero, so that the contract is always guaranteed to pay off in the gain space, then the LA-model simplifies to a version of the EU-model where  $\gamma = 1 - \alpha$ . So, for our parameterization we then have  $\gamma = 0.12$ . Conversely, if  $w^R$  becomes so high that the probability of a loss is close to one, then the LA-model predicts even risk-seeking behavior. By contrast, loss aversion becomes most important if the probability distribution is centered around the point where  $w = w^R$ . At this point, the value function (2) is not differentiable and therefore infinitely concave, implying a strong disinclination to bearing risk. It is therefore intuitive that whenever the probability of a loss is in the intermediate range (40% to 50%) then the LA-model has greater difficulty in predicting observed contracts.<sup>17</sup> We investigate this problem further below for the case of general non-linear contracts.

One implication of this discussion is that we cannot find an optimal parameterization of either model purely by looking at the distance  $D_{Lin}$  between observed contracts and model contracts as this distance converges to zero as loss aversion, respectively, risk aversion, becomes smaller. We therefore need to restrict what we consider to be reasonable parameter ranges based on considerations outside the scope of our models. A large literature in economics and finance investigates risk aversion, unfortunately without establishing a consensus on plausible parameter ranges for the coefficient of relative risk aversion  $\gamma$ . The literature on executive compensation has often discussed values for  $\gamma$  in the range between 2 and 3.<sup>18</sup> A useful point of

<sup>17</sup>Note that the probability of loss is not zero for  $\theta = 0$  as  $w^R(\theta = 0)$  is equal to the base salary of the previous period. Hence, whenever the base salary of a CEO is lower in 2005 than it was in 2004, there is some probability that the observed contract pays off below the previous base salary.

<sup>18</sup>Hall and Murphy (2000) use these values that seem to go back to Lambert, Larcker, and Verrecchia (1991).

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.<sup>19</sup> 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 0.2, even though we regard such values as implausible. For the LA-model, no such comparison is available since the notion of framing in this context suggests a separate analysis of portfolio behavior and of the evaluation of compensation contracts. However, based on our arguments in Section 4.2 above we regard values of  $\theta$  outside the unit interval as implausible and inconsistent with the notion of a reference wage.

The strong dependence of the fit of the two models to their parameterizations implies that comparisons of the LA-model and the EU-model have to be based on *comparable* parameterizations as these are critical to the performance of both models. We therefore need an additional variable that we hold constant across models. We propose to compare parameterizations that generate the same valuation of the observed contract by the same CEO. More specifically, we define the certainty equivalent of model  $M$  from:  $E(V^M(w^d(P_T))) = CE^M$ . We fix  $\theta$  to establish the reference wage of each CEO and then define an equivalent  $\gamma$  by:

$$CE^{LA}(\theta) \equiv CE^{EU}(\gamma_e) . \quad (19)$$

We refer to the value of  $\gamma_e$  that satisfies (19) 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  $\theta$  we calculate the equivalent  $\gamma_e$  and the optimal EU-contract with  $\gamma = \gamma_e$ . Table 3 compares the two models.

[Insert Table 3 here]

Table 3 reports mean and median of the distance metric  $D_{Lin}$  for both models and the difference  $D_{Lin}^{EU} - D_{Lin}^{LA}$ . We test whether this difference is significantly different from zero with the standard t-test, the Wilcoxon signed rank test, and the sign test. All tests are applied to the

---

Lambert and Larcker (2004) more recently proposed a value as low as 0.5.

<sup>19</sup>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.,  $\gamma = 0.2$ , the lowest value considered in Table 2, would imply that the CEO invests five times her wealth in the stock market.

differences  $D_{Lin}^{EU} - D_{Lin}^{LA}$  and test the null hypothesis that the mean or, respectively, the median is equal to zero.

The distribution of  $D_{Lin}$  is skewed, so we sometimes obtain conflicting results for the means and for the medians. The LA-model outperforms the EU-model for 45% - 80% of the CEOs depending on the parameterization, and at the median this is statistically significant for all parameterizations (except  $\theta = 0.2$ , the only value where the EU-model dominates) based on the Wilcoxon test, while the sign test is also insignificant for  $\theta = 0.3$ . The average distance between observed contracts and model contracts is generally larger for the LA-model, but statistically significant only for  $\theta$ -values between 0.1 and 0.4. Overall, this suggests that the LA-model dominates the EU-model in most cases. However, when it fails, then its failure is often more extreme than that of the EU-model, which gives rise to a more skewed distribution of  $D_{Lin}$ . This motivates our further analysis of the comparative strengths and weaknesses of both models. The equivalent  $\gamma_e$ 's are generally very low and below the range we regard as plausible (see above, particularly footnote 19). We note also that they are non-monotonic in  $\theta$ : 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.

[Insert Table 4 here]

Table 4 investigates how the metric  $D_{Lin}^M$  is correlated with the option holdings and the stock holdings of the CEOs for different parameterizations. The correlations between  $D_{Lin}^{EU}$  and option holdings are large and significantly different from zero. Hence, the EU-model turns out to be particularly bad for CEOs who own many options. This is simply a reflection of the fact that the EU-model consistently underpredicts options. On the other hand, the LA-model performs particularly poorly for CEOs with high stock holdings, because it often underpredicts stock in an attempt to insure the CEO against downside risk.

Figure 1 illustrates the comparative strengths and weaknesses of both models by looking at two extreme cases. The left hand panel of the figure shows a CEO (Warren E. Buffet, Berkshire Hathaway) with no options and a large share ownership of his company. Here the EU-model predicts the optimal contract correctly, whereas the LA-model suggests to increase his meagre base salary from \$309,000 to \$8.2 bn, to reduce his stock holdings from 32.7% to 21.4%, and to replace the stock with options on 15.4% of the firm. The right hand panel of the figure shows a CEO (R. Chad Dreier, Ryland Group Inc.) who owns only 0.75% of his firm but holds about

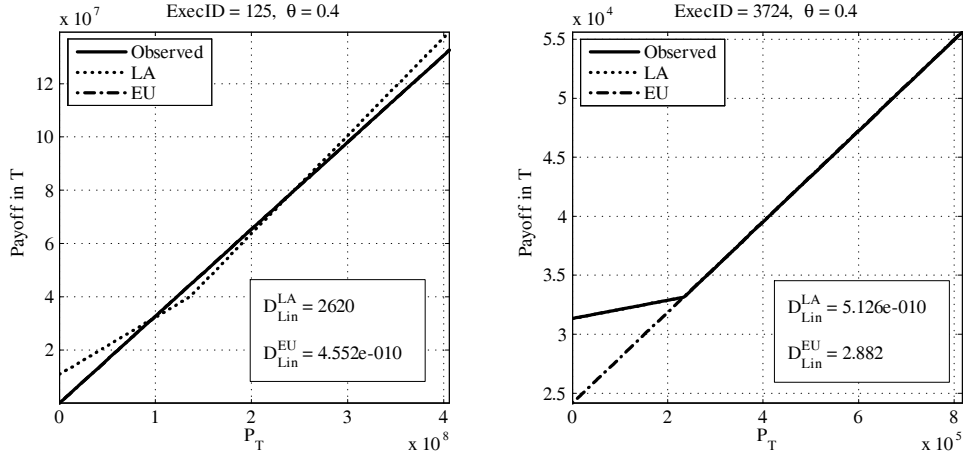


Figure 1: The left hand panel shows the CEO with ExecuComp ID #3724 (R. Chad Dreier, Ryland Group Inc.), who has parameters  $n_S = 0.75\%$ ,  $n_O = 3.08\%$  and  $\phi = 27,013$ . The right hand panel shows the CEO with ID #125 (Warren E. Buffet, Berkshire Hathaway), who has parameters  $n_S = 32.68\%$ ,  $n_O = 0\%$  and  $\phi = 309$ .

four options for every share he owns. This contract is predicted correctly by the LA-model, whereas the EU-model suggests to eliminate all options, to increase the stock holdings from 0.75% to 3.79%, and to reduce the base salary by 23% from \$27.0m to \$20.8m.

In our view it seems implausible that a principal-agent model should be able to generate contracts like those of Warren Buffet who are effectively owner-managers of their companies rather than salaried agents of outside shareholders who need to be incentivized to provide adequate effort. The LA-model generates contracts where shareholders insure the CEO against downside risk, but such protection is generally not available to owner-managers. The ownership structure and governance structure of these companies is arguably outside the scope of a simple principal-agent model.

[Insert Table 5 here]

We investigate this issue further by splitting the sample into those CEOs who have large stakes in excess of 5% of all shares in their own firms (96 cases, or 10.5% of the sample) and those who have positive option holdings (817 cases, or 89.5% of the sample). Table 5, Panel A shows that the results for the subsample of CEOs with large shareholdings (on average 14.4% of their companies). The overprediction of their base salaries is often very large (up to a factor of about 100) and the LA-model suggests that contracts that replace up to one third of their shares with options would be beneficial. The statistics for  $D_{Lin}$  are accordingly large, so that these

results have a strong impact on the t-tests for the whole sample. Panel B of Table 5 reports the results for the subsample with small shareholdings (on average 0.62%, compared to average option holdings of 1.37%). The EU-model underpredicts options for both subsamples, but the relative weight is much larger in the subsample with small stock ownership. In many cases the EU-model generates corner solutions at  $n_O = 0$ . We will study these corner solutions separately below.

[Insert Table 6 here]

The apparent success of the EU-model also deserves further investigation, since it rests critically on its ability to predict contracts without options, which is a corner solution. Table 6, Panel A documents the frequency with which each model predicts either positive base salaries (so that the non-negativity constraint on  $\phi$  is not binding), or positive option holdings (hence, the constraint on  $n_O$  is not binding), or both. We can see that for the EU-model one of the two constraints is almost always binding. This suggests that positive option holdings result primarily from the downward constraint on base salaries, as options then cannot be exchanged any longer for a combination of shares and salary cuts. Conversely, for the LA-model we obtain mostly interior solutions for moderate levels of the reference wage.

Panel B of Table 6 reports the same results for a model where the non-negativity constraints are not imposed. Instead, we require that the net slope of the wage function cannot become negative and that the total gross payoff ( $\phi + W_0$ ) cannot become negative.<sup>20</sup> Clearly, neither negative option holdings nor negative base salaries are parts of observed contracts, but we reiterate that a good model should generate these results from the assumptions of the model itself and not from the additional restrictions we impose. Interestingly, the LA-model can generate positive option holdings in 31% to 81% of the cases, where the higher option holdings correspond to lower reference wages. Also, the EU-model predicts negative base salaries in more than 97% of all cases, whereas the LA-model predicts positive base salaries for the majority of all CEOs for moderate levels of the reference wage. Hence, from the perspective of the EU-model it would be optimal to have the CEOs invest a significant part of their wealth into their firms' stock, whereas the LA-model implies this only for high levels of  $w^R$ .

The distance metric  $D_{Lin}$  shows a dramatic shift in favor of the LA-model when we move from the restricted model to the unrestricted model. Now  $D_{Lin}$  obtains almost the same mag-

---

<sup>20</sup>The first constraint implies that  $n_O \geq -n_S \exp(dT)$  since we need to adjust the number of shares for the dividend yield, assuming that the options are not dividend protected.



nitudes for observations that are problematic for the EU-model as the LA-model does for the owner-CEOs we study in Table 5. The most striking observation is that the EU-model can almost never predict positive base salaries and positive option holdings simultaneously. We therefore attribute the remarkable ability of the EU-model to correctly predict the contracts of those CEOs without options to the non-negativity constraint on option holdings, which prevents the algorithm from generating concave contracts.

**Preliminary conclusion.** Overall, we conclude that the LA-model outperforms the EU-model by a number of criteria and for most parameterizations if we compare matched results for both models on the basis of equivalent risk premia. However, if the LA-model fails, then it fails spectacularly as it cannot generate contracts for CEOs with high shareholdings and only few options. By contrast, the EU-model fails to predict options and generates contracts with negative option holdings unless we impose a non-negativity constraint on options or on base salaries, so its comparative success at explaining the observations that are problematic for the LA-model in Tables 2, 3, and 5 above seems to be an artificial outcome.

## 6 Extensions and Robustness Checks

### 6.1 General nonlinear contracts

One drawback of the methodology in the previous section is that it relies on a stylized representation of the contracts. However, our theoretical analysis above shows that the optimal contract is highly non-linear, and some of the results on stylized piecewise linear contracts might be an artefact of the restrictions our stylized representation imposes on the optimal contracts generated by the models. In this section we therefore analyze the optimal nonlinear contracts generated by both models.

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  $w^R$ . We interpret this jump as firing the CEO if the stock price falls below  $\hat{P}$ . Dismissal is not an explicit part of the CEO's contract with the firm. Rather, contracts are negotiated for a limited period of time and not extended, or terminated prematurely as the result of negotiations between the board of directors and the CEO. In these cases the governance structure of the company basically provides the legal context, and we include this in our concept of the optimal contract. We therefore define the dismissal probability  $p$  of the optimal model contract as  $p(\hat{P}) \equiv \int_0^{\hat{P}} f(P_T) dP_T$ . For the EU-model we use

an analogous definition where  $\hat{P}$  is defined as the highest stock price where the wage function (12) drops to the lower bound  $\underline{w}$ . A good model of efficient contracting should generate realistic dismissal probabilities. While we do not have dismissal probabilities for individual CEOs, we can establish realistic ranges from the data. For this we look at the unconditional probability for a CEO on the ExecuComp database to leave the company for reasons other than retirement over any four-year period between 1992 and 2004 and establish that this equals 7.4%. However, not all of these are disciplinary dismissals due to underperformance and the literature on CEO turnover has not always found strong connections between stock price performance and CEO dismissals, so this number has to be regarded as an upper bound on a reasonable dismissal probability for our sample.<sup>21</sup>

The conceptual difficulty in comparing general nonlinear contracts to the data lies in the fact that contracts like (10) cannot be implemented using 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 contracts for both models have ranges where they are convex and ranges where they are concave, and the concave parts 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 rather complex contracts observed in practice to a stylized representation in terms of base salaries, stock, and one option grant.

We address these issues in two steps by first developing some heuristics that allow us to compare the model contracts to the observed contracts. In a second step we will then develop formal distance measures analogous to (18). The first set of measures looks at the average slopes of the nonlinear contracts. We define:

$$\Delta_{Low} \equiv \int_0^K \frac{\partial w^*(P_T)}{\partial P_T} \frac{f(P_T)}{F(K)} dP_T, \quad (20)$$

$$\Delta_{High} \equiv \int_K^\infty \frac{\partial w^*(P_T)}{\partial P_T} \frac{f(P_T)}{1 - F(K)} dP_T. \quad (21)$$

Here  $\Delta_{Low}$  is the average slope in the region below the strike price of the option, which can

---

<sup>21</sup>See Weisbach (1988) and Kaplan (1994) for earlier papers in this literature and Engel, Hayes, and Wang (2003) and Farrell and Whidbee (2003) for more recent contributions. Brickley (2003) states in the discussion of the last two papers that he is "struck by the limited explanatory power of the various performance measures in the CEO turnover regressions," which emphasizes our point that performance-related dismissal probabilities are low.

be compared to the number of shares  $n_S$ .  $\Delta_{High}$  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) and (12) we can determine the inflection point  $P_I$  of each contract, so that the contract is convex for all terminal stock prices below  $P_I$ , and we use the probability that the model contract pays off in the convex range,  $\Pr(w^*(P_T) \leq P_I)$  as another descriptive statistic.<sup>22</sup>

[Insert Table 7 here]

Table 7 reports the average slopes  $\Delta_{Low}$  and  $\Delta_{High}$ , 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,  $\Delta_{High}$  is almost always higher than the slope in the lower range,  $\Delta_{Low}$ . 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. By contrast, the EU-model generates contracts that are on average convex only over the lower 20% to 30% of the probability distribution and concave otherwise.

The dismissal probabilities are unrealistically high for the LA-model once the reference point becomes sufficiently high ( $\theta$ -values above 0.4). As the reference wage increases, the threat of dismissals becomes more important. In some sense, 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  $\Delta_{Low}$  declines as  $w^R$  increases) and to a larger extent through the threat of dismissals. For the EU-model, the probability of dismissals is above the realistic range for all parameterizations and increases with risk aversion, but we should be careful in interpreting these results since the EU-model does not have a discontinuous jump (the wage function (12) is continuous at  $\bar{P}$ ), so there are no additional incentives from dismissals.

Our next step is to compare the nonlinear contracts for both models. Unfortunately, there is no obvious and intuitive metric comparable to (18) that can be used for this purpose, so we

---

<sup>22</sup>There are some CEOs where  $\hat{P} \leq P_I$ , so the LA-contract for these has a slope of zero up to the discontinuity and then becomes concave. For these we evaluate  $\Pr(w^*(P_T) \leq \hat{P})$ . A similar case frequently occurs for the EU-model when  $\bar{P} \leq P_I$ .

develop three alternative measures that focus on different aspects of the contract. Based on (20) and (21) we can define a nonlinear analogue to  $D_{Lin}$  as:

$$D_{NonLin} = \frac{1}{N} \sum_{i=1}^N \sqrt{\left( \frac{\Delta_{Low}^{*,i} - \Delta_{Low}^{d,i}}{\sigma_{Low}} \right)^2 + \left( \frac{\Delta_{High}^{*,i} - \Delta_{High}^{d,i}}{\sigma_{High}} \right)^2} \quad (22)$$

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

Here,  $\Delta_{Low}^{d,i}$  and  $\Delta_{High}^{d,i}$  represent the slopes of the observed contract corresponding to (20) and (21) and  $\bar{\Delta}_{Low}^d$  and  $\bar{\Delta}_{High}^d$  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.

One disadvantage of  $D_{NonLin}$  as defined in (22) is that it refers only to the slope and not to the base salary of the contract, so it does not capture differences in levels. A natural metric that avoids this shortcoming is:

$$D_{Level} = \frac{1}{N} \sum_{i=1}^N \left[ \int_0^\infty \left( \frac{w^{*,i}(P_T) - w^{d,i}(P_T)}{w_0^{d,i}} \right)^2 f^i(P_T) dP_T \right]^{1/2}, \quad (23)$$

$$\text{where : } w_0^{d,i} = \int_0^\infty e^{-rT^i} w^{d,i}(P_T) f^i(P_T) dP_T. \quad (24)$$

The best way to interpret  $D_{Level}$  is as an extension of the notion of a mean squared error to a function space.<sup>23</sup> We also evaluate the following analogue to  $D_{Level}$ :

$$D_{Slope} = \frac{1}{N} \sum_{i=1}^N \left[ \int_0^\infty \left( \frac{\partial w^{*,i}(P_T)}{\partial P_T} - \frac{\partial w^{d,i}(P_T)}{\partial P_T} \right)^2 f^i(P_T) dP_T \right]^{1/2}. \quad (25)$$

[Insert Table 8 here]

$D_{Slope}$  also focuses on the slope of the wage function, but it does not depend on the strike price  $K$  of the representative option and on the scaling by sample standard deviations. Table 8 reports  $D_{NonLin}$ ,  $D_{Level}$ , and  $D_{Slope}$  for the nonlinear contracts of the LA-model and the EU-model. We observe that  $D_{Slope}$  always favors the LA-model, whereas  $D_{NonLinear}$  favors the

<sup>23</sup>Note that (23) is also close to the definition of a norm in a space of random variables. If  $x$  is a random variable that is distributed with some probability law  $g(x)$ , then we can define  $\|x\|_g = (\int x^2 g(x) dx)^{1/2}$  and the squared expression in (23) then equals  $\left\| (w^{*,i}(P_T) - w^{d,i}(P_T)) / w_0^{d,i} \right\|_{f^i}$ .

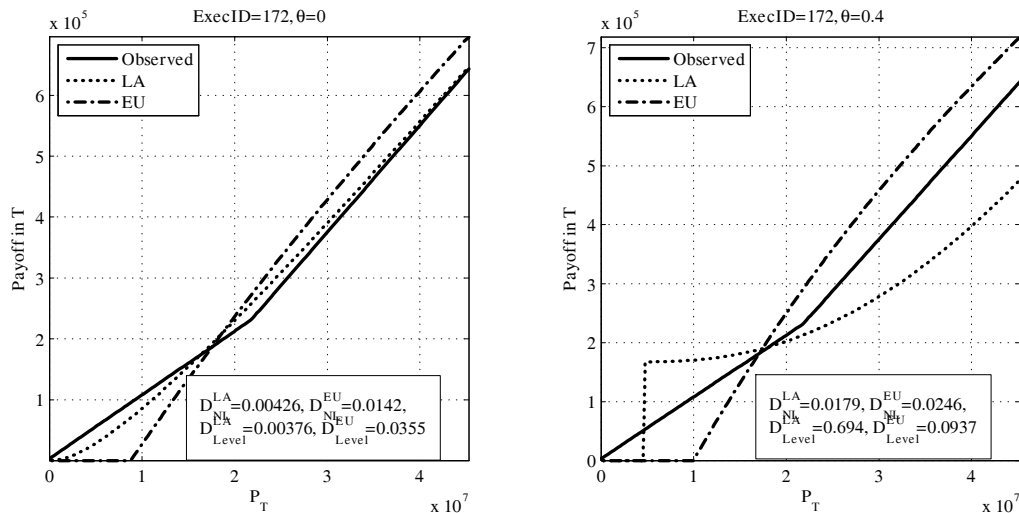


Figure 2: The figure shows CEO # 172 for  $\theta = 0$  (left panel) and for  $\theta = 0.4$  (right panel). This CEO has parameters  $n_S = 1.04\%$ , and  $n_O = 0.71\%$ .

EU-model for high reservation wages, and  $D_{Level}$  favors the EU-model for high and also for moderate reservation wages.

Hence, the EU-model seems to be better at capturing the level of payouts, whereas the LA-model performs better at capturing the slopes of the observed contracts. Figure 2 illustrates this aspect for a typical case. For  $\theta = 0$  the LA-model tracks the observed contract much better than the EU-model, but for  $\theta = 0.4$  the LA-model's performance deteriorates sharply. Here the LA-model performs worse in terms of levels because it features the possibility of firing the CEO, which is not part of our representation of the observed contract. As a result, the model contract is below the observed level for low stock prices (where the CEO is fired) and then jumps discretely to value above  $w^R$  that is significantly above the observed level. However, on either side the deviation from the observed contract can be significant and is in most cases larger than for the EU-model, which does not share this feature.

[Insert Table 9 here]

Compared to the previous section, where the LA-model mostly dominates the EU-model when we restrict the contract space to piecewise linear contracts, we do not find such clear results for the general contract in Tables 7 and 8. Therefore, we compare the piecewise linear contracts and the nonlinear contracts for the LA-model in Table 9. The table shows the averages for stock holdings  $n_S$  and compares those to the average slope in the low stock price range  $\Delta_{Low}$ ,

and similarly the average for stocks and options combined, which we compare with  $\Delta_{High}$ . We report the metric  $D_{NonLin}$  for both cases and also introduce the new variable "Savings," defined as

$$Savings = \frac{E(w^d(P_T)) - E(w^*(P_T))}{E(w^d(P_T))}, \quad (26)$$

or, in words, the percentage reduction in the costs of the optimal contract predicted by the model compared to the observed contract.

We can see that the performance of the nonlinear model becomes poorer as the reference wage increases, whereas the quality of approximation of the piecewise linear model is monotonic, as noted before. Interestingly, the average security holdings predicted by the piecewise linear contract are not affected much by changes in the reference wage, whereas the slopes  $\Delta_{Low}$  and  $\Delta_{High}$  both decline in the reference wage, consistent with the argument above that an increased threat of dismissal replaces incentives through security holdings as the reference wage increases.

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, which indirectly 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 tight monitoring of the CEO in order to enforce dismissals and the fall in wages associated with a significant drop in the stock price. E.g., boards should not endorse significant severance payments. The difference in savings between the piecewise linear contract and the general nonlinear contract would have to be related to the costs of implementing such a governance structure. These savings are in the range of 0.1% to 7.3%, where the higher estimates correspond to the highest (and least plausible) assumptions on the reference wage. We are not aware of estimates of the costs of enforcing CEO turnover, but we suspect that these costs are higher than these estimates of the potential savings from recontracting.

## 6.2 Comparative Static Analysis

Finally, we investigate to what extent our results are sensitive to parameter assumptions. We have based our discussion on the estimates of  $\alpha$ ,  $\beta$ , and  $\lambda$  on the experimental literature, and this may well be inappropriate for the study of CEOs.

[Insert Table 10 here]

Table 10 reports the results of a comparative static analysis in terms of the preference parameters. We report the same parameters as before. The results for the piecewise linear model are hardly affected by changes in  $\beta$  and  $\lambda$ , whereas higher values for  $\alpha$  are associated with lower values for  $D_{Lin}$  and also substantially lower savings. As  $\alpha$  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: (1) For the median CEO, the Loss Aversion-model predicts observed contracts more closely compared to the Expected Utility-model based on our tests. (2) The Loss Aversion-model can explain the prevalence of stock options in observed compensation contracts. It generates interior solutions for option holdings and the base salary for realistic parameterizations, whereas the Expected Utility-model does not. (3) 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) and for unrealistically low levels of risk aversion.

However, the strength of the Loss Aversion-model turns into a weakness for those CEOs who have no stock options and own large fractions of the stock of their companies. We suggest that the contracts of these CEOs should not be analyzed in the context of a principal-agent model based on independent shareholders who negotiate a contract with a salaried agent.

Our results are of particular importance to the large literature on the design and the valuation of executive stock options that relies on versions of the Expected Utility-model so far (see Footnote 2 in the Introduction). We therefore suggest that for all these applications to the typical CEO who is a salaried agent, choosing the Loss Aversion-model is more useful than

relying on the Expected Utility-model. Our analysis shows that for these applications it is useful to choose relatively low reference wages that are set close to the previous fixed salary (including bonus payments) and assume that the CEO applies steep discounts to her deferred compensation in this respect.

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.

We have used only some of the components of prospect theory by using the value function proposed by Kahneman and Tversky. We have neglected the other component, namely the probability weighting function. From the point of view of prospect theory this is a compromise since risk aversion is modeled through the decision weights as well as through the value function. However, at this point an inclusion of the probability weighting function appears analytically intractable as we would have to find conditions that preserve the monotone likelihood ratio property after transforming the decision weights.



## A Appendix

### Proof of Proposition 1:

In order to prove the proposition, we will prove two useful lemmas first. To simplify notation slightly, we define

$$U_l(w^R - w(P_T)) = \lambda(w^R - w(P_T))^\beta \quad (27)$$

$$U_g(w(P_T) - w^R) = (w(P_T) - w^R)^\alpha. \quad (28)$$

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

### Proof of Lemma 1:

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

$$g(P_T)U_l(w^R - w^R) + (1 - g(P_T))U_l(w^R - \underline{w}) = U_l(w^R - w(P_T)). \quad (29)$$

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

$$g(P_T)U_l(w^R - w^R) + (1 - g(P_T))U_l(w^R - \underline{w}) \leq U_l(w^R - (g(P_T)w^R + (1 - g(P_T))\underline{w})). \quad (30)$$

Combining equations (29) and (30) yields:

$$U_l(w^R - w(P_T)) \leq U_l(w^R - (g(P_T)w^R + (1 - g(P_T))\underline{w})). \quad (31)$$

$U_l$  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  $\{g(P_T), w^R; 1 - g(P_T), \underline{w}\}$  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  $\{g(P_T)p, w^R; (1-g(P_T))p, \underline{w}\}$ .

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

From Lemma 1, we can write contracts as a combination of a payoff function in the gain space and a lottery over the minimum wage and the reference wage,  $\{g(P_T), w_g(P_T)\}$ . The corresponding optimization problem then becomes:

$$\min_{g(P_T), w_g(P_T), I(P_T)} \int [I(P_T)w_g(P_T) + (1 - I(P_T))(g(P_T)w^R + (1 - g(P_T))\underline{w})] f(P_T|\bar{e})dP_T \quad (32)$$

$$\begin{aligned} \text{s.t. } & \int [I(P_T)U_g(w_g(P_T) - w^R) - (1 - I(P_T))(1 - g(P_T))U_l(w^R - \underline{w})] f(P_T|\bar{e})dP_T \\ & \geq \underline{V} + C(\bar{e}), \end{aligned} \quad (33)$$

$$\int [I(P_T)U_g(w_g(P_T) - w^R) - (1 - I(P_T))(1 - g(P_T))U_l(w^R - \underline{w})] \Delta f(P_T|e)dP_T \geq \Delta C, \quad (34)$$

where  $I(P_T)$  is an indicator function which is one if the contract pays off in the gain space and zero otherwise.

**Lemma 2.** (i) Whenever the optimal contract pays off in the gain space it satisfies the condition

$$\frac{1}{U'_g(w_g^*(P) - w^R)} \geq \mu_{PC} + \mu_{IC} \frac{\Delta f(P|e)}{f(P|\bar{e})}, \quad (35)$$

where (35) holds as an equality for all interior wages  $w_g^*(P) > w^R$  and  $w_g^*(P) = w^R$  if the inequality is strict.  $w_g^*(P_T)$  is monotonically increasing in  $P_T$ .

(ii) If the optimal contract pays off in the loss space, then the manager receives  $\underline{w}$  for all  $P_T \leq P^R$  and she receives  $w^R$  for all  $P_T \geq P^R$ , where  $P^R$  is a uniquely defined cutoff value.

### Proof of Lemma 2:

(i) If the contract pays off in the gain space, then  $I(P_T) = 1$  and the first order condition

for  $w_g(P_T)$  of the corresponding Lagrangian is:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial w_g(P_T)} &= f(P_T|\bar{e}) - \mu_{PC} U'_g(w_g^*(P_T) - w^R) f(P_T|\bar{e}) - \mu_{IC} U'_g(w_g^*(P_T) - w^R) \Delta f(P_T|e) \\ &= U'_g(w_g^*(P_T) - w^R) f(P_T|\bar{e}) \left[ \frac{1}{U'_g(w_g^*(P_T) - w^R)} - \mu_{PC} - \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\bar{e})} \right] \geq 0 . \end{aligned} \quad (36)$$

Note that  $U'_g(w_g^*(P_T) - w^R) f(P_T|\bar{e}) > 0$ . Then the condition has to hold as an equality for all  $w_g^*(P_T) > w^R$ . Otherwise, if  $\frac{\partial \mathcal{L}}{\partial w} > 0$  over the entire gain space, then the solution is at the lowest possible value at the constraint  $w_g^*(P_T) \geq w^R$  is binding. From MLRP and (35) we can infer directly that the optimal contract is monotonically increasing in the gain space.

(ii) If the contract pays off in the loss space, then  $I(P_T) = 0$  and the first order condition for  $g(P_T)$  of the corresponding Lagrangian is:

$$\frac{\partial \mathcal{L}}{\partial g(P_T)} = f(P_T|\bar{e}) U_l(w^R - \underline{w}) \left[ \frac{w^R - \underline{w}}{U_l(w^R - \underline{w})} - \mu_{PC} - \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\bar{e})} \right] = 0 . \quad (37)$$

We have  $f(P_T|\bar{e}) U_l(w^R - \underline{w}) > 0$  by assumption. The only part of the expression in brackets that depends on  $P_T$  is  $\Delta f(P_T|e)/f(P_T|\bar{e})$ , which is increasing in  $P_T$  from assuming MLRP, hence there can be at most one cut-off point  $P^R$  that satisfies (37) as an equality. For any point  $P_T > P^R$  we have  $\frac{\partial \mathcal{L}}{\partial g(P_T)} < 0$ , so that  $\mathcal{L}$  is minimized by increasing  $g$  to its upper limit, so  $g = 1$ . Conversely, for any point  $P_T < P^R$  we have  $\frac{\partial \mathcal{L}}{\partial g(P_T)} > 0$ , so that  $\mathcal{L}$  is minimized by reducing  $g$  to its lower limit, so  $g = 0$ . Hence, interior probabilities  $0 < g < 1$  are never optimal and the optimal lottery is always degenerate. Then the resulting contract is deterministic with a cut-off value  $P^R$ .

Now we are in a position to prove Proposition 1. For notational convenience define  $x \equiv w_g(P_T) - w^R$ , and  $y = w^R - w_l(P_T)$ , i.e.  $y = w^R - \underline{w}$  if  $P < P^R$  and  $y = 0$  if  $P > P^R$ . Then, the Lagrangian becomes

$$\begin{aligned} \mathcal{L} &= \int [(1 - I(P_T)) w_l(P_T) + I(P_T) w_g(P_T)] f(P_T|\bar{e}) dP_T \\ &\quad + \mu_{PC} \left[ \underline{V} + C(\bar{e}) + \int [(1 - I(P_T)) U_l(y) - I(P_T) U_g(x)] f(P_T|\bar{e}) dP_T \right] \\ &\quad + \mu_{IC} \left[ \Delta C + \int [(1 - I(P_T)) U_l(y) - I(P_T) U_g(x)] \Delta f(P_T|e) dP_T \right] . \end{aligned} \quad (38)$$

Differentiating with respect to  $I(P_T)$  yields

$$\frac{\partial \mathcal{L}}{\partial I(P_T)} = f(P_T|\bar{e}) [U_l(y) + U_g(x)] \left[ \underbrace{\frac{x+y}{U_l(y) + U_g(x)}}_{>0} - \mu_{PC} - \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\bar{e})} \right] \quad (39)$$

As  $f(P_T|\bar{e}) [U_l(y) + U_g(x)] > 0$ , the term in the large brackets determines the sign of equation (39). Now we have to consider two cases:

**Case 1:**  $\mu_{PC} + \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\bar{e})} < 0$ . Since we assume MLRP, this can only be the case for all  $P_T$  smaller than some  $\tilde{P}_T$  for which  $\mu_{PC} + \mu_{IC} \frac{\Delta f(\tilde{P}_T|e)}{f(\tilde{P}_T|\bar{e})} = 0$ . In this case we then have from equation (39) that  $\frac{\partial \mathcal{L}}{\partial I(\tilde{P}_T)} > 0$ . Hence for all  $P_T < \tilde{P}_T$  it is optimal to set  $I(P_T)$  to its lowest possible level, zero. But this implies by construction that the contract always pays off in the loss space for all  $P_T \in (0, \tilde{P}_T)$ .

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

$$\frac{1}{U'_g(x)} = \mu_{PC} + \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\bar{e})}. \quad (40)$$

For all  $P_T$  where the contract pays off in the gain space, this is the exactly the condition for the optimal contract as established in Lemma 2. However, it should be noted that equation (40) is defined over all  $P_T \in (\tilde{P}_T, \infty)$  and not just over the gain space, which by Case 1 must be a subset of  $(\tilde{P}_T, \infty)$ . Hence at this point we presume nothing about whether the contract actually pays off in the loss space, or in the gain space, for any given  $P_T > \tilde{P}_T$ . Now, using (40) in (39) we get

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

where  $z(x, y) \equiv U'_g(x)(x+y) - U_l(y) - U_g(x)$ . Note that  $y$  is constant on the intervals  $(-\infty, P^R)$  and  $(P^R, \infty)$ . Hence,  $z(x, y)$  is a strictly decreasing function in  $x$  because  $z'(x) = U''_g(x)(x+y) < 0$  as  $U_g(\cdot)$  is concave. As  $x(P_T)$  defined by (40) is strictly increasing in  $P_T$ ,  $z(x, y)$  is strictly decreasing in  $P_T$  on these two intervals. Consequently, there can be at most two solutions to the first order condition  $\frac{\partial \mathcal{L}}{\partial I(P_T)} = 0$ : one for  $y = 0$  and one for  $y = w^R - \underline{w}$ . In the

first case,  $z(x, y) = 0$  is equivalent to  $U'_g(x)x - U_g(x)$  which smaller than zero. Consequently, there is at most one solution to the first order condition that defines a unique value  $\widehat{P}$  for which it holds that

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

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

$$U'_g(w^*(\widehat{P}) - w^R)(w^*(\widehat{P}) - \underline{w}) - [\lambda U_l(w^R - \underline{w}) + U_g(w^*(\widehat{P}) - w^R)] = 0 \quad (41)$$

Hence, we have established in Case 1 and Case 2, that loss space and gain space are non-empty intervals,  $(0, \widehat{P}_T)$  and  $(\widehat{P}_T, +\infty)$ . To establish that the optimal contract cannot feature a region in the loss space where  $w_l^*(P_T) = w^R$ , look again at equation (37) from the Proof of Lemma 2, which we state here again for convenience

$$\frac{\partial \mathcal{L}}{\partial g(P_T)} = f(P_T|\bar{e})U_l(w^R - \underline{w}) \left[ \frac{w^R - \underline{w}}{U_l(w^R - \underline{w})} - \mu_{PC} - \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\bar{e})} \right]. \quad (42)$$

This is zero if the term in the square brackets is zero, which can only be the case for  $\mu_{PC} + \mu_{IC} \frac{\Delta f(P_T|e)}{f(P_T|\bar{e})} > 0$ . By the same logic as before, we can rewrite this for  $P_T > \widetilde{P}_T$ , using (40) as

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

Comparing the term in square brackets in this equation with  $z(x, y)$ , and using the fact that  $U'_g(x)x < U_g(x)$  for all  $x \geq 0$ , we have that  $z(x, y)$  is always zero before the jump in the loss space from  $\underline{w}$  to  $w^R$  occurs, which is just what equation (37) determines. Hence the optimal contract pays off  $\underline{w}$  in the loss space for all  $P_T < \widehat{P}_T$ , and  $w_g^*(P_T)$  in the gain space for  $P_T > \widehat{P}_T$ , where  $w_g^*(P_T)$  can be found by solving equation (40) for  $w_g(P_T)$ . ■

### Proof of Proposition 2:

Shareholders' problem if they wish to minimize the contracting costs for implementing effort

level  $\hat{e}$  can be written as:

$$\min_{w(P_T) \geq \underline{w}} \int w(P_T) f(P_T|\hat{e}) dP_T \quad (43)$$

$$s.t. \quad - \int (1 - I(P_T)) U_l(w^R - w(P_T)) f(P_T|\hat{e}) dP_T \quad (44)$$

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

$$- \int (1 - I(P_T)) U_l(w^R - w(P_T)) f_e(P_T|\hat{e}) dP_T \quad (45)$$

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

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 (43) to (45) is pointwise, the only changes with respect to program (4) to (6) are: replace  $\Delta C$  with  $C'$ , which is a constant for a given level of effort in both programs; replace  $f(P_T|\bar{e})$  with  $f(P_T|\hat{e})$ , which is just a density that has the same properties in both programs; replace  $\Delta f(P_T|e)$  with  $f_e(P_T|\hat{e})$ , which also has the same properties in both programs as we assume MLRP in both cases. Hence, the same arguments as in Lemmas 1 and 2 and in Proposition 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  $\ln(P_T)$  is distributed normal with mean  $\mu(e) = \ln(P_0(e)) + \left(r_f - \frac{\sigma^2}{2}\right)T$  and standard deviation  $\sigma\sqrt{T}$ . The density  $f(P_T|e)$  of the lognormal distribution is then:

$$f(P_T|e) = \frac{1}{P_T\sqrt{2\pi T}\sigma} \exp\left\{-\frac{[\ln P_T - \mu(e)]^2}{2\sigma^2 T}\right\} \quad , \quad (46)$$

and the likelihood ratio is

$$\frac{\partial f(P_T|e)/\partial e}{f(P_T|e)} = \frac{P'_0(e)}{P_0(e)} \frac{\ln P_T - \mu(e)}{\sigma^2 T} \quad . \quad (47)$$

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

$$\gamma_1 = \alpha \mu_{IC} \frac{P'_0(e)}{P_0(e) \sigma^2 T} , \quad (48)$$

$$\gamma_0 = \alpha \left( \mu_{PC} - \mu_{IC} \frac{P'_0(e) \mu(e)}{P_0(e) \sigma^2 T} \right) = \alpha \mu_{PC} - \gamma_1 \mu(e) , \quad (49)$$

allows us to write:

$$\alpha \left( \mu_{PC} + \mu_{IC} \frac{P'_0(e) \ln P_T - \mu(e)}{P_0(e) \sigma^2 T} \right) = \gamma_0 + \gamma_1 \ln P_T . \quad (50)$$

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 (41), by

$$U'_g \left( w^* \left( \widehat{P} \right) - w^R \right) \left( w^* \left( \widehat{P} \right) - \underline{w} \right) - \left[ \lambda U_l \left( w^R - \underline{w} \right) + U_g \left( w^* \left( \widehat{P} \right) - w^R \right) \right] = 0 . \quad (51)$$

Substituting  $U_g$  and  $U_l$  by their definitions in equations (27) and (28) 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, Kouwenberg and Post, 2004, Optimal Portfolio Choice under Loss Aversion, *The Review of Economics and Statistics*, November, 86(4): 973–987
- [10] Bettis, J. Carr; John M. Bizjak, and Michael L. Lemmon, 2003, The Cost of Employee Stock Options, Working Paper, Arizona State University, (March)
- [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, 2005, Incentive Design Under Loss Aversion, (July)



- [17] Dittmann, Ingolf, and Ernst. Maug, 2006, Lower Salaries and No Options? On the Optimal Structure of Executive Pay, *Journal of Finance*, no. forthcoming (July)
- [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] Rabin, Matthew, 2000, Risk Aversion and Expected-Utility Theory: A Calibration Theorem , *Econometrica* 68, no. 5 (September), pp. 1281-1291
- [55] Rogerson, William P., 1985, The First-Order Approach to Principal-Agent Problems, *Econometrica* 53, no. 6 (November), pp. 1357-1367
- [56] Samuelson, W., and Richard Zeckhauser, 1988, Status quo bias in decision making, *Journal of Risk and Uncertainty*, 1, 7—59.
- [57] Segal, U., and A. Spivak, 1990, First Order versus Second Order Risk Aversion, *Journal of Economic Theory* 51, 111–125.
- [58] Shumway, T., 1997, Explaining returns with loss aversion. Working paper, University of Michigan.
- [59] Thaler, Richard H., 1980, Toward a Positive Theory of Consumer Choice, *Journal of Economic Behavior and Organization*, 1, pp. 39-60.
- [60] ———, 1999, Mental Accounting Matters, *Journal of Behavioral Decision Making* 12, pp. 183-206
- [61] Tversky, Amos, and Daniel Kahneman, 1986, Rational Choice and the Framing of Decisions, *The Journal of Business*, 59(4), 251—278.
- [62] ———, 1991, Loss Aversion in Riskless Choice: A Reference-Dependent Model, *The Quarterly Journal of Economics*, 106, no. 4 (November), pp. 1039-1061
- [63] ———, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, 5, pp. 297-323.
- [64] 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.

<b>Variable</b>		<b>Mean</b>	<b>Std. Dev.</b>	<b>10% Quantile</b>	<b>Median</b>	<b>90% Quantile</b>
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)	$\phi$	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	$\sigma^2$	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 subject to the constraint that option holdings and salaries must be non-negative ( $n_O \geq 0, \phi \geq 0$ ). It shows the mean of the three contract parameters base salary  $\phi^*$ , stock holdings  $n_S^*$  and option holdings  $n_O^*$  together with mean and median of the errors  $error(\phi) = (\phi^* - \phi^d)/\sigma_\phi$ ,  $error(n_S) = (n_S^* - n_S^d)/\sigma_S$ , and  $error(n_O) = (n_O^* - n_O^d)/\sigma_N$ . 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 six different reference wages parameterized by  $\theta$ . Panel B shows the results for the Expected Utility Model for five levels of the risk aversion parameter  $\gamma$ . The last row in each panel shows the corresponding values of the observed contract.

**Panel A: Loss Aversion Model**

$\theta$	Obs.	Avg. Prob. of Loss	Salary ( $\phi$ )			Stock ( $n_S$ )			Options ( $n_O$ )			$D_{Lin}$	
			Mean	Error		Mean	Error		Mean	Error		Mean	Median
				Mean	Median		Mean	Median		Mean	Median		
0.0	912	5.05%	731	-0.5566	-0.3521	0.0234	0.0506	0.0203	0.0111	-0.1907	-0.0931	0.7425	0.4890
0.2	913	26.41%	10018	2.6494	-0.0372	0.0163	-0.0797	0.0013	0.0203	0.4023	-0.0068	3.6283	0.7011
0.4	911	40.75%	22987	7.1156	-0.2885	0.0176	-0.0572	0.0191	0.0197	0.3612	-0.0897	8.4251	0.7625
0.6	910	50.51%	23206	7.1870	-0.4249	0.0212	0.0099	0.0271	0.0146	0.0418	-0.1225	8.6753	0.6986
0.8	908	58.17%	14224	4.0887	-0.4440	0.0236	0.0528	0.0285	0.0109	-0.1992	-0.1318	5.6143	0.6429
1.0	907	64.62%	672	-0.5740	-0.4661	0.0240	0.0607	0.0284	0.0101	-0.2478	-0.1345	0.9744	0.6346
<b>Data</b>	916	N/A	2332	N/A	N/A	0.0208	N/A	N/A	0.0141	N/A	N/A	N/A	N/A

**Panel B: Expected Utility Model**

$\gamma$	Obs.	Salary ( $\phi$ )			Stock ( $n_S$ )			Options ( $n_O$ )			$D_{Lin}$	
		Mean	Error		Mean	Error		Mean	Error		Mean	Median
			Mean	Median		Mean	Median		Mean	Median		
0.2	916	225	-0.7277	-0.4840	0.0245	0.0687	0.0307	0.0094	-0.2987	-0.1388	0.8932	0.6461
0.5	916	244	-0.7210	-0.4806	0.0247	0.0714	0.0315	0.0087	-0.3424	-0.1536	0.9223	0.6668
1.0	915	293	-0.7034	-0.4761	0.0246	0.0695	0.0325	0.0077	-0.4075	-0.1793	0.9599	0.6958
2.0	915	418	-0.6605	-0.4292	0.0241	0.0594	0.0338	0.0060	-0.5185	-0.2506	1.0101	0.7766
3.0	915	578	-0.6046	-0.3851	0.0236	0.0502	0.0290	0.0045	-0.6127	-0.3522	1.0378	0.8103
<b>Data</b>	916	2332	N/A	N/A	0.0208	N/A	N/A	0.0141	N/A	N/A	N/A	N/A

**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  $\gamma$  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 ( $n_o \geq 0, \phi \geq 0$ ). The table shows the average equivalent  $\gamma$  and the mean and median of the distance metric  $D_{Lin}$  for the two models. The table also displays the mean and the median of the difference between the two distance metrics and the frequency that it 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. Results are shown for eleven different reference wages parameterized by  $\theta$ . Some observations are lost because of numerical problems.

$\theta$	Obs.	Average equivalent $\gamma$	$D_{Lin}^{EU}$		$D_{Lin}^{LA}$		$D_{Lin}^{EU} - D_{Lin}^{LA}$			P-values $D_{Lin}^{EU} - D_{Lin}^{LA} = 0$		
			Mean	Median	Mean	Median	Percent $> 0$	Mean	Median	T-test	Wilcoxon	Sign-Test
0.0	912	0.1783	0.9007	0.6525	0.7425	0.4890	55.04%	0.1582	0.0071	0.0000	0.0000	0.0026
0.1	913	0.2353	0.8995	0.6458	1.7211	0.5046	52.90%	-0.8216	0.0056	0.0069	0.0000	0.0852
0.2	913	0.3272	0.9065	0.6519	3.6283	0.7011	45.45%	-2.7218	-0.0001	0.0011	0.0062	0.0066
0.3	911	0.4333	0.9166	0.6639	7.1403	0.7878	50.93%	-6.2237	0.0000	0.0082	0.0069	0.5961
0.4	911	0.5551	0.9391	0.6769	8.4251	0.7625	62.79%	-7.4860	0.0005	0.0269	0.0031	0.0000
0.5	909	0.6821	0.9305	0.6759	7.9026	0.7500	70.41%	-6.9720	0.0012	0.1354	0.0000	0.0000
0.6	908	0.7935	0.9277	0.6774	8.6841	0.6980	75.99%	-7.7564	0.0024	0.2228	0.0000	0.0000
0.7	905	0.8798	0.9209	0.6771	10.1477	0.6653	78.90%	-9.2268	0.0034	0.2730	0.0000	0.0000
0.8	903	0.9256	0.9236	0.6769	5.6303	0.6433	80.07%	-4.7067	0.0040	0.2867	0.0000	0.0000
0.9	903	0.9168	0.9203	0.6735	0.9860	0.6401	80.29%	-0.0657	0.0035	0.2831	0.0000	0.0000
1.0	903	0.8569	0.9169	0.6694	0.9726	0.6338	79.51%	-0.0558	0.0028	0.4334	0.0000	0.0000

**Table 4: Correlations between distance metrics and data**

This table shows the correlations between the distance metric  $D_{Lin}$  on the one hand and the observed stockholdings and the observed option holdings on the other hand. Correlations are shown for the optimal Loss Aversion contract and the equivalent optimal Expected Utility contract where each CEO's risk aversion parameter  $\gamma$  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 ( $n_o \geq 0, \phi \geq 0$ ). The table also shows the average equivalent  $\gamma$ . \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% level, respectively.

$\theta$	Average equivalent $\gamma$	Correlation between $D_{Lin}^{LA}$ and ...		Correlation between $D_{Lin}^{EU}$ and ...	
		$n_S^d$	$n_O^d$	$n_S^d$	$n_O^d$
0.0	0.1783	0.21%	16.84%***	-3.94%	19.13%***
0.1	0.2353	25.09%***	-1.73%	-4.11%	19.95%***
0.2	0.3272	23.58%***	-3.79%	-4.20%	22.09%***
0.3	0.4333	27.89%***	-5.07%	-4.25%	22.87%***
0.4	0.5551	26.16%***	-4.76%	-4.24%	22.98%***
0.5	0.6821	21.84%***	-3.55%	-3.92%	25.43%***
0.6	0.7935	20.56%***	-3.32%	-3.53%	25.03%***
0.7	0.8798	19.59%***	-3.14%	-3.06%	25.91%***
0.8	0.9256	19.18%***	-3.00%	-2.90%	25.67%***
0.9	0.9168	-3.04%	7.32%**	-2.96%	25.11%***
1.0	0.8569	-3.98%	6.73%**	-3.02%	24.71%***

**Table 5: Comparison of Loss Aversion model and Expected Utility model for subsamples**

This table compares the optimal Loss Aversion contract with the equivalent optimal Expected Utility contract where each CEO's risk aversion parameter  $\gamma$  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 ( $n_O \geq 0, \phi \geq 0$ ). The contracts are compared for two subsamples: Panel A displays results for CEOs who own more than 5% of their firm's equity, while Panel B displays the corresponding results for the remaining CEOs in our sample. The table shows the mean of the three contract parameters base salary  $\phi^*$ , stock holdings  $n_S^*$  and option holdings  $n_O^*$ . It also displays the mean between the two distance metrics together with the results of the t-test for zero mean, the median of the difference between the two distance metrics together with the results of the Wilcoxon signed rank test for zero median, and the frequency that this difference is positive together with the sign test for the frequency being 50%. Results are shown for six different reference wages parameterized by  $\theta$ . \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% level, respectively. Some observations are lost because of numerical problems.

**Panel A: Observations for owner executives ( $n_S^d \geq 5\%$ )**

$\theta$	Obs.	Mean Contract Parameters						$D_{Lin}^{EU} - D_{Lin}^{LA}$		
		LA-Model			EU-Model			Percent > 0	Mean	Median
		$\phi$	$n_S$	$n_O$	$\phi$	$n_S$	$n_O$			
0.0	96	462	0.1479	0.0119	332	0.1484	0.0113	18.75%***	0.0497	0.0000***
0.2	96	50638	0.0978	0.0788	341	0.1484	0.0111	12.50%***	-17.6100***	-5.8726***
0.4	96	166830	0.0993	0.0841	353	0.1485	0.0107	26.04%***	-58.0395*	-2.2587***
0.6	96	208173	0.1249	0.0486	391	0.1485	0.0102	42.71%	-71.6593	0.0000
0.8	95	129185	0.1443	0.0178	444	0.1488	0.0099	54.74%	-44.0133	0.0002***
1.0	95	1217	0.1473	0.0127	381	0.1488	0.0101	56.84%	0.0720	0.0016***
Data	96	2127	0.1438	0.0169	2127	0.1438	0.0169	N/A	N/A	N/A

**Panel B: Observations for non-owner executives ( $n_S^d < 5\%$ )**

$\theta$	Obs.	Mean Contract Parameters						$D_{Lin}^{EU} - D_{Lin}^{LA}$		
		LA-Model			EU-Model			Percent > 0	Mean	Median
		$\phi$	$n_S$	$n_O$	$\phi$	$n_S$	$n_O$			
0.0	816	763	0.0087	0.0110	203	0.0097	0.0093	59.31%***	0.1709***	0.0511***
0.2	817	5245	0.0068	0.0135	216	0.0099	0.0089	49.33%	-0.9724	0.0000
0.4	815	6043	0.0079	0.0121	295	0.0099	0.0085	67.12%***	-1.5313*	0.0009***
0.6	812	1392	0.0089	0.0106	243	0.0098	0.0082	79.93%***	-0.2014***	0.0029***
0.8	808	793	0.0093	0.0101	255	0.0098	0.0081	83.04%***	-0.0853	0.0041***
1.0	808	611	0.0094	0.0098	250	0.0098	0.0083	82.18%***	-0.0708	0.0028***
Data	817	2360	0.0062	0.0137	2360	0.0062	0.0137	N/A	N/A	N/A



**Table 6: Comparison of Loss Aversion model and Expected Utility model with and without restrictions**

This table compares the optimal Loss Aversion contract with the equivalent optimal Expected Utility contract where each CEO's risk aversion parameter  $\gamma$  is chosen such that both models predict the same certainty equivalent for the observed contract. Contracts are piecewise linear. Panel A shows the results for the restricted models where option holdings and salaries must be non-negative ( $n_O \geq 0, \phi \geq 0$ ), while Panel B shows the results for the unrestricted models where options and salary can become negative ( $n_O \geq -n_S \exp(r_f T), \phi \geq -W_0$ ). The table shows the average equivalent  $\gamma$  and the frequencies that optimal option holdings are positive, that the optimal salary is positive, and that both (options and salary) are positive. The table also displays mean and median of the difference between the two distance metrics. Behind the mean the result of the t-test for zero mean, and behind the median the result of the sign test for zero median are given. Results are shown for six different reference wages parameterized by  $\theta$ . \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% level, respectively.

**Panel A: Model with restricted salary and restricted option holdings**

$\theta$	Obs.	Average equivalent	Percent with positive Option holdings		Percent with positive fixed salary		Percent with positive options and salary		$D_{Lin}^{EU} - D_{Lin}^{LA}$	
		$\gamma$	EU	LA	EU	LA	EU	LA	Mean	Median
0.0	912	0.1783	82.35%	87.83%	17.76%	63.16%	0.44%	51.21%	0.1582***	0.0071***
0.2	913	0.3272	80.18%	94.19%	19.17%	67.03%	0.11%	61.23%	-2.7218***	-0.0001***
0.4	911	0.5551	78.38%	90.45%	20.53%	44.24%	0.33%	34.69%	-7.4860**	0.0005***
0.6	908	0.7935	77.53%	87.44%	21.59%	29.52%	0.11%	17.07%	-7.7564	0.0024***
0.8	903	0.9256	77.41%	84.72%	22.04%	24.70%	0.11%	9.63%	-4.7067	0.0040***
1.0	903	0.8569	77.96%	83.72%	21.71%	22.70%	0.11%	6.53%	-0.0558	0.0028***

**Panel B: Model with unrestricted salary and unrestricted option holdings**

$\theta$	Obs.	Average equivalent	Percent with positive Option holdings		Percent with positive fixed salary		Percent with positive options and salary		$D_{Lin}^{EU} - D_{Lin}^{LA}$	
		$\gamma$	EU	LA	EU	LA	EU	LA	Mean	Median
0.0	913	0.1783	28.37%	81.27%	1.97%	58.05%	0.44%	51.15%	15.3707***	4.0812***
0.2	908	0.3263	26.32%	78.52%	1.76%	64.10%	0.22%	61.12%	12.9648***	3.0291***
0.4	909	0.5550	23.21%	56.11%	1.87%	36.74%	0.33%	34.43%	6.3083***	0.7955***
0.6	908	0.7928	18.39%	41.30%	1.98%	19.16%	0.11%	16.41%	3.0577*	0.1588***
0.8	902	0.9242	15.63%	33.81%	2.44%	11.86%	0.11%	9.09%	2.9660***	0.1173***
1.0	902	0.8570	17.29%	31.37%	2.33%	7.32%	0.00%	5.32%	2.3199*	0.0734***

**Table 7: Optimal nonlinear contracts**

This table compares the optimal Loss Aversion contract with the equivalent optimal Expected Utility contract where each CEO's risk aversion parameter  $\gamma$  is chosen such that both models predict the same certainty equivalent for the observed contract. Contracts are estimates of the respective general nonlinear contract and subject to the constraint that the smallest possible wage must be positive ( $\underline{w} = 0$ ). The table shows the average slope of the wage function below the observed strike price  $\Delta_{Low}$ , the average slope of the wage function above the observed strike price  $\Delta_{High}$ , and the frequency with which  $\Delta_{High} > \Delta_{Low}$ . In addition, the table shows the average dismissal probability which is the probability with which the contract pays the minimum wage  $\underline{w}$ , and the mean inflection quantile, which is the quantile at which the curvature of the optimal wage function changes from convex to concave. Panel A displays the results for the Loss Aversion Model for six different reference wages parameterized by  $\theta$ . Panel B shows the results for the Expected Utility Model, and here the table also reports the equivalent  $\gamma$ . For the Loss Aversion Model, Panel A also shows the incentives from dismissals that are generated by the drop to the minimum wage  $\underline{w}$ . For the Expected Utility Model this quantity cannot be calculated as the contract falls continuously to the minimum wage  $\underline{w}$ . Some observations are lost because of numerical problems.

**Panel A: Loss Aversion Model**

$\underline{w}$	$\theta$	Obs.	Mean $\Delta_{Low}$	Mean $\Delta_{High}$	Percent $\Delta_{High} > \Delta_{Low}$	Mean Dismissal Probability	Incentives from Dismissals	Mean Inflection Quantile
0	0.0	909	2.81%	3.43%	90.10%	1.12%	0.40%	84.92%
0	0.2	904	1.68%	3.20%	98.45%	5.43%	4.43%	97.42%
0	0.4	896	0.95%	2.90%	98.88%	11.68%	13.23%	98.68%
0	0.6	882	0.56%	2.49%	99.09%	18.32%	25.35%	99.05%
0	0.8	878	0.38%	2.04%	99.32%	24.96%	38.77%	99.18%
0	1.0	862	0.32%	1.61%	99.30%	30.54%	50.50%	99.29%

**Panel B: Expected Utility Model**

$\underline{w}$	$\theta$	Obs.	Average equivalent $\gamma$	Mean $\Delta_{Low}$	Mean $\Delta_{High}$	Percent $\Delta_{High} > \Delta_{Low}$	Mean Dismissal Probability	Mean Inflection Quantile
0	0.0	650	0.1799	3.60%	3.56%	65.38%	11.40%	33.00%
0	0.2	787	0.3300	3.34%	3.49%	64.55%	18.13%	30.10%
0	0.4	801	0.5642	3.79%	3.23%	50.44%	18.71%	22.47%
0	0.6	800	0.8085	4.17%	2.88%	37.50%	18.66%	19.83%
0	0.8	791	0.9358	4.31%	2.70%	35.40%	18.95%	19.85%
0	1.0	794	0.8592	4.17%	2.80%	40.05%	19.16%	20.36%

**Table 8: Comparison of general nonlinear Loss Aversion  
and Expected Utility models**

This table compares the optimal Loss Aversion contract with the equivalent optimal Expected Utility contract where each CEO's risk aversion parameter  $\gamma$  is chosen such that both models predict the same certainty equivalent for the observed contract. Contracts are estimates of the respective general nonlinear contract and subject to the constraint that the smallest possible wage must be positive ( $w = 0$ ). The table displays mean and median of the difference between the two models according to three distance metrics:  $D_{NonLin}$ ,  $D_{Level}$ , and  $D_{Slope}$ . Behind the means the result of the t-test for zero mean, and behind the medians the result of the sign test for zero median are given. Results are shown for six different reference wages parameterized by  $\theta$ . \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% level, respectively. Some observations are lost because of numerical problems.

$w$	$\theta$	Obs.	$D_{NonLin}^{EU} - D_{NonLin}^{LA}$		$D_{Level}^{EU} - D_{Level}^{LA}$		$D_{Slope}^{EU} - D_{Slope}^{LA}$	
			Mean	Median	Mean	Median	Mean	Median
0	0.0	648	0.1444***	0.0431***	0.14235**	0.0184***	0.00693	0.00002***
0	0.2	782	0.0931**	0.0384***	0.58233*	-0.0310***	0.00794	0.00002***
0	0.4	787	0.0416	0.0254***	-0.55179	-0.8790***	0.04981	0.00001***
0	0.6	774	0.0257	0.0177***	-1.59592***	-1.6421***	0.08389	0.00000**
0	0.8	766	-0.0047	0.0096***	-1.86386***	-1.5970***	0.05422	0.00000**
0	1.0	757	-0.0596	-0.0117***	-1.88794***	-1.3015***	0.04337	0.00000**

**Table 9: Comparison of linear and nonlinear Loss Aversion models**

This table compares the optimal linear Loss Aversion contract with the optimal nonlinear Loss Aversion contract. The nonlinear model assumes that the minimum wage is equal to  $\underline{w}$  (which is either zero or  $-W_0$ ). The corresponding linear model is subject to the same restriction for the fixed salary,  $\phi > \underline{w}$ , while options can become negative ( $n_O \geq -n_S \exp(r_f T)$ ). For both models, the table shows the average slope of the wage function below the observed strike price,  $n_S$  and  $\Delta_{Low}$ , respectively, the average slope of the wage function above the observed strike price,  $n_S + n_O$  and  $\Delta_{High}$ , 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 six different reference wages parameterized by  $\theta$  and for two different levels of the minimum wage  $\underline{w}$ . Some observations are lost because of numerical problems.

$\underline{w}$	$\theta$	Obs.	Linear Option Contract				General Nonlinear contract			
			Mean $n_S$	Mean $n_S + n_O$	Mean Savings	Mean $D_{NonLin}$	Mean $\Delta_{Low}$	Mean $\Delta_{High}$	Mean Savings	Mean $D_{NonLin}$
0	0.0	909	0.0238	0.0335	0.0014	0.0901	0.02815	0.03425	0.0029	0.1593
0	0.2	904	0.0166	0.0361	0.0049	0.1899	0.01677	0.03202	0.0201	0.2226
0	0.4	896	0.0191	0.0369	0.0082	0.2054	0.00949	0.02895	0.0411	0.3516
0	0.6	882	0.0228	0.0355	0.0104	0.1642	0.00561	0.02489	0.0582	0.4381
0	0.8	878	0.0255	0.0342	0.0122	0.1277	0.00383	0.02041	0.0732	0.4975
0	1.0	861	0.0264	0.0336	0.0134	0.1184	0.00320	0.01614	0.0867	0.5384
$-W_0$	0.0	913	0.0277	0.0335	0.0017	0.1446	0.03093	0.03509	0.0047	0.1743
$-W_0$	0.2	912	0.0264	0.0352	0.0102	0.3629	0.01884	0.03292	0.0298	0.2040
$-W_0$	0.4	904	0.0380	0.0336	0.0237	0.5663	0.01091	0.02955	0.0614	0.3282
$-W_0$	0.6	900	0.0504	0.0300	0.0371	0.6875	0.00660	0.02512	0.0916	0.4265
$-W_0$	0.8	866	0.0550	0.0236	0.0482	0.8402	0.00319	0.01740	0.1188	0.5291
$-W_0$	1.0	867	0.0586	0.0253	0.0549	0.7342	0.00368	0.01562	0.1408	0.5443

**Table 10: Comparative statics for the parameters of the value function**

This table describes the piecewise linear optimal contract subject to the constraint that option holdings and salaries must be non-negative ( $n_O \geq 0$ ,  $\phi \geq 0$ ) for different values of the parameters  $\alpha$ ,  $\beta$ , and  $\lambda$  of the value function. The table shows the average slope of the wage function below the observed strike price  $n_S$ , the average slope of the wage function above the observed strike price  $n_S + n_O$ , 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. If not otherwise stated in the table, the remaining parameters are set at their base value:  $\alpha = 0.88$ ,  $\beta = 0.88$ ,  $\lambda = 2.25$ ,  $\theta = 0$ . Some observations are lost because of numerical problems.

Parameter	Value	Obs.	Linear Option Contract			
			Mean	Mean	Mean	Mean
			$n_S$	$n_S + n_O$	Savings	$D_{NonLin}$
$\lambda$	1.00	912	0.0242	0.0338	0.0018	0.0694
$\lambda$	1.50	912	0.0237	0.0340	0.0013	0.0624
$\lambda$	2.00	912	0.0234	0.0343	0.0012	0.0618
$\lambda$	2.25	912	0.0234	0.0345	0.0012	0.0622
$\lambda$	2.50	912	0.0233	0.0346	0.0012	0.0626
$\lambda$	3.00	912	0.0232	0.0347	0.0012	0.0636
$\lambda$	4.00	912	0.0231	0.0350	0.0013	0.0649
$\alpha$	0.60	909	0.0248	0.0328	0.0210	0.0864
$\alpha$	0.70	910	0.0248	0.0331	0.0114	0.0830
$\alpha$	0.80	912	0.0245	0.0335	0.0046	0.0773
$\alpha$	0.88	912	0.0234	0.0345	0.0012	0.0622
$\alpha$	0.95	912	0.0223	0.0354	0.0004	0.0557
$\beta$	0.60	912	0.0230	0.0349	0.0014	0.0613
$\beta$	0.70	912	0.0231	0.0348	0.0013	0.0607
$\beta$	0.80	912	0.0232	0.0346	0.0012	0.0610
$\beta$	0.88	912	0.0234	0.0345	0.0012	0.0622
$\beta$	0.95	912	0.0235	0.0343	0.0013	0.0635