Sticks or Carrots? Optimal CEO Compensation when Managers are Loss-Averse

Finance Working Paper N°. 193/2007

November 2007

Ingolf Dittmann
Erasmus University Rotterdam

Ernst Maug
University of Mannheim and ECGI

Oliver Spalt
University of Mannheim

ECGI Working Paper Series in Finance

Sticks or Carrots? Optimal CEO Compensation when Managers are Loss-Averse

Working Paper N°. 193/2007
November 2007

Ingolf Dittmann
Ernst Maug
Oliver Spalt

We are grateful to seminar participants at the University of Cologne, Frankfurt, Georgia State, Humboldt, Mannheim, Maryland, Tilburg, the JFI-Conference on "Financial Contracting", the 6th Oxford Finance Symposium, the European Finance Association meeting in Ljubljana, and to Axel Börsch-Supan, Gerard Hoberg, Matjaz Koman, Christian Laux, and David De Meza for their feedback. We also thank the collaborative research centers SFB 649 on "Economic Risk" in Berlin and the SFB 504 "Rationality Concepts, Decision Making and Economic Modeling" for financial support. Ingolf Dittmann acknowledges financial support from NWO through a VIDI grant.

Sticks or Carrots?
Optimal CEO Compensation when Managers are Loss-Averse*

Ingolf Dittmann† Ernst Maug‡ Oliver Spalt§

December 14, 2007

Abstract

This paper analyzes optimal executive compensation contracts when managers are loss averse. We establish the general optimal contract analytically and calibrate the model to the observed contracts of 595 CEOs. We find that the Loss Aversion-model explains the observed structure of executive compensation contracts significantly better than the Risk Aversion-model. This holds especially for the mix of stock and options. The Loss Aversion-model predicts convex contracts with substantial option holdings that provide a stronger upside ("carrots"). By contrast, the optimal contract is concave for the standard Risk Aversion-model where it provides a significant downside ("sticks"). Our results suggest that loss aversion is a better paradigm for analyzing design features of stock options and for developing preference-based valuation models than the conventional model used in the literature.

JEL Classifications: G30, M52

Keywords: Stock Options, Executive Compensation, Loss Aversion

*We are grateful to seminar participants at the University of Cologne, Frankfurt, Georgia State, Humboldt, Mannheim, Maryland, Tilburg, the DGF-conference in Dresden, the GEABA-conference in Tübingen, the JFI-Conference on "Financial Contracting", the 6th Oxford Finance Symposium, the European Finance Association meeting in Ljubljana, and to Bo Becker, Axel Börsch-Supan, Xavier Gabaix, Gerard Hoberg, Andreas Knabe, Matjaz Koman, Roy Kouwenberg, David Larcker, Christian Laux, David De Meza, and Werner Neus for their feedback. We also thank the collaborative research centers SFB 649 on "Economic Risk" in Berlin and the SFB 504 "Rationality Concepts, Decision Making and Economic Modeling" for financial support. Ingolf Dittmann acknowledges financial support from NWO through a VIDI grant.

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

‡Corresponding author. University of Mannheim, D-68131 Mannheim, Germany. Email: maug@bwl.uni-mannheim.de, Tel: +49 621 181 1952.

§University of Mannheim, D-68131 Mannheim, Germany. Email: spalt@bwl.uni-mannheim.de, Tel: +49 621 181 1973.

Sticks or Carrots?
Optimal CEO Compensation 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 calibrate the model to the observed contracts of 595 CEOs. We find that the Loss Aversion-model explains the observed structure of executive compensation contracts significantly better than the Risk Aversion-model. This holds especially for the mix between stock and options. The Loss Aversion-model predicts convex contracts with substantial option holdings that provide a strong upside ("carrots"). By contrast, optimal contracts are concave for the standard Risk Aversion-model where they feature a significant downside ("sticks"). Our results suggest that loss aversion is a better paradigm for analyzing the design features of stock options and for developing preference-based valuation models than the conventional model used in the literature.

JEL Classifications: G30, M52

Keywords: Stock Options, Executive Compensation, Loss Aversion

1 Introduction

In this paper we explain salient features of observed compensation contracts with a simple contracting model where the manager is loss averse. We parameterize this model using standard assumptions and then compare the contracts generated by the model with those actually observed for a large sample of U.S. CEOs. Our main conclusion is that a principal agent-model with loss-averse agents can approximate observed contracts far better than the standard model based on risk aversion used in the literature. In particular, the Loss Aversion-model can explain the prevalence of stock options, a feature that is inconsistent with the standard Risk Aversion-model.

The theoretical literature on executive compensation contracts is largely based on contracting models where shareholders (principal) are risk-neutral and where the manager (agent) is risk averse, which is modeled with a concave utility function. Some highly stylized models can explain option-type features, but quantitative approaches rely more or less entirely on a standard model with constant relative risk aversion, lognormally distributed stock prices, and effort aversion.[1] However, Hall and Murphy (2002) and Dittmann and Maug (2007) show that the standard CRRA-lognormal model cannot explain observed compensation practice if companies and managers can bargain over all components of CEO compensation packages.[2] Dittmann and Maug find that the optimal predicted contract almost never contains any options and typically features negative base salaries. These results raise a concern for the widespread application of the model to the valuation of executive stock options and to the analysis of their design (strike price, indexing, reloading, and repricing).[3]

In this paper we suggest a different approach to explaining the almost universal presence of stock options by assuming that managers' preferences exhibit loss aversion as described by Kahneman and Tversky (1979) and Tversky and Kahneman (1991, 1992). On the basis of experimental evidence they argue that choices under risk exhibit three features: (i) reference dependence, where agents do not value their final wealth levels, but evaluate outcomes relative to some benchmark or reference level; (ii) loss aversion, which adds the notion that losses (measured relative to the reference level)

[1] A model that can explain the use of options is Feltham and Wu (2001) who assume that the effort of the agent affects the risk of the firm, and Oyer (2004), who models options as a device to retain employees when recontracting is expensive. Inderst and Müller (2005) explain options as instruments that provide outside shareholders with better liquidation incentives. In Oyer (2004) and Inderst and Müller (2005), options do not provide incentives to exert effort. The applications by Haubrich (1994), Haubrich and Popova (1998), and by Margiotta and Miller (2000) use constant absolute risk aversion when calibrating a principal-agent model. Calibration exercises with CRRA preferences and lognormal distributed stock prices include Lambert, Larcker, and Verrecchia (1991), Hall and Murphy (2000, 2002), Hall and Knox (2002), and Lambert and Larcker (2004).

[2] Hall and Murphy (2002) establish this for the case with adjustable base salaries, where the optimal strike price of stock options becomes zero. Then the optimal contract features only restricted stock but no options.

[3] 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 Dybvig and Loewenstein (2003) on reloading.

loom larger than gains; (iii) diminishing sensitivity, so that individuals become progressively less sensitive to incremental gains and incremental losses. These assumptions accord well 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.[4]

The main drawback of risk aversion-approaches in explaining the prevalent use of stock options in compensation contracts is the fact that risk averse managers gain little utility from payoffs when the value of the firm is high.[5] 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 the same level of incentives at a lower cost. The Risk Aversion-model therefore predicts contracts with much higher stock holdings combined with zero or even negative salaries and option holdings. However, these predictions are at odds with observed compensation practice, where managers are paid with options, have guaranteed base salaries and entitlements to severance payments, which protect them even in case of dismissal. By comparison, loss aversion implies that managers are more averse to losses than they are attracted by gains, so they demand a premium for being exposed to losses and value the downside protection provided by options. Shareholders will therefore offer a contract that pays at least the reference wage most of the time in order to avoid paying this premium. The Loss Aversion-model therefore suggests contracts that reward good outcomes rather than penalize bad outcomes and combine positive option holdings with positive fixed salaries.

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

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

[5] This assessment relies on the standard implementation of the Risk Aversion-model, see Footnote 1 for the relevant literature. Other authors have pursued larger deviations from the Risk Aversion–model, which can accomodate options in a stylized setup, see for example Hemmer, Kim, and Verrecchia (1999) and the approaches by Feltham and Wu (2001) and Oyer (2004) cited above.

In the second step of our analysis we parameterize both models using assumptions that are based on available compensation data and on prior research, especially experimental evidence on preference-parameter values. Then we calibrate the models for 595 CEOs for whom we have complete data. We first restrict contracts to be piecewise-linear and represent them as consisting of base salary, stock, and stock options. We compute the optimal contract for each CEO for the Loss Aversion-model and for the Risk Aversion-model for a range of plausible parameterizations and assess how well each model predicts the observed contract. We consider two specifications of the Risk Aversion-model – the constant absolute risk aversion model and the constant relative risk aversion model – as these cover virtually the entire literature on compensation.

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

We also drop the simplifying assumption that the contract is piecewise-linear and calculate the optimal nonlinear contracts for each CEO in our sample. This approach allows us to perform a robustness check on our stylized representation of contracts. Above some threshold level, the general nonlinear contracts are mostly convex, and at the threshold level they feature a discontinuous drop to the lowest feasible wage, which is reminiscent of a dismissal of the CEO. For plausible parameterizations of the Loss Aversion-model we estimate that shareholders would save an additional 0.4% to 4.6% of current compensation costs if they would replace the optimal piecewise linear contract with the optimal nonlinear contract, including the discontinuous drop below a critical stock price. We therefore suggest that the governance costs of incentive provision through CEO dismissals (with big drops in compensation, i.e. without severance pay) rather than through high-powered wage functions is probably not worth the additional costs for most companies. The ability to quantify these effects based on data is the strength of our approach, which calibrates the model to each individual CEO.

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. Gomes (2005) and Berkelaar, Kouwenberg, and Post (2004) apply the model to portfolio choice. Barberis and Huang (2001) and Barberis, Huang and Santos (2001) apply loss aversion to the explanation of the value premium. Haigh and List (2005) find that CBOT-traders are loss averse, and more so than inexperienced students, contradicting the effect List (2004) found earlier for consumers. Coval and Shumway (2005) support the same conclusion in their study of intraday risk-taking of CBOT-traders. Kouwenberg and Ziemba (2004) study the incentives and investment decisions of hedge-fund managers, and Ljungqvist and Wilhelm (2005) base their measure of issuer satisfaction in initial public offerings on loss aversion. The only application that fails to support loss aversion to the best of our knowledge is Massa and Simonov (2005) in their study of individual investor behavior. Despite the usefulness of loss aversion to analyze risk taking incentives in many areas of finance, the only paper so far that rigorously applies loss aversion to principal-agent theory is de Meza and Webb (2007). However, they do not apply their argument to executive compensation contracts and explore a different specification from ours. To the best of our knowledge, ours is the first paper that explores empirically the potential of loss aversion to explain observed compensation contracts.

In the following Section 2 we develop the model and discuss the main assumptions. In Section 3 we characterize the optimal contract analytically. Section 4 develops our empirical methodology in detail. Section 5 analyzes contracts that consist of fixed salaries, stock, and options. Section 6 extends this analysis to general nonlinear contracts. Section 7 documents the robustness of our approach. Section 8 concludes. All proofs and derivations are deferred to the appendix.

2 The Model

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

Contracts and technology. The contract is a wage function $w(P_T)$ that specifies the wage of the manager for a given realization of the company value P_T at time T. Contract negotiations take place at time 0. At the end of the contracting period, T, the value of the firm P_T is commonly observed and the wage is paid according to $w(P_T)$. P_T depends on the CEO's effort e and the state of nature.

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

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

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

where $C(e)$ is an increasing and convex cost function. The assumption of additive separability in effort and income is conventional in the literature, and our strategy is to follow conventions in the literature for all aspects other than the modeling of preferences.[6] For this we assume preferences over wage income, $w(P_T)$, of the form[7]

$$V\left(w\left(P_T\right)\right) = \begin{cases} \left(w\left(P_T\right) - w^R\right)^{\alpha} & if \quad w\left(P_T\right) \geq w^R \\ -\lambda\left(w^R - w\left(P_T\right)\right)^{\beta} & if \quad w\left(P_T\right) < w^R \end{cases}, \text{ where } 0 < \alpha, \beta < 1 \text{ and } \lambda \geq 1. \tag{2}$$

Here, w^R denotes the reference wage. If the payoff of the contract at time T exceeds the reference wage, then the manager codes this as a gain, whereas a payoff lower than w^R is coded as a loss. We will refer to the range of the wage above w^R as the *gain space* and to the range below w^R as the *loss space*. There are three aspects that set this specification apart from standard concave utility specifications. First, the parameter $\lambda > 1$ 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.[8] Formally, this introduces a kink in the value function at w^R and thus locally infinite risk-aversion.[9] Second, the manager treats her income from the firm separately from income from other sources, a phenomenon that is often referred to as "framing" or "mental accounting" (Thaler, 1999). 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 the LA-model to the Risk Aversion-model (RA-model).

The standard implementation in the literature on executive compensation features preferences

[6] Edmans, Gabaix, and Landier (2007) argue for multiplicative preferences, which makes an important difference for their calibrations of the optimal level of incentives.

[7] 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 Langer and Weber (2001), Berkelaar, Kouwenberg, and Post (2004), and Barberis and Huang (2005).

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

[9] This characteristic is also called 'first-order risk aversion' (Segal and Spivak, 1990).

with constant relative risk aversion, but some papers also use constant absolute risk aversion:

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

$$V^{CARA}\left(w\left(P_T\right)\right) = -\exp\left(-\rho\left(W_0 + w\left(P_T\right)\right)\right) \ , \tag{4}$$

where W_0 denotes wealth, γ represents the coefficient of relative risk aversion and ρ the coefficient of absolute risk aversion. Our theoretical analysis focuses on the LA-model only as the RA-model has been analyzed in many places in the literature (see Footnote 1 in the Introduction). In the empirical part we calibrate both models to the data.

We assume that the reference point w^R is exogenous in two respects. First, the reference point does not depend on any of the parameters of the contract. Alternative assumptions would relate the reference point to the median or the mean payoff of the contract $w\left(P_T\right)$, which would increase the mathematical complexity of the argument substantially. De Meza and Webb (2007) focus on this aspect of applying loss aversion to principal-agent theory. Second, the reference point is also independent of the level of effort. This is defensible if the cost of effort is non-pecuniary and if the manager separates the costs of effort from the pecuniary wage. However, this is potentially a strong assumption if the costs are pecuniary and the manager frames the problem so that she feels a loss if her payoff does not exceed w^R plus any additional expenses for exerting effort. In the second case, $C\left(e\right)$ should simply be added to the reference point w^R. We do not pursue this route here for mathematical tractability. With an exogenous reference point the distinguishing feature of the Loss Aversion-model is that the attitude to risk is not a global property but is different for wage distributions centered around the reference point compared to distributions where most of the probability mass is far away from the reference point.

The manager has some outside employment opportunity that provides her with a value net of effort costs $\underline{V}$, so any feasible contract must satisfy the ex ante participation constraint $E\left[V\left(w\left(P_T\right)\right)\right] - C\left(e\right) \geq \underline{V}$. We assume that the principal cannot pay a wage below some lower bound $\underline{w}$ on the wage function such that $\underline{w} \leq w\left(P_T\right)$ for all P_T, where $\underline{w} < w^R$. If the manager would be required to invest all her private wealth in the securities of the firm, then her total payoff cannot fall below $-W_0$ in any state of the world, and this would happen only if these securities expired worthless at the end of the period. This makes $\underline{w} = -W_0$ a natural choice, but higher values of $\underline{w}$ may also be plausible.

3 Analysis

3.1 Discrete effort

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

$$\min_{w(P_T)\geq \underline{w}} \int w(P_T) f(P_T|\overline{e}) dP_T \tag{5}$$

$$s.t. \int V(w(P_T)) f(P_T|\overline{e}) dP_T \geq \underline{V} + C(\overline{e}) \quad , \tag{6}$$

$$\int V(w(P_T)) \Delta f(P_T|e) dP_T \geq \Delta C \ , \tag{7}$$

where $\Delta C = C(\overline{e}) - C(\underline{e})$. We denote the Lagrange multiplier on the participation constraint (6) by μ_{PC} and the Lagrange multiplier on the incentive compatibility constraint (7) by μ_{IC} and can now characterize the optimal contract.

Proposition 1. ***(Optimal contract):*** *Given the preference structure in (1) and (2) and assuming that the monotone likelihood ratio property holds for $f(P_T|e)$ the optimal contract $w^*(P_T)$ for the principal agent problem (5) to (7), is:*

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

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

The details of the proof of Proposition 1 and an implicit definition of $\widehat{P}$ are deferred to Appendix A. The proof involves three steps. The first step shows that the optimal contract can never pay off in the interior of the loss space, so $w^*(P_T)$ cannot lie strictly between $\underline{w}$ and w^R. The reason is that the agent is risk loving in the loss space, so any payment in the loss space can be improved upon by replacing it with a lottery between the lowest possible wage $\underline{w}$ and a payoff for some wage $w \geq w^R$ in the gain space. The second step shows that such lotteries are not optimal. Instead, incentives are improved if the contract always pays $\underline{w}$ if the stock price falls below some critical value $\widehat{P}$, and pays off in the gain space otherwise. The third step derives the Lagrangian and maximizes it pointwise with respect to $w(P_T)$. Equation (8) shows that for the gain space, where $P_T > \widehat{P}$, we

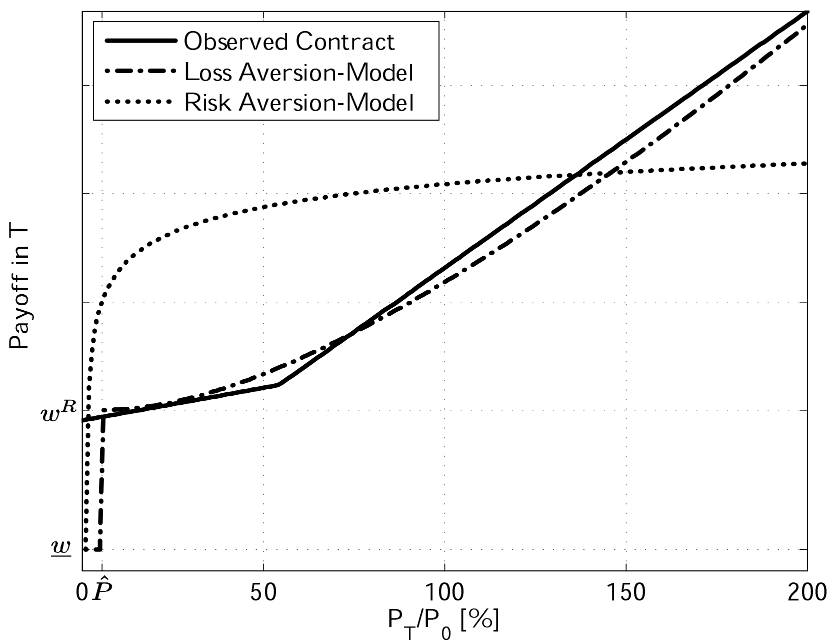

Figure 1: The Figure plots the Loss-Aversion-contract (equation (8)), the Risk-Aversion-contract, and the observed contract for a representative CEO in our sample, assuming a lognormal distribution of the terminal stock price P_T.

obtain a result very similar to the familiar Holmström condition (Holmström, 1979, equation (7)) 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.

Proposition 1 provides us with a general characterization of the optimal contract with a loss-averse manager. Figure 1 illustrates this contract for a typical parameterization and contrasts it with the corresponding RA-contract. For some region $P_T > \widehat{P}$ the optimal contract is continuous, monotonically increasing, and pays off only in the gain space. For $P_T \leq \widehat{P}$ the optimal contract pays off the lowest possible wage $\underline{w}$. The contract features a discontinuity at $\widehat{P}$ where the manager's wage jumps discretely from $\underline{w}$ to some value $w^*(P_T) \geq w^R > \underline{w}$.[10]

Under the assumption that stock prices are lognormal the LA-contract is convex above $\widehat{P}$ but has an inflection point above which it becomes concave. The RA-contract is always concave for a coefficient of relative risk aversion greater than 1. Hence, the optimal LA-contract (8) provides the

[10] De Meza and Webb (2007) find a similar discontinuity in a principal agent model with loss aversion. In their specification, however, the payoff jumps from $\underline{w}$ to w^R and is flat at w^R before it possibly increases continuously. A flat payout at the reference wage w^R occurs if the slope of the line that connects $(0, \underline{w})$ and $(\widehat{P}, w^R)$ is steeper than the slope of the utility function entering the gain space. With the Kahneman and Tversky (1992) value function (2), this cannot occur because the slope entering the gain space is infinite, so that the agent prefers a fair gamble over $\underline{w}$ and $w^R + \varepsilon$ to w^R for ε sufficiently small.

manager with significant downside protection, punishments for extreme declines in the stock price, and increasing marginal rewards as the share price increases. By contrast, the RA-contract provides high-powered incentives for low and intermediate stock prices and decreasing marginal rewards as the share price increases. These qualitative features drive our empirical results for the general nonlinear contracts as well as for the piecewise linear contracts that can be implemented with stock and options.

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 discrete case, we have to apply the first-order approach, i.e., we replace the agent's incentive compatibility constraint (7) (more precisely, its analogue for continuous effort) with the first order condition for (7). 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 condition uniquely identifies the maximum of her objective function.[11] In our case, this requires that

$$\frac{\partial^2 E\left(V\left(w\left(P\right)\right)|e\right)}{\partial e^2} = \int V\left(w\left(P_T\right)\right)\frac{\partial^2 f\left(P_T\,|e\right)}{\partial e^2}dP_T - \frac{\partial^2 C\left(e\right)}{\partial e^2} < 0\ . \tag{9}$$

This condition will not hold generally. In our setting, one issue is the convexity of the function $V\left(P_T\right)$ over the loss space. Moreover, the optimal contract $w\left(P_T\right)$ may be convex over some regions of the gain space. However, we can ensure that condition (9) holds for some cost functions C and some density functions in two ways. Firstly, equation (9) shows that this condition will be satisfied for sufficiently convex cost functions, so that $\partial^2 C\left(e\right)/\partial e^2$ is bounded from below such that (9) holds. Secondly, if the production function $P_T\left(e\right)$ is sufficiently concave (such that $\partial^2 P_T\left(e\right)/\partial e^2$ is sufficiently small for all effort levels), then (9) will also be satisfied. In the remainder of this paper we will assume that equation (9) 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 (9) holds for each effort level. Then, the results from Proposition 1 continue to hold when the likelihood ratio for the discrete case, $\Delta f\left(P_T\,|e\right)/f\left(P_T\,|\bar{e}\right)$, is replaced by its continuous equivalent, $f_e(P_T|e)/f(P_T|e)$.*

[11] The literature on the principal-agent model has identified conditions where this "first-order approach" is valid in a risk aversion framework. See, for example, Jewitt (1988) and Rogerson (1985).

4 Implementation and data

4.1 Implementation

The general Loss Aversion-contract. In our empirical implementation, we assume that the stock price is lognormally distributed:[12]

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

where r_f is the risk-free rate of interest, σ^2 the variance of the returns on the stock, T the time horizon, 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 $P_0 = E[P_T \exp\{-r_f T\}]$.[13] 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 observed 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 (5) to (7) can then be written as:

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

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

$$\alpha\left(w^R - \underline{w}\right) = \left(\gamma_0 + \gamma_1 \ln \widehat{P}\right)\lambda\left(w^R - \underline{w}\right)^{\beta} + (1-\alpha)\left(\gamma_0 + \gamma_1 \ln \widehat{P}\right)^{\frac{1}{1-\alpha}}. \tag{12}$$

Hence, we can represent the nonlinear LA-contract by the coefficients γ_0 and γ_1 and write it as $\mathcal{C}^{LA} = \{\gamma_0, \gamma_1\}$. This specification implies that the contract predicted by the model is strictly increasing in P_T and that it is convex as long as $P_T \leq \exp\{\alpha/(1-\alpha) - \gamma_0/\gamma_1\}$. Above this value $w^*(P_T)$ is concave. It is therefore an empirical question whether the contract described in equation (11) can explain option contracts, because the concave region may or may not be empirically relevant.

All parameters of the model given by equations (11) and (12) except γ_0 and γ_1 can be determined from standard data sources and from experimental results in the literature (see Section 4.2).

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

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

We determine the remaining two parameters γ_0 and γ_1 numerically as described in the following paragraph.

Finding optimal contracts. Our null hypothesis is that the observed contract $w^d(P_T)$ is an optimal contract. Here and in the following we use the superscript 'd' in order to refer to observed values or 'data.' Since $w^d(P_T)$ is optimal under the null, it can be rationalized as the outcome of an optimization program, where we assume that preferences are parameterized as in (1) and that the technology is parameterized as in (10). (The program is specified in equations (46) to (48) in Appendix A.) If $w^d(P_T)$ is indeed optimal, then it should not be possible to find another contract that (i) provides the same incentives as the observed contract, (ii) provides the same utility to the CEO as the observed contract, and (iii) costs less to shareholders compared to the observed contract. We therefore determine the contract parameters by solving the following program numerically:

$$\min_{\mathcal{C}} \ \pi\left(w(P_T\,|\mathcal{C})\right) \equiv \int w(P_T\,|\mathcal{C})f(P_T)dP_T \tag{13}$$

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

$$\int V\left(w(P_T\,|\mathcal{C})\right)\frac{\partial f(P_T)}{\partial P_0}dP_T \geq \int V(w^d\left(P_T\right))\frac{\partial f(P_T)}{\partial P_0}dP_T \quad . \tag{15}$$

This program uses a slightly more general notation as we write the wage function as $w(P_T\,|\mathcal{C})$, where $\mathcal{C}$ can refer to different types of contracts. For the time being, we only consider $\mathcal{C} = \mathcal{C}^{LA} = \{\gamma_0, \gamma_1\}$.[14] By writing P_T as in (10) 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 focus only on the first stage by solving program (13) to (15) as it does not depend on knowledge of the cost function $C(e)$ or of the production function $P_0(e)$. We therefore do not consider the second stage. This implies also that we cannot analyze the optimal *level* of incentives (pay for performance sensitivity) for a compensation contract, which would invariably depend on

[14] The optimal contract (11) is completely determined by the two parameters γ_0 and γ_1. As the constraints (14) and (15) always bind in the optimum, these constraints uniquely define the optimal contract, and no further optimization is necessary. Hence, the optimal general contract can be calculated with a system of two equations (14) and (15) in two unknowns γ_0 and γ_1. The piecewise linear contract (16) has three parameters, so for this contract we solve the complete problem (13) to (15).

this information. However, with our approach we can analyze the optimal *structure* of compensation contracts for any given level of incentives.

Proposition 1 provides only necessary but not sufficient conditions. We therefore solve the optimization problem for different starting values in order to find the global optimum.[15] For none of the CEOs in our sample and none of the parameter constellations considered did we find any indication that there is more than one local optimum.

Program (13) to (15) generates a new contract $w^* (P_T)$ that is less costly to shareholders. Condition (15) ensures that the CEO has at least the same incentives under the new contract as she had under the observed contract, so that the contract found by the program will not result in a reduced level of effort (assuming the validity of condition (9)). Similarly, condition (14) 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 $w^d (P_T)$ with the optimal contract $w^* (P_T)$ generated by program (13) to (15).

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

$$w^d (P_T) = \phi^d e^{r_f T} + n_S^d P_T + n_O^d \max (P_T - K, 0) \ , \tag{16}$$

where K is the strike price of the option. We abstract from other details of observed contracts and consolidate each CEO's portfolio of options into one representative option (see Section 4.2 for details). The main reason is that different option grants have different maturities and can therefore not be modeled within the standard one-period principal agent model. We comment on the restrictions imposed by this simplification in the conclusion.

Since the observed contract is piecewise linear and expressed as a tuple of the fixed salary ϕ, the number of shares n_S, and the number of options n_O, we also calculate optimal contracts that are restricted to be piecewise linear. We will compute the piecewise linear LA-contract as the solution to program (13) to (15) and denote this contract by $\mathcal{C}_{Lin}^{LA} = \{\phi^{LA}, n_S^{LA}, n_O^{LA}\}$. Here the strike price K and the maturity T of the option grant are set equal to the strike price and maturity of the representative option that is estimated from the data. We also compare the LA-model with the RA-model and

[15] We calculate all our results for three different starting values and have experimented with six different starting values for a subset of our dataset.

calculate optimal piecewise linear RA-contracts, which we denote by $\mathcal{C}_{Lin}^{RA} = \{\phi^{RA}, n_S^{RA}, n_O^{RA}\}$. This is the solution to program (13) to (15) with $\mathcal{C} = \mathcal{C}_{Lin}^{RA}$, where the wage function is again piecewise linear as in (16) and preferences are given by (3) or (4).

Comparing model contracts with observed contracts. We use metrics that measure the average distance between optimal contracts and observed contracts. We want to analyze to what extent the model predicts the observed composition of the contract between stock and options, so we define the metric D_{Lin} as:

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

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

Here summation is over all N CEOs in the sample. Arithmetic means over all CEOs are denoted by a bar. This metric measures the distance between the observed contract and the model contract and gives more weight to those parameters that have lower cross-sectional dispersion. D_{Lin} does not take into account fixed salaries, because these may be determined by considerations outside the model, in particular the CEO's bargaining power. In our formalization of the game shareholders have all the bargaining power, but this assumption does not affect the shape of the optimal contract.[16] If the CEO had some or all of the bargaining power then the shape of the optimal contract would still be dictated by optimal risk sharing considerations and the CEO would extract a bargaining rent through a higher base salary. For these reasons the accurate prediction of base salaries is a less important feature of the model than the prediction of the mix of stock and options. Still, we want to investigate to what extent both models predict base salaries correctly and therefore define a second metric D_{LinS} analogous to D_{Lin}, where D_{LinS} also includes the squared deviations of the base salary, $error(\phi) = \frac{\phi_i^* - \phi_i^d}{\sigma_\phi}$, where σ_ϕ is the cross-sectional standard deviation of base salaries in the sample.

A similar approach to ours was used in Carpenter (1998) and Bettis, Bizjak, and Lemmon (2005).[17] To check the robustness of this approach, we experimented with alternative metrics obtained by different weighting schemes and different approaches to scaling the squared or absolute

[16] This statement is strictly true for preferences with constant absolute risk aversion. With constant relative risk aversion, bargaining power affects CEOs' wealth, and thereby their attitude to risk as well as the shape of the contract. This effect is ignored in the discussion above.

[17] The main difference between their approach and ours is that we calibrate our model to individual observations, whereas they calibrate their models to sample averages.

differences between model parameters and observed parameters. We found that all plausible approaches yield qualitatively similar results. This is not surprising because the incentive compatibility constraint (15) and the participation constraint (14) ensure that deviations from the observed value for one parameter result in deviations for the other two parameters as well. For example, an increase in the number of options increases incentives and therefore generates a lower number of shares. Hence, large deviations for one parameter result in similarly large deviations for the other parameter (or one of the other two parameters in case of D_{LinS}), so that the scaling and weighting of any single parameter relative to the other is largely inconsequential.

4.2 Data

We identify all CEOs in the ExecuComp database who are CEO for the entire fiscal years 2004 and 2005. We also delete all CEOs who where executives in more than one company in either 2004 or 2005 and separately estimate CEOs' contracts in 2004 and in 2005. The 2004 contracts are only needed to construct the reference wage for 2005. We set P_0 equal to the market capitalization at the end of 2004 and take the dividend rate d, the stock price volatility σ, and the proportion of shares owned by the CEO n_S^d from the 2004 data, while the fixed salary ϕ^d is calculated from 2005 data.[18]

Option portfolios. 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 as the estimated option portfolio. We take into account the fact that most CEOs exercise their stock options before maturity by multiplying the maturity of the individual options in the estimated portfolio by 0.7 before calculating the representative option (see also Huddart and Lang, 1996, and Carpenter, 1998). The maturity T determines the contracting period and the risk-free rate r_f is the U.S. government bond rate from January 2005 with maturity closest to T.

Minimum wage. For the minimum wage we rely on the argument above that the CEO's wage cannot drop below $-W_0$. Such a contract requires that the CEO invests all her non-firm wealth in securities of her firm. There is anecdotal evidence that newly hired executives are asked to invest

[18] This reflects the fact that stock and options are stock variables measured at the end of the period whereas base salary is a flow measured during the period. ϕ is the sum of the following four ExecuComp data types: Salary, Bonus, Other Annual, and All Other Total. We do not include LTIP (long-term incentive pay), as these are typically not awarded annually.

some of their private wealth into their new company. In our base case, we therefore set the minimum wage $\underline{w}$ equal to $-W_0$. We argue that we should not exclude contracts with negative payouts just because we rarely observe them. Instead, a good model should *generate* contracts with non-negative payouts. Nevertheless, we also repeat our analysis with the minimum wage set equal to zero, an assumption that is more commonly made in the literature.

Wealth. We need an estimate of the CEO's non-firm wealth to evaluate relative risk aversion for the RA-model and the lower bound $\underline{w}$ on the wage function for both models. We estimate the portion of each CEO's wealth that is not tied up in securities of his or her company from historical data. 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. In order to obtain meaningful wealth estimates, we delete all CEOs with less than five years history as executive of any firm in the database. After deleting 4 CEOs of firms with stock volatility exceeding 250%, our data set contains 595 CEOs.

[Insert Table 1 here]

Table 1A provides descriptive statistics for all variables in our data set. The median CEO receives a fixed salary of \$1.7m, owns 0.3% of the firm's equity and has options on another 1% of the firm's equity. The median firm value is \$2.3bn and the median moneyness K/P_0 is 0.7, so most options are clearly in the money. The median maturity is 4.4 years. The distributions of the contract parameters are highly skewed, so their means are substantially larger than their medians. We also provide the same data for 576 CEOs in 1997 in order to show that the observed parameters are broadly similar to those observed in 2005. Apart from lower firm values and lower fixed salaries we see that volatility was lower, moneyness higher and option maturities somewhat longer compared to the 2005 data set. Option holdings have almost doubled over the interval from 1997 to 2005. We conduct our analysis for the 2005 dataset and provide the key results also for the 1997 dataset as a robustness check.[19]

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

[19] Option pay did not vanish in recent years after the burst of the new-economy bubble and recent changes in accounting rules. New option grants amounted to 21.6% of total CEO compensation in 1992 and increased steadily during the 1990s with a peak at 42.3% in 2001. They came down since then and accounted for 21.2% in 2005, effectively returning to their 1992 level. These numbers are calculated from the ExecuComp database and not shown in the tables.

and assume that the reference wage reflects expectations the CEO forms based on her previous year's (i.e. 2004) compensation package. It seems natural that the CEO regards a total compensation (fixed and variable) below the fixed salary of the previous year as a loss and we use this as a lower bound. In addition, she may also build in some part of her deferred compensation into her reference wage. Most likely, she will evaluate her securities at a substantial discount relative to their value for a well-diversified investor. This discount depends on her attitude to risk and on her framing of the wage-setting process. We therefore regard the market value of her existing 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.[20] We denote the market value of her deferred compensation in 2005 based on the number of shares and options she held in 2004 by MV and write:

$$w^R_{2005}(\theta) = \phi_{2004} + \theta \cdot MV(n^S_{2004}, n^O_{2004}, P_{2005}) \ , \tag{18}$$

The parameter θ is an index of the discount the CEO applies to her deferred compensation. If $\theta = 0$, then the reference wage for 2005 equals her base salary for 2004. If $\theta = 1$, then the reference wage equals the market value of her total compensation in the previous year, valued at current market prices and without a discount for risk. We will look at a grid of alternative values for θ. The distribution of deferred compensation is highly skewed. If CEOs set their reference points based on the median of the distribution, then CEOs' reference points will be below the market value, i.e. $\theta < 1$.

Preference parameters. For the preference parameters α, β, and λ we rely on the experimental literature for guidance. We therefore use $\alpha = \beta = 0.88$ and $\lambda = 2.25$ as our baseline values.[21]

5 Contracts with restricted stock and options

We now describe the piecewise linear contracts predicted by the LA-model and compare them to the contracts predicted by the standard RA-model. Minimization of program (13) to (15) is subject to two additional constraints: First, option awards can become negative (i.e. managers can be required to write options), but the manager's short position in options cannot exceed her stock

[20] DeMeza and Webb (2007) develop a related argument why this discount may be substantial.

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

holdings $n_S \exp(dT)$, so $n_O > -n_S \exp(dT)$.[22] This restricts the wage function to be non-decreasing. Similarly, we assume limited liability, so the base salary is limited by the manager's non-firm wealth ($\phi > -W_0$). For each CEO, we compare the observed contract with the optimal piecewise linear contract for the LA-model and for the RA-model.

[Insert Table 2 here]

Table 2 Panel A summarizes the results for the LA-Model for seven different levels of the reference wage as parameterized by θ (see equation (18)). Panel B shows the results for the RA-model for seven values of the coefficient of relative risk-aversion γ.[23] Panel C shows the same results for the RA-model for constant absolute risk aversion preferences (equation (4)), where the coefficient of absolute risk aversion is chosen so that relative risk aversion corresponds to the values in Panel B.[24] For each model we show the means and medians of the contract parameters predicted by the models and the scaled mean deviations of these predicted parameters from their observed counterparts (referred to as errors in equation (17)).

Both parameterizations of the RA-model predict negative base salaries and negative option holdings, so optimal RA-contracts are concave, confirming what we expected based on the theoretical analysis above (see also Figure 1). Both versions of the RA-model predict larger stock holdings, although the scaled deviations are smaller here because the cross-sectional standard deviations of stockholdings is 5.2% and therefore almost four times as large as the standard deviation of option holdings, which is 1.4% (see Table 1). Given the similarity of the two parameterizations of the RA-model, we will focus on one model from now on. The CRRA-model performs better than the CARA-model in terms of the metric D_{LinS} for all levels of risk aversion, and also better in terms of D_{Lin} for lower levels of risk aversion. We want to make sure not to bias our analysis in favor of the LA-model and therefore focus our analysis and all comparisons on the CRRA-version from now on.

[22] If the dividend yield $d = 0$, then this constraint becomes $n_O > -n_S$. We abstract from dividends in our theoretical analysis, but we do consider them in our empirical work.

[23] We do not consider values of γ below 0.1 in Table 2 as they lead to numerical problems. When the manager is risk-neutral, then the optimal contract is indeterminate and the numerical problems for low values of γ reflect this indeterminacy. The literature on executive compensation has often discussed values for γ in the range between 2 and 3. Hall and Murphy (2000) use these values that seem to go back to Lambert, Larcker, and Verrecchia (1991). Lambert and Larcker (2004) more recently proposed a value as low as 0.5. A useful point of reference here is the portfolio behavior of the CEO, since very low levels of risk aversion (below 1) imply that CEOs have implausibly highly leveraged investments in the stock market. Ingersoll (2006) develops a parameterization of the RA-model that is sufficiently similar to ours but includes investments in the stock market. Using his equation (8) and assuming a risk premium on the stock market as low as 4% and a standard deviation of the market return of 20% gives an investment in the stock market (including exposure to the stock market through holding securities in his own firm) equal to $1/\gamma$. E.g., $\gamma = 0.1$, the lowest value considered in Table 2, would imply that the CEO invests ten times her wealth in the stock market. We do not wish to take a restrictive stance in order not to bias our analysis in favor of the LA-model and therefore allow for levels of risk aversion as low as 0.1, even though we regard such values as highly implausible.

[24] The coefficient of absolute risk aversion ρ is calculated from γ as: $\rho = \gamma/(W_0 + \pi_0)$, where π_0 is the market value of the manager's contract (i.e., the costs of the contract to the firm).

The performance of the LA-model is very sensitive to the assumed reference wage. For lower values of the reference wage ($\theta = 0$ to $\theta = 0.2$) the LA-model predicts values for all contract parameters that are broadly consistent with the data. The scaled deviations are below 0.5 in absolute value for $\theta = 0.1$ and $\theta = 0.2$ for all three contract parameters. While the option holdings are smaller than observed, the predicted magnitudes are similar to the observed magnitudes and median option holdings are positive for all values of the reference wage up to and including $\theta = 0.4$. Overall, the LA-model performs well as long as we assume that managers have reference points that are closer to their fixed salaries (which, in our simplification, includes bonus payments) than to the market value of their total compensation.

The fit of the LA-model deteriorates markedly for high values of the reference point ($\theta > 0.4$). It then becomes similar to both parameterizations of the RA-model and predicts negative median option holdings and negative median base salaries, with scaled deviations in excess of 1 in absolute value. The reason is that options can limit losses only if the reference wage w^R is sufficiently low. Both models feature higher base salaries if incentives are provided with options, and lower base salaries if incentives are provided with shares because shares are worth more to the manager than options for the same level of incentives, and the participation constraint then requires that base salaries are adjusted accordingly. With a low reference wage, option compensation together with a high base salary ensures that total compensation almost never falls below the reference wage. However, with a high reference wage this is not the case and then the manager incurs large losses when the options expire out of the money, and then incentive provision through shares becomes optimal.

For very low reference wages any feasible contract will only pay off in the gain space and the loss space becomes irrelevant. As the manager is slightly risk-averse in the gain space, the optimal contract would then contain no options and only stock (assuming this is feasible) just as in the RA-model with a low value of γ. This is the reason why the LA-model predicts the largest option holdings for $\theta = 0.1$, where it is also most accurate.

We can illustrate this point with the help of Figure 1 above. Effectively, the piecewise linear contract attempts to approximate the general nonlinear contract as well as possible. As we increase the reference wage, the discontinuity of the general nonlinear contract moves to the right, i.e., $\hat{P}$ becomes larger and moves more towards the center of the distribution. This is reflected in the average probability of loss, $\Pr(w(P_T) \leq w^R)$, in Table 2 Panel A. The optimal nonlinear contract is locally concave at $\hat{P}$: It jumps discretely and then has a very small, positive slope. If the jump at $\widehat{P}$ is in the center of the distribution, this local concavity is important and the best approximation with a piecewise linear contract is achieved through a concave contract with negative option holdings.

We further analyze the relationship between the optimal general contract and the optimal piecewise linear contract in Section 6 below.

These qualitative observations are also reflected in the metric D_{Lin} computed from (17). Its median is above one in absolute value for both versions of the RA-model and for all parameterizations of the LA-model with high reference points ($\theta \geq 0.6$). This confirms our conclusion that the LA-model works well for low reference wages. It achieves the optimum at $\theta = 0.1$, whereas the RA-model works best if risk aversion is either very low or very high. The lowest distances between observed contracts and RA-model contracts occur for the highest levels of risk aversion. Recall that the RA-model always replaces all options with shares. High risk aversion reduces the incentives from options more than those from stock, so optimal contracts feature fewer additional shares to replace the existing options compared to lower levels of risk aversion. The accuracy of the model is therefore higher for higher levels of risk aversion.

The RA-model also becomes slightly more accurate if risk-aversion decreases and converges to zero. This reflects the fact that *any* observed contract is optimal (i.e. cost minimizing) if the agent is risk-neutral ($\gamma = 0$), because subjective values are then identical to market values and all contracts that generate the same incentives are equally costly. The values for the metric D_{LinS}, which also considers base salaries, are larger than those for D_{Lin} by construction and the qualitative results are similar to those for D_{Lin}.

An important limitation of the analysis in Table 2 is the fact that it confounds two aspects of our problem. First, we analyze and compare different approaches to modeling attitudes to risk. Second, we also vary the overall attitude to risk as we change the reference wage, respectively, the degree of relative risk aversion. It therefore does not seem warranted to compare *all* parameterizations of the LA-model with *all* parameterizations of the RA-model. Instead, it is more sensible to compare the two models based on *comparable* parameterizations that hold the overall attitude to risk constant in a meaningful way. Then we can be sure that differences between the models do not reflect implicit differences in the overall attitude to risk. We therefore compare parameterizations that generate the same valuation of the observed contract by the same CEO. We define the certainty equivalent of model M, CE^M, from $E\left(V^M\left(w^d\left(P_T\right)\right)\right) = V(CE^M)$. We fix θ to determine the reference wage of each CEO and then define an equivalent degree of relative risk aversion γ_e from

$$CE^{LA}(w^d, \theta) \equiv CE^{RA}(w^d, \gamma_e) \; . \tag{19}$$

We refer to the value of γ_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 paid by shareholders, $E(w^d) - C(w^d)$, constant for both models. For each CEO and for each θ we calculate the equivalent γ_e and the optimal RA-contract with $\gamma = \gamma_e$. Table 3 compares the two models.

[Insert Table 3 here]

Table 3, Panel A reports the mean and the median difference $D_{Lin}^{RA} - D_{Lin}^{LA}$ of the distance metric D_{Lin} between the two models (as of now the RA-model refers to the CRRA-preferences). The verdict based on the mean and median of D_{Lin} as well as that based on the median of D_{LinS} is clear and independent of the overall attitude to risk: The LA-model dominates the RA-model for the entire range of reference wages. The distribution of D_{LinS} is skewed, so we sometimes obtain different indications for means and medians. Note that the mean of $D_{LinS}^{RA} - D_{LinS}^{LA}$ is never significantly different from zero when it is negative, so the RA-model never dominates the LA-model for any parameterization and any test. However, the RA-model fits the data better than the LA-model according to D_{LinS} for a small number of observations (3% - 23% of the sample), some of which generate extreme deviations for the LA-model. We investigate this in more detail in Table 10 below and show that the large deviations occur primarily for owner-CEOs who own a large fraction of their companies.

The equivalent γ_e's are generally very low and below the range we regard as plausible (see Footnote 23). They are also non-monotonic in θ: As the reference wage increases or decreases far enough, the kink of the value function moves into the tails of the payoff distribution for the CEO, so that overall risk aversion (which is captured by γ_e) becomes smaller.

Table 3, Panel B reports how successful the two models are in explaining the two stylized facts that fixed salaries and option holdings are almost always positive for observed CEO pay contracts. The LA-model predicts positive option holdings for 91% of the sample for $\theta = 0.1$, the value that also yields the best approximation overall. Moreover, the LA-model predicts positive salaries for the majority of all CEOs when $\theta \leq 0.2$ and then it also predicts simultaneously positive option holdings and positive base salaries. By contrast, the number of cases where the RA-model predicts simultaneously positive option holdings and positive salaries is virtually zero. The model reduces options and exchanges them for more stock and lower salaries until either the restriction on salaries ($\phi \geq -W_0$) or the restriction on option holdings ($n_O \geq -n_S \exp(dT)$) binds. This model can therefore never explain positive option holdings and positive salaries simultaneously, while more than

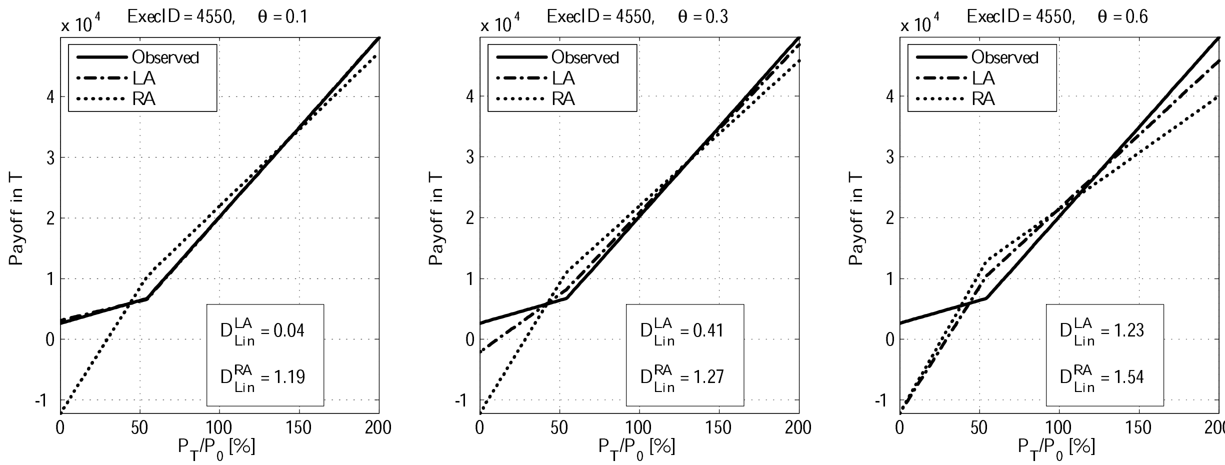

Figure 2: The figure shows the observed contract, the LA-contract and the RA-contract for the CEO with ExecID # 4550 for $\theta = 0.1$, $\theta = 0.3$, and $\theta = 0.6$. The horizontal axis shows the terminal stock price P_T as a percentage of the current stock price P_0. The vertical axis displays the total payoff $w(P_T)$ for each type of contract.

99% of the CEOs in our sample have such a contract. Altogether, the LA-model can generate the qualitative characteristics of observed contracts for the majority of the CEOs in our sample, provided we parameterize the model appropriately. The RA-model is clearly inferior on this dimension.

Figure 2 illustrates the results from Tables 2 and 3 for the case of a typical CEO and provides also a visual impression of the distance metric and the corresponding observed and predicted wage functions. The figure shows the optimal LA-contract, the optimal RA-contract and the observed contract for the same CEO for $\theta = 0.1$, $\theta = 0.3$, and for $\theta = 0.6$. For $\theta = 0.1$, the LA-contract and the observed contract are visually indistinguishable with a value of $D^{LA}_{Lin} = 0.04$ for the distance metric. By contrast, the corresponding RA-contract is concave and differs substantially from the observed contract, which is reflected in a higher value of the distance metric of $D^{RA}_{Lin} = 1.19$. For $\theta = 0.3$, the LA-model predicts a convex contract with positive option holdings, but with a negative base salary. Here the LA-model still performs much better than the RA-model. For $\theta = 0.6$ both models perform poorly, but the deterioration is somewhat stronger for the LA-model than it is for the RA-model, even though the LA-model still dominates.

Finally, we observe that our results on optimal contracts rely entirely on risk-sharing considerations. In particular, shareholders' objective to reduce the CEO's rents never plays a role in our analysis. Both, the theoretical contract (11) and the observed contract may provide the agent with a positive rent [25]. Any rent the agent receives in the observed contract is preserved in our calibrations,

[25] Our preference specification (2) implies that the agent's lowest possible utility is bounded away from minus infinity, so rents cannot be precluded (see Proposition 2 in Grossman and Hart, 1983). In the observed contract, rents could

because the participation constraint in our numerical work (14) ensures that the agent's utility from the optimal contract is never lower than her utility from the observed contract. Empirically, constraint (14) is always binding in our sample, so the optimal contract provides the agent with exactly the same rent as the observed contract.

6 General non-linear Loss-Aversion contracts

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

One feature of the optimal nonlinear contract in the LA-model is the discrete jump at the point $\hat{P}$ from $\underline{w}$ to some number above w^R. This jump can be interpreted as a dismissal of the manager, and we will also use the word "dismissal" in the tables for brevity. In practice, however, dismissals do not always generate a sharp drop in the payoff function, for example when managers receive sufficient severance pay to compensate them for their loss of compensation. We do not have data on severance pay and we therefore abstract from this aspect.

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

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

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

Here Δ_{Low} is the average slope in the region below the strike price of the option, which can be compared to the number of shares n_S. Δ_{High} is the average slope in the region above the strike price and can be compared to shares and options combined.

We are also interested in the convexity and the concavity of the optimal contracts and we analyze this in two ways. First, we ask if the slope in the high range of terminal stock prices, Δ_{High}, exceeds the slope in the lower range, Δ_{Low}. This would correspond to positive option holdings. Second, from (11) 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 concave above this point. We use the probability that the predicted contract pays off in the convex range, $\Pr (w^* (P_T) \leq P_I)$ as another descriptive statistic.[26]

additionally be caused by rigid salaries (i.e. liquidity constraints) or managerial power.

[26] There are some CEOs where $P_I \leq \hat{P}$, so the LA-contract has a slope of zero up to the discontinuity and then

Finally, we define the dismissal probability p of the optimal model contract as

$$p(\hat{P}) \equiv \int_0^{\hat{P}} f(P_T)\, dP_T. \tag{22}$$

We have no reliable method to evaluate individual dismissal probabilities for CEOs. We estimate the average probability of dismissal by calculating the frequency with which CEOs in the ExecuComp database leave the company within a given four-year period, where the recorded reason is 'resigned.' We repeat this for all four-year periods between 1995 and 2004 and obtain an average dismissal probability of 7.4%. This number is inferred from a cross-section and the *ex ante* probabilities may well vary across CEOs. However, we have no reliable way of modeling this heterogeneity here, so we can only compare the mean generated by the model with the mean in the data.

[Insert Table 4 here]

Table 4 reports the average slopes Δ_{Low} and Δ_{High}, the dismissal probability, and the quantile of the inflection point for different parameterizations. We also report the percentage of those CEOs where $\Delta_{High} > \Delta_{Low}$. The contracts predicted by the LA-model are mostly convex by both measures of convexity. The slope in the upper range, Δ_{High} is almost always higher than the slope in the lower range, Δ_{Low}. Similarly, almost all of the probability mass for this contract lies to the left of the inflection point, rendering the concave part of the contract irrelevant.

The dismissal probabilities are unrealistically high for the LA-model once the reference point becomes sufficiently high ($\theta-$values above 0.5). This aspect underlines our earlier assessment that high parameterizations with high reference wages lead to poor performance of the LA-model. As the reference wage increases, the threat of dismissals becomes more important. Intuitively, CEOs with a higher reference wage demand a higher compensation, and they receive it in the sense that their compensation while they are employed is larger. However, then incentives are provided to a lesser extent through the slope of the wage function (note how Δ_{Low} and Δ_{High} both tend to decline as w^R increases) and to a larger extent through the threat of dismissals (column seven in Table 4).

In principle, the optimal nonlinear contract (11) could be approximated with a sufficiently large number of options with different strike prices, where option holdings are negative for some strike prices to approximate the discrete jump and the concave part of the wage function for very high wages. In practice however, we do not observe contracts with negative option holdings. This raises the question how costly it is to restrict the contract shape to being piecewise linear, i.e. implementable

becomes concave. For these CEOs we calculate $\Pr\left(w^*(P_T) \leq \hat{P}\right)$.

by fixed salary, stock and one option grant. In Table 5 we therefore compare the optimal non-linear contract (11) with the optimal piecewise linear contract. For both contracts, the table shows the average slopes Δ_{Low} and Δ_{High} and the distance metric D_{NonLin}, which parallels our definition of D_{Lin}:[27]

$$D_{NonLin} = \left[\left(\frac{\Delta^*_{Low} - \Delta^d_{Low}}{\sigma_{Low}} \right)^2 + \left(\frac{\Delta^*_{High} - \Delta^d_{High}}{\sigma_{High}} \right)^2 \right]^{1/2} \tag{23}$$

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

Here, $\Delta^{d,i}_{Low}$ and $\Delta^{d,i}_{High}$ represent the slopes of the observed contract corresponding to (20) and (21) and $\bar{\Delta}^d_{Low}$ and $\bar{\Delta}^d_{Low}$ denote their sample averages. In addition, Table 5 shows how much shareholders could save (as a proportion of total observed compensation) if they could recontract and replace the observed contract with the contract predicted by the models. These savings from recontracting are defined as

$$Savings = \frac{E\left(w^d\left(P_T\right)\right) - E\left(w^*\left(P_T\right)\right)}{E\left(w^d\left(P_T\right)\right)} \ , \tag{24}$$

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

[Insert Table 5 here]

Table 5 shows that the accuracy (i.e. the negative of the average D_{NonLin}) of the general contract is higher than the accuracy of the piecewise linear contract except for $\theta = 0$. For low reference wages the difference is small, but it increases as the reference wage increases. By construction, the savings relative to the status-quo of the optimal general contract are higher than the savings of the piecewise linear contract.

The savings are not substantial for either version of the contract. This is important, because it shows that even where the distance between the observed contracts and the predicted contracts appears large in terms of the metrics developed above, the savings are insubstantial, particularly for the piecewise linear contract. The difference in savings between the piecewise linear contract and the general nonlinear contract is small: It is 0.4% for $\theta = 0$ of total compensation costs and 4.6% for $\theta = 0.4$, or \$1.37 million for the median CEO with a pay package worth \$29.8 million. This is about

[27] Note that for the piecewise linear contract, $\Delta_{Low} = n_S \exp(dT)$ and $\Delta_{High} = n_S \exp(dT) + n_O$.

0.06% of the value of the median company. These savings have to be related to the costs of writing and enforcing such a general contract. We conclude that the benefits of incentive provision through CEO dismissals with big drops in compensation rather than through high-powered wage functions is negligible for most companies.

7 Robustness checks

The measurement of wealth. The measurement of non-firm wealth cumulates the CEO's past income and adjusts for purchases and sales of securities. The actual wealth may be higher than this (e.g., if the CEO has saved income earned before she enters the database) or lower (e.g., if the savings rate was less than 100% and some income was consumed). We therefore check the robustness of our results for measurement errors in CEO wealth.

[Insert Table 6 here]

Table 6 reports the main results of Table 3 if we reduce the estimate of wealth by 50% (Panel A) and if we increase it by 100% (Panel B). The results are qualitatively very similar to those reported in Table 3 for the base case. The mean and median difference of $D_{Lin}^{RA} - D_{Lin}^{LA}$ is significantly positive for all levels of the reference wage. The results are more pronounced compared to those in Table 3 if wealth is higher and somewhat less strong but statistically still highly significant if wealth is lower. This is so because the values for the average equivalent γ are increased if wealth is higher and reduced if wealth is lower. In the CRRA-model absolute risk aversion is lower if wealth is higher, so the equivalent γ must be higher with higher wealth, and we observe already in the discussion of Table 2 that very low levels of risk aversion improve the performance of the RA-model. The mean and median differences of D_{LinS} exhibit the same patterns as in the base case and the median difference favors the LA-model in all cases. Also, the percentage of CEOs for whom the fixed salary as well as option holdings are positive is hardly affected by the changes in wealth considered. Overall, none of our results seems to be affected by measurement errors of CEO wealth.

Restrictions on the wage function. Our analysis of the base case allows for negative salaries and option holdings. However, many previous authors have imposed tighter restrictions and we therefore repeat our analysis and require that salary and option holdings cannot become negative, i.e. $\phi \geq 0$ and $n_O \geq 0$.

[Insert Table 7 here]

Table 7 reports the results for the model with tighter restrictions and has the same structure as Table 3. Comparison of the restricted model with the unrestricted base case from Panel B of Tables 3 shows that the restrictions have a much stronger impact on the RA-model than they have on the LA-model. This is not surprising given that the LA-model already generates non-negative base salaries and option holdings in most cases, so that tightening the constraints has no impact. However, the RA-model is still not able to generate positive salaries and positive option holdings simultaneously. One of the two new constraints always binds: Either option holdings are equal to zero, then salary is positive, or the predicted salary is zero and option holdings are positive. The median values of D_{Lin} and D_{LinS} in Table 7, Panel A show that for most CEOs the LA-model still dominates the RA-model for almost all values of the reference wage w^R. In terms of the means of $D_{Lin}^{RA} - D_{Lin}^{LA}$ the accuracy of the RA-model increases and is higher than the accuracy of the LA-model on average in most cases. This shows that ruling out concave contracts and negative base salaries improves the performance of the RA-model significantly and for a minority of cases the RA-model now dominates. We conclude that the RA-model is only able to generate positive salaries or positive option holdings if we impose this as a restriction on the maximization problem, but even with these assumptions the LA-model still dominates the RA-model for the typical CEO.

Data from 1997. The data for the 2005 cross-section of CEOs on ExecuComp may be special. As a robustness check we repeat our analysis for 1997 (see Table 1, Panel B for descriptive statistics on these CEOs).

[Insert Table 8 here]

Table 8 shows the results for 1997, which are very similar to those for the 2005 sample in Table 3. The percentage of CEOs where $D_{Lin}^{RA} > D_{Lin}^{LA}$ and $D_{LinS}^{RA} > D_{LinS}^{LA}$ depend less on the reference wage than they do for the 2005 sample. Both models are now better at predicting positive option holdings and positive fixed salaries, but the RA-model still cannot predict both contract features simultaneously, while the results for the LA-model are better for the 1997 dataset than they are for the 2005 dataset in this respect. The 1997 data therefore lead to very similar conclusions and, if anything, strengthen the case for the LA-model.

Preference parameters. We check to what extent our results are sensitive to our assumptions on the preference parameters. We have based our discussion on the estimates of α, β, and λ from the experimental literature. These estimates might be inappropriate for the study of CEOs, so we check the robustness of our results with respect to different values for the preference parameters.

[Insert Table 9 here]

Table 9 reports the results of a comparative static analysis in terms of the preference parameters where the reference wage w^R is set to last year's fixed salary plus 10% of the risk-neutral value of last year's stock and option holdings (i.e. $\theta = 0.1$). We report only the results for the piecewise linear model. From the metric D_{Lin} we can see that the LA-model performs better if we increase the loss aversion-parameter λ, whereas the performance of the model deteriorates for increases in the curvature of the value function, i.e., for reductions in α and β. Increases in α and β make the value function locally risk-neutral, so this result is similar to the improvement with convergence to risk neutrality noted earlier. For high α-values and β-values the attitude to risk depends then only on the degree of loss aversion λ, but unlike risk aversion loss aversion is a local property of the value function in the neighborhood of the reference point. The results of Table 9 therefore show that it is this local property that is responsible for the better performance of the LA-model, which improves further if this aspect is emphasized (higher λ, α and β). Conversely, for a lower degree of loss aversion and stronger curvature of the value function (lower λ, α and β) the value function becomes more similar to that of the standard CRRA-model with $\gamma = 1 - \alpha$ in the gain space, where more than 95% of the probability mass lies for the base scenario in Table 9 (Table 2 Panel A). The performance of the LA-model deteriorates accordingly and becomes more similar to that of the RA-model.

Owners versus managers. As a last robustness check we try to identify those observations where the LA-model performs consistently poorly. We split the sample into a subsample with the 54 owner-executives who own 5% or more of the shares of their firm and a subsample with the remaining 541 CEOs who own less than 5% of their firm. Table 10 displays the results for the two subsamples; it provides a breakdown of the results shown in Table 3.

[Insert Table 10 here]

For both subsamples the LA-model performs better than the RA-model and the median results are not strongly affected. However, for the metric D_{LinS}, which also accounts for prediction errors of the base salary, the average difference $D_{LinS}^{RA} - D_{LinS}^{LA}$ becomes negative and very large in magnitude in Panel A, for a large range of values of the reference wage. Closer inspection of the data shows that these results are driven by those owner-manager CEOs who have no options (one example in our dataset is Warren Buffett). We conclude from this that the LA-model should not be applied to these CEOs. Their relationship to the firm cannot be described by a principal-agent relationship as they are not salaried agents of outside shareholders.[28]

8 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 on several dimensions compared to the Risk Aversion-model.

- Contracts predicted by the Loss Aversion-model are much closer to observed contracts than contracts predicted by the Risk Aversion-model.
- The Loss Aversion-model predicts positive option holdings in line with observed contracts for most CEOs, whereas the Risk Aversion-model always predicts concave contracts with negative option holdings.
- The Loss Aversion-model predicts positive base salaries, whereas the Risk Aversion-model implies that the majority of CEOs should invest some of their private wealth in their firms without receiving a base salary.

Our results are of particular importance to the substantial literature on the design and the valuation of executive stock options that relies on variants of the Risk Aversion-model (see Footnote 3 in the Introduction). Our analysis suggests that for these applications the Loss Aversion-model is more relevant than the Risk Aversion-model. Our analysis also gives some guidance regarding relevant ranges of the reference wage: Predicted contracts most closely resemble observed contracts for relatively low reference wages that are set close to the previous fixed salary.

Our analysis relies on stylized contracts that abstract from a number of features of observed contracts. The simplest and probably most innocuous assumption restricts the number of option grants to one. Multiple strike prices would allow for a better approximation of the piecewise linear contract to the optimal nonlinear contract, and we have shown that the benefits from such a better approximation are small. We also ignore pension commitments, the use of perks, and loans the

[28] The agency problem in these companies is more likely that between the inside blockholder and minority shareholders, and this problem cannot be captured by a model based on effort aversion.

corporation extends to its officers, largely because we do not have data on these items. These compensation items are not related to stock price performance, so they only bias our estimate of fixed compensation downward. The Risk Aversion-model predicts lower levels of fixed compensation compared to the Loss Aversion-model, so such a downward bias in estimating fixed compensation biases our results against the Loss Aversion-model. Finally, we ignore severance provisions, again for lack of data, but our discussion in Section 6 suggests that our analysis can potentially help to explain the widespread use of severance arrangements. If we assume that the Loss-Aversion-model is correct, then the benefits from threatening the CEO with dismissal and an associated drop in compensation are small and probably outweighed by the costs of a governance structure that could enforce such a contract.

While our results demonstrate that the Loss Aversion-model is better at explaining the structure of observed CEO compensation contracts than the Risk Aversion-model, it is still subject to important limitations. The most crucial aspect of both models may be the fact that they are both static, whereas shareholders and CEOs typically revise their contracts repeatedly over a number of periods. Research in contract theory shows that in such a context the surplus may be appropriated by the agent even when the principal has all of the bargaining power (Ray, 2002). Then the contractual structure may serve to allocate the surplus of the contractual relationship between the CEO and shareholders over time, an aspect that is absent from static models. Exploration of these aspects of the structure of compensation contracts is left for future research.

A Proofs

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

For Lemma 1, we extend the set of permissible contracts to contracts that pay out lotteries. The agent is risk-seeking over losses, so lotteries might be part of the optimal contract. Lemma 2 shows, however, that the optimal contract does not contain lotteries.

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

Proof of Lemma 1: Consider first the contract $w(P_T)$ that pays off $\underline{w} < w(P_T) < w^R$ at some price P_T with certainty. Since the value function in the loss space, $-\lambda\left(w^R - w(P_T)\right)^\beta$, is monotonically increasing in $w(P_T)$, there exists a unique number $g(P_T) \in (0,1)$ for each $w(P_T)$ such that

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

Note that since $0 < \alpha, \beta < 1$,

$$g(P_T)\lambda\left(w^R - w^R\right)^\beta = g(P_T)\left(w^R - w^R\right)^\alpha = 0.$$

This implies that replacing the payoff $w(P_T)$ with the lottery $\left\{g(P_T), w^R; 1-g(P_T), \underline{w}\right\}$ leaves the participation constraint (6) and the incentive compatibility constraint (7) unchanged. From equation (25) and the strict concavity of $\lambda\left(w^R - w(P_T)\right)^\beta$ we have:

$$\begin{aligned} \lambda\left(w^R - w(P_T)\right)^\beta &= (1-g(P_T))\lambda\left(w^R - \underline{w}\right)^\beta \\ &< \lambda\left(w^R - \left(g(P_T)w^R + (1-g(P_T))\underline{w}\right)\right)^\beta, \end{aligned}$$

which implies that

$$g(P_T)w^R + (1-g(P_T))\underline{w} < w(P_T).$$

Hence the lottery $\left\{g\left(P_T\right), w^R; 1-g\left(P_T\right), \underline{w}\right\}$ improves on the original contract $w\left(P_T\right)$ in the sense that it provides the same incentives and utility to the manager and costs less to the firm.

Finally, consider a contract that pays off w' with $\underline{w} < w' < w^R$ with some probability p less than one. Then we can use the same argument as above, but we replace the random payoff w' with the lottery $\left\{g\left(P_T\right)p, w^R; \left(1-g\left(P_T\right)\right)p, \underline{w}\right\}$. ■

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Lemma 2 allows us to rewrite the principal's program (5) to (7) as follows:

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

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

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

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

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

$$\frac{\partial \mathcal{L}}{\partial w(P_T)} = f(P_T|\overline{e}) - \mu_{PC} \cdot \alpha \left(w\left(P_T\right) - w^R\right)^{\alpha-1} f(P_T|\overline{e}) - \mu_{IC} \cdot \alpha \left(w\left(P_T\right) - w^R\right)^{\alpha-1} \Delta f(P_T|e)$$

$$= \alpha \left(w\left(P_T\right) - w^R\right)^{\alpha-1} f(P_T|\overline{e}) \left[\frac{1}{\alpha}\left(w\left(P_T\right) - w^R\right)^{1-\alpha} - \mu_{PC} - \mu_{IC}\frac{\Delta f(P_T|e)}{f(P_T|\overline{e})}\right] . \tag{39}$$

Setting this equal to zero and solving for $w\left(P_T\right)$ yields the expression for $P_T > \widehat{P}$ in (8):

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

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

$$\frac{\partial \mathcal{L}}{\partial \widehat{P}} = \left(\underline{w} - w(\widehat{P})\right) f(\widehat{P}|\overline{e}) + \mu_{PC}\left(w(\widehat{P}) - w^R)^\alpha + \lambda\left(w^R - \underline{w}\right)^\beta\right) f(\widehat{P}|\overline{e})$$

$$+ \mu_{IC}\left(w(\widehat{P}) - w^R)^\alpha + \lambda\left(w^R - \underline{w}\right)^\beta\right) \Delta f(\widehat{P}|\overline{e}) \tag{41}$$

$$= -\left(w(\widehat{P}) - w^R)^\alpha + \lambda\left(w^R - \underline{w}\right)^\beta\right) f(\widehat{P}|\overline{e}) \tag{42}$$

$$\left[\frac{\left(w(\widehat{P}) - \underline{w}\right)}{\left(w(\widehat{P}) - w^R)^\alpha + \lambda\left(w^R - \underline{w}\right)^\beta\right)} - \mu_{PC} - \mu_{IC}\frac{\Delta f(\widehat{P}|e)}{f(\widehat{P}|\overline{e})}\right] . \tag{43}$$

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

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

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

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

Proof of Proposition 2: The shareholders' problem if they wish to minimize the contracting costs for implementing effort level $\hat{e}$ can be written as:

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

$$s.t. \quad \int V\left(w\left(P_T\right)\right) f(P_T|\hat{e}) dP_T \geq \underline{V} + C\left(\hat{e}\right) \quad , \tag{47}$$

$$\int V\left(w\left(P_T\right)\right) f_e(P_T|\hat{e}) dP_T \geq C' \ , \tag{48}$$

where C' denotes the first derivative of C and f_e denotes the first derivative of f with respect to e. Since optimization of program (46) to (48) is pointwise, the only changes with respect to program (5) to (7) are: replace ΔC with C', which is a constant for a given level of effort in both programs; replace $f(P_T|\overline{e})$ with $f(P_T|\hat{e})$, which is just a density that has the same properties in both programs; replace $\Delta f(P_T|e)$ with $f_e(P_T|\hat{e})$, which also has the same properties in both programs as we assume MLRP in both cases. Hence, the same arguments as in Proposition 1 go 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 (10), 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:

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

and the likelihood ratio is

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

Using the continuous effort analogue of the optimal contract as given in equation (8), and defining

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

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

allows us to write:

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

From this and equation (8), equation (11) follows immediately.

The optimal cut-off point is derived in the proof of Proposition 1, equation (45).

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 51, (September), pp. 1384-1399.

[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] Bebchuk, Lucian Arye, and Jesse M. Fried, 2004, Pay Without Performance - The Unfulfilled Promise of Executive Compensation, Cambridge (MA) and London (Harvard University Press).

[8] 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.

[9] ———, 1999, Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments , Management Science 45, no. 3 (May), pp. 364-381.

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

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

[12] 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.

[13] 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.

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

[15] 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.

[16] Coval, Joshua D., and Tyler Shumway, 2005, Do Behavioral Biases Affect Prices?, Journal of Finance 60, no. 1 (February), pp. 1-34.

[17] de Meza, David, and David C. Webb, 2007, Incentive Design Under Loss Aversion, Journal of the European Economic Association, forthcoming.

[18] Dittmann, Ingolf, and Ernst. Maug, 2007, Lower Salaries and No Options? On the Optimal Structure of Executive Pay, Journal of Finance, no. 62, pp. 303-343.

[19] 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.

[20] Dybvig, Philip H., and Mark Loewenstein, 2003, Employee Reload Options: Pricing, Hedging, and Optimal Exercise, Review of Financial Studies 16, no. 1 (Spring), pp. 145-171.

[21] Edmans, Alex, Xavier Gabaix, and Augustin Landier, 2007, A Calibratable Model of Optimal CEO Incentives in Market Equilibrium, Working Paper, New York University (September).

[22] 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.

[23] 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.

[24] 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.

[25] Gomes, Francisco J., 2005, Portfolio Choice and Trading Volume With Loss Averse Investors, Journal of Business 78, no. 2 , pp. 675-706.

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

[27] 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.

[28] 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.

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

[30] ———, 2002, Stock Options for Undiversified Executives, Journal of Accounting and Economics 33, no. 2 (April), pp. 3-42.

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

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

[33] Hemmer, Thomas, Oliver Kim, and Robert E. Verrecchia, 1999, Introducing Convexity into Optimal Compensation Contracts, Journal of Accounting and Economics 28, no. 3 (December), pp. 307-327.

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

[35] Inderst, Roman, and Holger M. Müller, 2005, A Theory of Broad-Based Option Pay, CEPR Discussion Paper, no. 4878 (January).

[36] Ingersoll, Jonathan E. Jr., 2006, The Subjective and Objective Evaluation of Incentive Stock Options, Journal of Business 79, no. 2 (March).

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

[38] 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.

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

[40] ———, 1984, Choices, Values, and Frames, American Psychologist, 39(4), pp. 341-350.

[41] 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.

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

[43] 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.

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

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

[46] 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.

[47] 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.

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

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

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

[51] 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.

[52] Massa, Massimo, and Andrei Simonov, 2005, Behavioral Biases and Investment, Review of Finance 9, pp. 483-507.

[53] 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.

[54] 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.

[55] 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.

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

[57] 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.

[58] Post, Thierry, Marijn van den Assem, Guido Baltussen, and Richard Thaler, 2007, Deal or No Deal? Decision making under risk in a large-payoff game show, American Economic Review, forthcoming.

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

[60] Ray, Debraj, 2002, The Time Structure of Self-Enforcing Agreements, Econometrica 70, no. 2 (March), pp. 547-582.

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

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

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

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

[65] ———, 1999, Mental Accounting Matters, Journal of Behavioral Decision Making 12, pp. 183-206.

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

[67] ———, 1991, Loss Aversion in Riskless Choice: A Reference-Dependent Model, The Quarterly Journal of Economics, 106, no. 4 (November), pp. 1039-1061.

[68] ———, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty, 5, pp. 297-323.

[69] 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 eleven variables for our main sample of 595 CEOs from 2005 (panel A) and a second sample of 576 CEOs from 1997 (Panel B). "Value of Contract" is the market value of the compensation package $\pi = \phi + n_S{*}P_0 + n_O{*}BS$, where BS is the Black-Scholes option value. All dollar amounts are given in thousands ('000).

Panel A: Sample for 2005

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Stock	n_S	1.87%	5.18%	0.04%	0.31%	3.78%
Options	n_O	1.44%	1.42%	0.15%	1.03%	3.24%
Fixed Salary	ϕ	2,496	3,107	594	1,675	4,694
Value of Contract	π	178,966	1,887,655	5,523	29,837	157,961
Non-firm Wealth	W_0	33,285	113,239	2,268	10,298	60,858
Firm Value	P_0	10,650,934	30,260,334	342,422	2,274,781	19,810,415
Strike Price	K	8,243,201	26,213,423	242,240	1,479,528	13,915,001
Moneyness	K/P_0	70.06%	20.54%	40.26%	70.81%	98.94%
Maturity	T	4.58	1.30	3.39	4.44	6.01
Stock Volatility	σ	42.83%	21.42%	22.90%	36.10%	75.10%
Dividend Rate	d	1.24%	2.70%	0.00%	0.61%	3.28%

Panel B: Sample for 1997

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Stock	n_S	2.50%	6.01%	0.02%	0.28%	8.32%
Options	n_O	1.01%	1.35%	0.00%	0.56%	2.54%
Fixed Salary	ϕ	1,786	4,454	459	1,141	2,966
Value of Contract	π	118,319	1,046,636	2,409	15,528	93,686
Non-firm Wealth	W_0	15,270	67,782	1,186	4,253	25,807
Firm Value	P_0	5,236,535	11,209,383	258,109	1,540,377	11,284,427
Strike Price	K	3,777,856	8,251,907	192,662	1,085,677	8,186,544
Moneyness	K/P_0	76.27%	22.43%	47.93%	77.15%	100.00%
Maturity	T	5.58	1.86	4.10	5.22	7.34
Stock Volatility	σ	29.28%	13.11%	16.20%	26.00%	47.40%
Dividend Rate	d	1.83%	1.90%	0.00%	1.46%	4.42%

Table 2: Optimal piecewise linear contracts

This table describes the optimal piecewise linear contract for the base model where options and salary can become negative ($n_O \geq -n_S \exp(dT)$, $\phi \geq -W_0$). It shows mean and median of the three contract parameters base salary ϕ^*, stock holdings n_S^* and option holdings n_O^* together with the mean of the errors $error(\phi)=(\phi^* - \phi^d)/\sigma_\phi$, $error(n_S)=(n^*_S - n^d_S)/\sigma_S$, and $error(n_O)=(n^*_O - n^d_O)/\sigma_O$. The table also shows mean and median of the two distance metrices D_{Lin} and D_{LinS}, and the average probability of a loss, i.e., Prob($w^*(P_T) < w^R$). Panel A displays the results for the Loss Aversion Model for seven different reference wages parameterized by θ. Panels B and C show the results for the CRRA Model and, respectively, the CARA Model for six levels of the CRRA risk aversion parameter γ. For the results in Panel C, we calculate each CEO's coefficient of absolute risk aversion ρ as $\rho = \gamma / (W_0 + \pi_0)$, where π_0 is the market value of her observed compensation package. The last row in Panel A shows the corresponding values of the observed contract.

Panel A: Loss Aversion Model

θ	Obs.	Avg. Prob. of Loss	Salary (ϕ) Mean	Salary (ϕ) Median	Salary (ϕ) Mean Deviation	Stock (n_S) Mean	Stock (n_S) Median	Stock (n_S) Mean Deviation	Options (n_O) Mean	Options (n_O) Median	Options (n_O) Mean Deviation	D_{Lin} Mean	D_{Lin} Median	D_{LinS} Mean	D_{LinS} Median
0.0	594	4.07%	-2,458	292	-1.595	0.024	0.005	0.103	0.008	0.007	-0.429	0.543	0.163	1.946	0.737
0.1	578	13.60%	3,597	1,468	0.346	0.019	0.005	0.015	0.014	0.009	-0.022	0.713	0.148	2.226	0.674
0.2	571	20.07%	2,356	1,293	-0.049	0.022	0.006	0.050	0.012	0.007	-0.135	1.438	0.400	5.298	1.565
0.4	585	31.27%	11,980	-2,891	3.027	0.034	0.011	0.285	-0.002	0.001	-1.136	2.395	0.865	14.187	3.285
0.6	586	41.13%	9,816	-6,744	2.337	0.046	0.017	0.526	-0.018	-0.002	-2.271	3.069	1.132	19.766	4.286
0.8	585	51.05%	-7,800	-8,264	-3.294	0.053	0.018	0.647	-0.027	-0.003	-2.921	3.321	1.234	18.241	4.724
1.0	582	58.35%	-31,120	-8,894	-10.729	0.056	0.019	0.708	-0.032	-0.005	-3.292	3.471	1.275	12.484	4.906
Data	595	N/A	2,496	1,675	N/A	0.019	0.003	N/A	0.014	0.010	N/A	N/A	N/A	N/A	N/A

Panel B: Constant Relative Risk Aversion (CRRA) Model

γ	Obs.	Salary (ϕ)			Stock (n_S)			Options (n_O)			D_{Lin}		D_{LinS}	
		Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median	Mean	Median
0.1	595	-27,754	-8,682	-9.738	0.055	0.018	0.699	-0.029	-0.004	-3.083	3.163	1.086	11.350	4.785
0.2	593	-28,448	-8,833	-9.948	0.056	0.018	0.724	-0.032	-0.005	-3.265	3.351	1.167	11.680	4.819
0.5	595	-29,035	-8,845	-10.150	0.058	0.020	0.749	-0.037	-0.007	-3.625	3.709	1.395	12.102	5.084
1	593	-27,864	-8,271	-9.770	0.055	0.021	0.702	-0.040	-0.011	-3.827	3.894	1.680	11.843	5.086
3	594	-18,277	-5,166	-6.682	0.044	0.016	0.494	-0.042	-0.014	-3.953	3.991	2.014	9.106	4.148
6	585	-8,668	-1,229	-3.582	0.033	0.010	0.260	-0.033	-0.011	-3.384	3.400	1.826	5.982	2.815
20	487	1,017	971	-0.418	0.025	0.005	0.050	-0.026	-0.006	-2.779	2.807	1.538	3.000	1.687

Panel C: Constant Absolute Risk Aversion (CARA) Model

γ	Obs.	Salary (ϕ)			Stock (n_S)			Options (n_O)			D_{Lin}		D_{LinS}	
		Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median	Mean Deviation	Mean	Median	Mean	Median
0.1	595	-31,632	-9,207	-10.986	0.059	0.019	0.777	-0.037	-0.005	-3.586	3.672	1.258	12.870	5.246
0.2	595	-31,390	-9,104	-10.908	0.058	0.019	0.752	-0.037	-0.006	-3.600	3.680	1.341	12.782	5.266
0.5	595	-30,392	-9,016	-10.587	0.055	0.020	0.700	-0.037	-0.008	-3.634	3.704	1.487	12.466	5.144
1	595	-28,458	-8,271	-9.964	0.052	0.020	0.633	-0.037	-0.010	-3.631	3.690	1.604	11.862	4.931
3	594	-21,073	-6,087	-7.583	0.044	0.016	0.492	-0.038	-0.013	-3.709	3.747	1.833	9.703	4.284
6	595	-13,772	-3,160	-5.237	0.037	0.012	0.346	-0.035	-0.011	-3.484	3.507	1.858	7.444	3.532
20	590	-1,530	539	-1.292	0.024	0.007	0.103	-0.025	-0.007	-2.793	2.797	1.525	3.594	1.941

Table 3: Comparison of Loss Aversion-model with matched Risk Aversion-model

This table compares the optimal Loss Aversion-contract with the equivalent optimal (Constant Relative) Risk Aversion-contract where each CEO has constant relative risk aversion with parameter γ, which is chosen such that both models predict the same certainty equivalent for the observed contract (equation (19)). Contracts are piecewise linear, and options and salary can become negative ($n_O \geq -n_S\exp(dT)$, $\phi \geq -W_0$). Panel A shows the average equivalent γ, mean and median of the difference between the two distance metrices D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequency of these differences being positive. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level. Panel B shows the frequency of positive optimal option holdings, the frequency of positive optimal salaries, and the frequency of both (options and salary) being positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

θ	Obs.	Average equivalent γ	$D_{Lin}^{RA} - D_{Lin}^{LA}$			$D_{LinS}^{RA} - D_{LinS}^{LA}$		
			Percent > 0	Mean	Median	Percent > 0	Mean	Median
0.0	594	0.21	96.63%	2.751***	0.923***	96.63%	9.682***	3.673***
0.1	578	0.28	97.23%	2.644***	0.875***	97.40%	9.424***	3.833***
0.2	571	0.41	91.77%	2.040***	0.625***	92.12%	6.499***	2.956***
0.3	577	0.52	87.69%	1.541***	0.441***	88.73%	0.909	1.552***
0.4	585	0.68	89.74%	1.196***	0.298***	88.72%	-2.263	0.591***
0.5	586	0.83	90.96%	0.946***	0.271***	88.40%	-4.111	0.248***
0.6	586	0.95	90.27%	0.720***	0.247***	86.01%	-7.722	0.092***
0.7	582	1.04	88.66%	0.590***	0.235***	83.33%	-11.767	0.070***
0.8	582	1.09	86.43%	0.589***	0.217***	79.04%	-6.276	0.047***
0.9	579	1.06	84.11%	0.504***	0.183***	78.07%	-0.207	0.033***
1.0	581	0.98	82.10%	0.380***	0.138***	76.94%	-0.586	0.024***

Panel B: Positive option holdings and positive salaries

θ	Percent with positive option holdings		Percent with positive fixed salary		Percent with positive options and salary	
	RA	LA	RA	LA	RA	LA
0.0	30.81%	83.33%	1.68%	59.60%	0.34%	52.53%
0.1	30.10%	91.00%	1.56%	77.51%	0.00%	74.22%
0.2	28.20%	81.96%	1.93%	62.70%	0.35%	60.25%
0.3	28.08%	68.28%	1.56%	46.79%	0.35%	44.02%
0.4	25.81%	56.92%	1.37%	32.65%	0.00%	30.60%
0.5	25.60%	48.29%	1.71%	20.65%	0.34%	19.28%
0.6	22.53%	41.30%	1.54%	12.80%	0.00%	11.09%
0.7	20.79%	36.60%	2.06%	8.59%	0.00%	6.36%
0.8	20.96%	33.68%	2.06%	6.53%	0.00%	4.12%
0.9	21.24%	32.47%	2.25%	4.15%	0.17%	2.59%
1.0	22.20%	31.50%	2.07%	3.27%	0.00%	1.89%

Table 4: Optimal nonlinear Loss Aversion contracts

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

θ	Obs.	Mean Δ_{Low}	Mean Δ_{High}	Percent $\Delta_{High} > \Delta_{Low}$	Mean Dismissal Probability	Incentives from Dismissals	Mean Inflection Quantile
0.0	571	2.06%	2.58%	90.37%	0.65%	1.78%	85.54%
0.1	571	1.57%	2.50%	95.27%	1.46%	3.97%	92.84%
0.2	570	1.07%	2.29%	97.37%	2.84%	7.90%	96.19%
0.3	574	0.88%	2.32%	97.56%	4.46%	12.94%	97.00%
0.4	572	0.73%	1.97%	98.08%	6.60%	19.21%	97.60%
0.5	573	0.69%	2.11%	98.25%	8.80%	25.79%	97.94%
0.6	573	0.52%	1.93%	98.25%	11.42%	33.60%	98.12%
0.7	574	0.40%	1.78%	98.08%	13.96%	41.07%	98.07%
0.8	569	0.35%	1.59%	98.24%	16.44%	48.31%	98.21%
0.9	563	0.31%	1.54%	98.93%	18.83%	54.43%	98.68%
1.0	547	0.28%	1.36%	98.90%	21.08%	59.85%	98.41%

Table 5: Comparison of linear and nonlinear Loss Aversion models

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

		Linear Option Contract				General Nonlinear contract			
θ	Obs.	Mean n_S	Mean $n_S + n_O$	Mean Savings	Mean D_{NonLin}	Mean Δ_{Low}	Mean Δ_{High}	Mean Savings	Mean D_{NonLin}
0.0	570	0.0186	0.0273	0.0015	0.1517	0.0206	0.0259	0.0051	0.2208
0.1	557	0.0155	0.0283	0.0041	0.2012	0.0158	0.0252	0.0153	0.1942
0.2	547	0.0186	0.0277	0.0099	0.3859	0.0109	0.0230	0.0335	0.2469
0.3	559	0.0268	0.0290	0.0165	0.4622	0.0089	0.0233	0.0515	0.2787
0.4	567	0.0319	0.0258	0.0228	0.6343	0.0073	0.0197	0.0689	0.4309
0.5	571	0.0410	0.0282	0.0296	0.6233	0.0070	0.0211	0.0844	0.4338
0.6	570	0.0466	0.0266	0.0372	0.7047	0.0052	0.0194	0.1015	0.4855
0.7	573	0.0497	0.0251	0.0434	0.7373	0.0040	0.0178	0.1159	0.5243
0.8	569	0.0516	0.0245	0.0495	0.7384	0.0035	0.0159	0.1298	0.5415
0.9	561	0.0546	0.0255	0.0533	0.7178	0.0031	0.0155	0.1406	0.5566
1.0	546	0.0553	0.0253	0.0564	0.7353	0.0028	0.0136	0.1502	0.5918

Table 6: Wealth Robustness Check

This table contains the main results from repeating our analysis shown in Table 3 when we decrease or increase our wealth estimates by a factor of two. For Panel A, our wealth estimate W_0 is multiplied by 0.5. For Panel B, it is multiplied by 2. Both panels show the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequencies that option holdings and salary are both positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level.

Panel A: Results for lower wealth (-50%)

θ	Obs.	Average equivalent	$D_{Lin}^{RA} - D_{Lin}^{LA}$		$D_{LinS}^{RA} - D_{LinS}^{LA}$		Percent with positive options and salary	
		γ	Mean	Median	Mean	Median	RA	LA
0.0	594	0.17	1.699***	0.535***	5.149***	2.009***	0.34%	53.20%
0.1	577	0.23	1.609***	0.444***	4.880***	2.141***	0.00%	74.52%
0.2	572	0.33	1.098***	0.269***	2.660***	1.355***	0.00%	59.79%
0.3	574	0.42	0.708***	0.164***	-2.406	0.452***	0.17%	44.25%
0.4	588	0.54	0.521***	0.125***	-4.898	0.111***	0.34%	30.10%
0.5	586	0.66	0.496***	0.141***	-6.259	0.074***	0.00%	18.94%
0.6	591	0.77	0.476***	0.165***	-8.734	0.066***	0.00%	11.34%
0.7	585	0.84	0.433***	0.170***	-12.122	0.057***	0.17%	6.32%
0.8	587	0.88	0.506***	0.174***	-6.056	0.054***	0.00%	3.92%
0.9	583	0.86	0.504***	0.160***	0.320***	0.048***	0.00%	2.57%
1.0	585	0.81	0.434***	0.135***	0.140	0.037***	0.00%	1.88%

Panel B: Results for higher wealth (+100%)

θ	Obs.	Average equivalent	$D_{Lin}^{RA} - D_{Lin}^{LA}$		$D_{LinS}^{RA} - D_{LinS}^{LA}$		Percent with positive options and salary	
		γ	Mean	Median	Mean	Median	RA	LA
0.0	592	0.27	4.004***	1.566***	16.516***	6.323***	0.51%	52.53%
0.1	576	0.38	3.853***	1.590***	16.151***	6.353***	0.17%	74.48%
0.2	568	0.57	3.117***	1.181***	12.290***	4.867***	0.00%	60.21%
0.3	577	0.71	2.476***	0.883***	6.270**	3.378***	0.17%	44.02%
0.4	585	0.93	1.867***	0.650***	1.745	1.824***	0.34%	30.43%
0.5	579	1.14	1.472***	0.520***	-1.197	0.730***	0.00%	19.34%
0.6	587	1.31	1.036***	0.370***	-6.131	0.200***	0.17%	11.24%
0.7	581	1.42	0.778***	0.306***	-11.383	0.062***	0.00%	6.37%
0.8	578	1.48	0.698***	0.256***	-6.242	0.024	0.17%	4.15%
0.9	577	1.43	0.461***	0.188***	-1.660	0.012	0.00%	2.60%
1.0	575	1.33	0.206**	0.142***	-2.376**	0.006**	0.00%	1.91%

Table 7: Restricted models with positive salaries and positive option holdings

This table contains the results from repeating the analysis shown in Table 3 with the stricter constraints that option holdings and salaries must be non-negative ($n_O \geq 0$, $\phi \geq 0$). The table compares the optimal Loss Aversion-contract with the equivalent optimal (Constant Relative) Risk Aversion-contract where each CEO's risk aversion parameter γ is chosen such that both models predict the same certainty equivalent for the observed contract (equation (19)). Panel A shows the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequency of these differences being positive. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level. Panel B shows the frequency of positive optimal option holdings, the frequency of positive optimal salaries, and the frequency of both (options and salary) being positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

θ	Obs.	Average equivalent γ	$D^{RA}_{Lin} - D^{LA}_{Lin}$			$D^{RA}_{LinS} - D^{LA}_{LinS}$		
			Percent > 0	Mean	Median	Percent > 0	Mean	Median
0.0	588	0.21	49.49%	-0.024	0.000***	52.04%	0.076	0.024***
0.1	574	0.28	50.87%	-0.221***	0.001	52.96%	-0.763**	0.030***
0.2	569	0.41	39.02%	-0.537***	-0.004***	41.12%	-2.186***	-0.001***
0.3	573	0.53	50.26%	-0.682***	0.000***	51.13%	-6.587**	0.000***
0.4	584	0.68	61.47%	-0.612***	0.003	61.99%	-8.148*	0.000
0.5	584	0.83	74.14%	-0.465***	0.010***	73.63%	-8.740	0.002***
0.6	586	0.95	78.50%	-0.394***	0.017***	77.65%	-10.436	0.003***
0.7	585	1.05	82.39%	-0.354***	0.022***	81.37%	-13.130	0.004***
0.8	582	1.09	83.16%	-0.189***	0.024***	81.96%	-6.626	0.005***
0.9	583	1.06	82.85%	-0.111**	0.022***	82.16%	-0.156***	0.004***
1.0	577	0.98	82.32%	-0.110**	0.018***	81.98%	-0.159***	0.003***

Panel B: Positive option holdings and positive salaries

θ	Percent with positive option holdings		Percent with positive fixed salary		Percent with positive options and salary	
	RA	LA	RA	LA	RA	LA
0.0	84.18%	89.80%	15.82%	65.31%	0.34%	58.33%
0.1	82.93%	94.43%	16.55%	81.71%	0.00%	78.22%
0.2	81.37%	94.73%	17.93%	67.14%	0.18%	64.50%
0.3	81.33%	92.50%	17.45%	53.23%	0.00%	49.91%
0.4	79.79%	90.58%	19.01%	39.73%	0.00%	37.33%
0.5	79.45%	89.55%	19.52%	29.11%	0.17%	27.05%
0.6	78.84%	87.71%	19.97%	20.48%	0.00%	18.09%
0.7	78.80%	87.01%	20.17%	15.21%	0.00%	12.65%
0.8	79.04%	86.08%	20.10%	14.78%	0.00%	12.20%
0.9	78.90%	85.59%	20.41%	12.18%	0.00%	9.61%
1.0	80.94%	85.10%	18.72%	9.01%	0.35%	6.93%

Table 8: Results for the 1997 sample

This table contains the results from repeating the analysis shown in Table 3 for data for 1997. The table compares the optimal Loss Aversion-contract with the equivalent optimal (Constant Relative) Risk Aversion-contract where each CEO's risk aversion parameter γ is chosen such that both models predict the same certainty equivalent for the observed contract (equation (19)). Contracts are piecewise linear, and options and salary can become negative ($n_O \geq -n_S\exp(dT)$, $\phi \geq -W_0$). Panel A shows the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequency of these differences being positive. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level. Panel B shows the frequency of positive optimal option holdings, the frequency of positive optimal salaries, and the frequency of both (options and salary) being positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

θ	Obs.	Average equivalent γ	$D^{RA}_{Lin} - D^{LA}_{Lin}$ Percent > 0	Mean	Median	$D^{RA}_{LinS} - D^{LA}_{LinS}$ Percent > 0	Mean	Median
0.0	569	0.20	95.08%	1.830***	0.430***	95.08%	6.031***	2.451***
0.1	545	0.22	97.43%	1.795***	0.511***	97.80%	3.604***	1.458***
0.2	547	0.26	92.87%	1.705***	0.406***	93.24%	3.058***	1.333***
0.3	557	0.33	89.23%	1.222***	0.306***	89.41%	0.905	1.058***
0.4	555	0.42	85.95%	0.769***	0.221***	85.77%	-1.168	0.681***
0.5	557	0.53	84.20%	0.445***	0.180***	83.84%	-3.182	0.389***
0.6	565	0.64	85.84%	0.469***	0.154***	85.84%	-4.526	0.279***
0.7	558	0.76	89.61%	0.674***	0.163***	89.43%	-4.755	0.194***
0.8	564	0.89	93.26%	0.922***	0.175***	92.20%	-1.749	0.209***
0.9	564	1.01	94.86%	1.023***	0.185***	93.79%	0.931***	0.184***
1.0	567	1.10	94.18%	0.920***	0.177***	93.47%	0.839***	0.138***

Panel B: Positive option holdings and positive salaries

θ	Percent with positive option holdings		Percent with positive fixed salary		Percent with positive options and salary	
	RA	LA	RA	LA	RA	LA
0.0	27.07%	70.47%	6.33%	50.26%	0.00%	37.79%
0.1	27.71%	86.61%	6.97%	85.87%	0.18%	74.31%
0.2	26.14%	88.30%	6.95%	88.85%	0.18%	80.07%
0.3	26.21%	84.38%	7.18%	82.23%	0.54%	75.04%
0.4	25.05%	77.84%	6.85%	70.09%	0.18%	63.78%
0.5	24.78%	70.74%	6.82%	59.25%	0.00%	53.50%
0.6	24.60%	63.72%	6.90%	47.96%	0.18%	41.95%
0.7	24.37%	57.53%	6.81%	39.07%	0.18%	33.69%
0.8	23.40%	51.06%	6.91%	30.85%	0.35%	26.06%
0.9	23.05%	44.86%	6.74%	22.70%	0.00%	18.26%
1.0	21.52%	40.74%	6.88%	18.34%	0.18%	13.93%

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

This table describes the optimal piecewise linear Loss Aversion contract for different values of the parameters α, β, and λ of the value function. The reference wage w^R is set equal to last year's fixed salary plus 10% of the risk-neutral value of last year's stock and option holdings, i.e. $\theta = 0.1$ in equation (18). Panel A shows the results for the parameter λ, Panel B for α, and Panel C for β. Options and salary can become negative ($n_O \geq -n_S\exp(dT)$, $\phi \geq -W_0$). The table shows mean and median of the three contract parameters base salary ϕ^*, stock holdings n_S^* and option holdings n_O^*. In addition, it displays mean and median of the distance metric D_{Lin}. Some observations are lost because of numerical problems.

Panel A: Loss aversion parameter λ

λ	Obs.	Salary (ϕ)		Stock (n_S)		Options (n_O)		D_{Lin}	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
1.00	413	-2,285	260	0.0404	0.0074	-0.0132	0.0035	2.0505	0.1673
1.50	452	1,533	1,342	0.0258	0.0058	0.0070	0.0075	0.7650	0.1344
2.00	459	2,364	1,581	0.0219	0.0051	0.0128	0.0093	0.5513	0.1259
2.25	578	3,597	1,468	0.0191	0.0047	0.0141	0.0095	0.7132	0.1478
2.50	471	2,607	1,628	0.0209	0.0055	0.0155	0.0099	0.6176	0.1242
3.00	465	2,825	1,684	0.0205	0.0048	0.0154	0.0098	0.6278	0.1271
4.00	466	2,972	1,770	0.0200	0.0051	0.0171	0.0102	0.6821	0.1292

Panel B: Gain space curvature α

α	Obs.	Salary (ϕ)		Stock (n_S)		Options (n_O)		D_{Lin}	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.60	312	-1,303	505	0.0235	0.0059	-0.0037	0.0022	1.2413	0.2759
0.70	362	-984	389	0.0231	0.0070	-0.0034	0.0031	1.3861	0.2581
0.80	394	87	809	0.0281	0.0070	0.0026	0.0062	1.0471	0.1800
0.88	578	3,597	1,468	0.0191	0.0047	0.0141	0.0095	0.7132	0.1478
0.95	546	2,974	1,930	0.0193	0.0035	0.0159	0.0107	0.7649	0.1122

Panel C: Loss space curvature β

β	Obs.	Salary (ϕ)		Stock (n_S)		Options (n_O)		D_{Lin}	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.60	267	-4,585	-1,329	0.0224	0.0070	-0.0016	0.0018	1.1690	0.2942
0.70	349	-6,244	-1,396	0.0341	0.0092	-0.0142	0.0015	2.3636	0.3923
0.80	388	163	940	0.0356	0.0057	-0.0053	0.0068	1.6406	0.1434
0.88	578	3,597	1,468	0.0191	0.0047	0.0141	0.0095	0.7132	0.1478
0.95	508	2,584	1,636	0.0204	0.0053	0.0153	0.0097	0.6294	0.1297

Table 10: Ownership Robustness Check

This table contains the main results from repeating our analysis shown in Table 3 when we split our sample according to the stock ownership of the CEOs. Panel A displays the 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. Both panels show the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequencies that option holdings and salaries are both positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level.

Panel A: Results for owner-managers ($n_S \geq 5\%$)

θ	Obs.	Average equivalent	$D^{RA}_{Lin} - D^{LA}_{Lin}$		$D^{RA}_{LinS} - D^{LA}_{LinS}$		Percent with positive options and salary	
		γ	Mean	Median	Mean	Median	RA	LA
0.0	54	0.16	11.057***	6.602***	25.808***	14.819***	0.00%	5.56%
0.1	51	0.19	11.025***	6.793***	26.889***	14.483***	0.00%	82.35%
0.2	54	0.26	7.241***	6.015***	20.333***	11.565***	0.00%	90.74%
0.3	53	0.34	6.197***	4.001***	-16.935	9.197***	0.00%	79.25%
0.4	54	0.46	5.130***	1.650***	-42.166	4.189***	0.00%	57.41%
0.5	54	0.62	4.291***	3.011***	-63.667	1.659***	0.00%	38.89%
0.6	54	0.77	3.049**	3.426***	-99.527	2.047**	0.00%	24.07%
0.7	54	0.86	2.718*	3.322***	-131.139	2.875**	0.00%	12.96%
0.8	54	0.92	3.492***	2.794***	-67.172	1.607**	0.00%	9.26%
0.9	53	0.92	3.436***	1.811***	-0.343	0.711**	0.00%	1.89%
1.0	54	0.86	2.759***	1.837***	-3.687	0.422**	0.00%	0.00%

Panel B: Results for non-owner managers ($n_S < 5\%$)

θ	Obs.	Average equivalent	$D^{RA}_{Lin} - D^{LA}_{Lin}$		$D^{RA}_{LinS} - D^{LA}_{LinS}$		Percent with positive options and salary	
		γ	Mean	Median	Mean	Median	RA	LA
0.0	540	0.21	1.920***	0.792***	8.070***	3.516***	0.37%	57.22%
0.1	527	0.29	1.833***	0.766***	7.734***	3.541***	0.00%	73.43%
0.2	517	0.43	1.497***	0.560***	5.054***	2.764***	0.39%	57.06%
0.3	524	0.54	1.070***	0.414***	2.713***	1.486***	0.38%	40.46%
0.4	531	0.70	0.796***	0.277***	1.795***	0.504***	0.00%	27.87%
0.5	532	0.85	0.607***	0.235***	1.934***	0.177***	0.38%	17.29%
0.6	532	0.97	0.484***	0.226***	1.597*	0.078***	0.00%	9.77%
0.7	528	1.06	0.373***	0.213***	0.442***	0.057***	0.00%	5.68%
0.8	528	1.11	0.292***	0.190***	-0.048	0.038***	0.00%	3.60%
0.9	526	1.07	0.209***	0.158***	-0.194*	0.025***	0.19%	2.66%
1.0	527	1.00	0.136***	0.124***	-0.269***	0.019***	0.00%	2.09%

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.org/wp) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html