Sticks or Carrots?

Optimal CEO Compensation when Managers are Loss Averse

INGOLF DITTMANN, ERNST MAUG and OLIVER SPALT*

forthcoming in the *Journal of Finance*

Abstract

This paper analyzes optimal executive compensation contracts when managers are loss averse. We calibrate a stylized principal-agent model to the observed contracts of 595 CEOs and show that this model can explain observed option holdings and high base salaries remarkably well for a range of parameterizations. We also derive and calibrate the general shape of the optimal contract that is increasing and convex for medium and high outcomes and drops discontinuously to the lowest possible payout for low outcomes. We identify the critical features of the loss-aversion model that render optimal contracts convex.

*Ingolf Dittmann is at the Erasmus University Rotterdam. Ernst Maug is at the University of Mannheim. Oliver Spalt is at Tilburg University. The authors are grateful to seminar participants at the University of Cologne, Frankfurt, Georgia State, Humboldt, Mannheim, Maryland, Tilburg, the European Finance Association meeting in Ljubljana, the NBER Behavioral Economics Working Group Meeting in Yale, the European Accounting Association meeting in Rotterdam, the JFI-Conference on "Financial Contracting", the 6th Oxford Finance Symposium, the DGF-conference in Dresden, the GEABA-conference in Tübingen, and to Bo Becker, Axel Börsch-Supan, Alex Edmans, Xavier Gabaix, Campbell Harvey (the Editor), Gerard Hoberg, Andreas Knabe, Matjaz Koman, Roy Kouwenberg, David Larcker, Christian Laux, David De Meza, Werner Neus, and an anonymous referee for valuable 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.

In this paper we analyze a simple contracting model where the manager is loss averse and explore to what extent its predictions are consistent with salient features of observed compensation contracts. We parameterize this model using standard assumptions and then compare the contracts generated by the model with those actually observed for a large sample of U.S. CEOs. Our main conclusion is that for a range of parameterizations a principal-agent model with loss-averse agents generates convex compensation contracts and can approximate observed contracts better than the standard model based on risk aversion that is widely used in the literature.

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. Quantitative studies in compensation research 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. Dittmann and Maug (2007) find that the optimal predicted contract almost never contains any options and typically features negative base salaries.[2]

In this paper we suggest a new 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) loom larger than gains; (iii) diminishing sensitivity, so that individuals become progressively less

[1] Calibration exercises with CRRA preferences and lognormal distributed stock prices include Lambert, Larcker, and Verrecchia (1991), Hall and Murphy (2000, 2002), and Hall and Knox (2004). Closely related to this are papers that combine CRRA-preference with geometric Brownian motions or a binomial approach in a dynamic model, see Huddart (1994), Carpenter (1998), Johnson and Tian (2000a, 2000b) and Ingersoll (2006). 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.

[2] There are a few extensions of the standard risk-aversion model that can explain option holdings. Feltham and Wu (2001) and Dittmann and Yu (2009) consider risk-taking incentives, Oyer (2004) models options as a device to retain employees when recontracting is expensive, and Inderst and Müller (2005) explain options as instruments that provide outside shareholders with better liquidation incentives. Hemmer, Kim, and Verrecchia (1999) assume gamma distributed stock prices and find convex contracts, but Dittmann and Maug (2007) show that these results are not robust. With the exception of Dittmann and Yu (2009), none of these models has ever been calibrated to data, and some models are too stylized to be calibrated at all.

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.[3]

We develop a stylized principal-agent model with a loss-averse agent and show how it can be calibrated to individual CEO data. In the first part of the paper, we consider piecewise linear contracts that consist of fixed salary, stock and one option grant. Our calibration method yields a predicted contract for each of the 595 U.S. CEOs in our data set, and we compare these predicted contracts with observed contracts in order to determine whether the model is descriptive of the data. The literature provides no guidance on the appropriate level of the reference point, so we repeat the analysis for a range of different levels of the manager's reference wage. We find that the loss-aversion model can explain observed contracts very well if the manager's reference wage is low, i.e. not much higher than last year's fixed salary and bonus. The model then predicts realistic option holdings and salaries. For higher reference wages however, the loss-aversion model cannot generate observed contracts; it then predicts negative option holdings and negative salaries, just like the benchmark risk-aversion model.

In order to understand these results, we derive, analyze, and calibrate the optimal nonlinear contract in the second part of the paper. It turns out that the optimal contract features two regions: First, above a certain critical stock price, it is increasing and convex and pays out at least the reference wage. In the second region below this critical stock price, compensation falls discontinuously to a lower bound, a feature that is reminiscent of performance-related dismissals. The intuition for this drop is that a loss-averse CEO is risk-loving whenever her pay falls below her reference wage. Therefore, whenever the contract pays off below the reference wage, it can only pay the lowest possible wage, because any intermediate payoff can be improved upon by a lottery over more extreme payoffs.

If the reference wage is low, the optimal contract features a low dismissal probability and is dominated by the region where the contract is increasing and convex. This shape can be approximated

[3]See Kahneman and Tversky (2000) and Post et al. (2008) as well as the papers they cite. There is broad experimental support for loss aversion in particular: Rabin (2000) calls loss aversion the "most firmly established feature of risk preferences." Recent experimental evidence in finance includes Gneezy, Kapetyn, and Potters (2003) and Haigh and List (2005). Chen, Lakshminarayanan and Santos (2006) experiment with capuchin monkeys and suggest that loss aversion is a basic evolutionary trait that extends beyond humans. 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. 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.

with high salaries and positive stock and option holdings that we observe in practice. As the reference wage increases, the dismissal probability increases while the contract becomes flatter in the region above the critical stock price where the CEO keeps his job. Consequently, effort incentives are provided increasingly through the threat of dismissal and less through pay-for-performance "on the job." For high reference wages, the piecewise linear contract therefore approximates the discontinuous drop in compensation with negative salaries, high stock holdings and negative option holdings.

The loss-aversion model can generate convex contracts, because the agent's aversion to risk is concentrated in the neighborhood of the reference point, where risk tolerance is almost zero. As the payout increases, the agent becomes rapidly more risk tolerant. Optimal risk sharing implies that the contract is more high-powered where the risk-tolerance of the CEO is higher, so the fast increase of risk-tolerance leads to optimal contracts that are convex. By contrast in the risk-aversion model, risk-tolerance is mainly determined by the agent's wealth and therefore increases only slowly with increasing payouts. As a consequence, the shape of the optimal risk-aversion contract is dominated by a second effect: the decreasing marginal utility that makes it more effective to provide incentives at low payout levels. This incentive effect by itself results in a concave contract. The incentive effect is also present in the loss-aversion model, but is dominated by the risk-tolerance effect. The ability to quantitatively disentangle the importance of these effects for individual CEO compensation contracts is a strength of our calibration approach.

A noteworthy cross-sectional implication of our model is that CEOs with a higher reference wage have less performance-related pay, but they are attracted to firms with better corporate governance standards where they have a higher probability of being fired. This prediction is supported by Fahlenbrach's (2009) finding that performance-related pay and good corporate governance are substitutes.

Our paper contributes to a growing literature that applies loss aversion successfully to problems 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. Coval and Shumway (2005) use loss aversion to explain patterns of intraday risk-taking in a field study of CBOT-traders. Genesove and Mayer (2001) find that sellers in the housing market exhibit loss aversion. Kouwenberg and Ziemba (2007) study the incentives and investment decisions of hedge-fund managers, and Loughran and Ritter (2002) and Ljungqvist and Wilhelm (2005) base their measure of issuer satisfaction in initial public offerings on loss aversion. Massa and Simonov's (2005) study of individual investor behavior is the only study in finance we found that fails to support loss aversion.

The only paper so far that rigorously applies loss aversion to principal-agent theory is de Meza and Webb (2007). They make two important contributions. First, they show that some part of the optimal compensation contract should be insensitive to firm performance and can therefore be represented by options. Second, they endogenize the reference wage by relating it, among others, to the median of the wage distribution. We build on their first contribution and show that optimal loss-aversion contracts can be implemented best with combinations of stock and options. We also find flat parts of the compensation contract, which become relevant if the reference wage is close to the median of the wage distribution. However, then the overall ability of the model to approximate observed contracts is low. To the best of our knowledge, ours is the first paper that explores empirically the potential of loss aversion to explain compensation contracts.

In the following Section I we develop the model and discuss the main assumptions. Section II explains our calibration method and the construction of the data set. Section III contains our calibration results for piecewise linear contracts that consist of base salary, stock and options. In Section IV, we theoretically derive and calibrate the general (nonlinear) shape of the optimal contract. Section V performs some robustness checks. Section VI develops the cross-sectional predictions of the model and indicates some avenues for future research. All proofs and derivations are deferred to the appendix.

I. 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 $e \in [0, \infty)$ and the state of nature u.

For our calibrations, we assume - as does most of the literature - that the end-of-period value is lognormally distributed with

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

where r_f is the risk-free rate of interest, σ^2 is the variance of the returns on the stock, u is a standard normal random variate, and $P_0(e)$ is a strictly increasing and concave function.[4]

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

$$V(w(P_T)) - C(e), \tag{2}$$

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.[5] Following Tversky and Kahneman (1992), we assume preferences over wage income, $w(P_T)$, of the form

$$V(w(P_T)) = \begin{cases} \left(w(P_T) - w^R\right)^\alpha & if \quad w(P_T) \geq w^R \\ -\lambda\left(w^R - w(P_T)\right)^\beta & if \quad w(P_T) < w^R \end{cases}, \text{ where } 0 < \alpha, \beta < 1 \text{ and } \lambda \geq 1. \tag{3}$$

Here, w^R denotes the reference wage.[6] 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

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

[5] To the best of our knowledge, Edmans, Gabaix, and Landier (2008) is the only paper in the compensation literature where the link between C and V is multiplicative rather than additive. This feature of their model seems to be critical for their calibrations of the optimal *level* of incentives. However, we are interested only in the structure, but not in the level of incentives.

[6] This preference specification was introduced into the finance literature by Benartzi and Thaler (1995) and was also used by Langer and Weber (2001), Berkelaar, Kouwenberg, and Post (2004), and Barberis and Huang (2008). De Meza and Webb (2007) use a generalized version of loss-aversion.

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. Formally, this introduces a kink in the value function at w^R and thus locally infinite risk-aversion. 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 this paper, we will refer to a CEO with preferences of the form (3) as *loss averse* and to the corresponding principal agent-model as the loss-aversion model or, for brevity, as the LA-model.[7]

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 $w(P_T)$, 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, $C(e)$ 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 $E[V(w(P_T))] - C(e) \geq \underline{V}$. We assume that the principal cannot pay a wage below some lower bound $\underline{w}$ on the wage function such that $\underline{w} \leq w(P_T)$ 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

[7]Strictly speaking, the term loss aversion refers to the fact that the agent is more averse to losses than he or she is attracted to gains, hence only to the first characteristic mentioned in the text. For brevity, we use loss aversion more comprehensively to refer to a model that comprises all three characteristics.

the period. This makes $\underline{w} = -W_0$ a natural choice, but higher values of $\underline{w}$ may also be plausible.

Optimization problem. The shareholders' problem can be written as (see Holmström (1979)):

$$\max_{e,\ w(P_T)\geq \underline{w}} E\left[P_T - w\left(P_T\right)|e\right] \tag{4}$$

$$s.t.\ E\left[V\left(w\left(P_T\right)\right)|e\right] \geq \underline{V} + C\left(e\right) \quad , \tag{5}$$

$$e \in \underset{\widetilde{e}\in[0,\infty)}{\operatorname{argmax}} \left\{E\left[V\left(w\left(P_T\right)\right)|\widetilde{e}\right] - C(\widetilde{e})\right\} \ , \tag{6}$$

We replace the incentive compatibility constraint (6) by the first order condition

$$\frac{d}{de} E\left[V\left(w\left(P_T\right)\right)|e\right] - C'(e) = 0. \tag{7}$$

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. This requires that the following condition holds for all effort levels, so that the manager's problem is globally concave:

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

Condition (8) can be used for the loss-aversion as well as for the risk-aversion model. It will not hold generally, because the function $V\left(w\left(P_T\right)\right)$ is convex in the loss space and because the optimal contract $w\left(P_T\right)$ may be convex. However, we can ensure that condition (8) holds for some cost functions C and some density functions in two ways. First, equation (8) 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 (8) holds. Second, 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 (8) will also be satisfied. We assume in the following that at least one of these conditions applies.

Benchmark. The standard implementation in the literature on executive compensation features preferences with constant relative risk aversion (see Footnote 1 in the Introduction):

$$V^{CRRA}\left(w\left(P_T\right)\right)=\frac{\left(W_0+w\left(P_T\right)\right)^{1-\gamma}}{1-\gamma}\ , \tag{9}$$

where W_0 denotes wealth and γ represents the coefficient of relative risk aversion. We use this parameterization as our benchmark risk-aversion model (RA-model). Some papers also use constant absolute risk aversion. We find very similar results if we compute our main calibration results with CARA-preferences instead of CRRA-preferences. These results can be found in Tables B.I and B.II in the Internet Appendix.[8]

II. Implementation and data

A. Implementation

This section describes our numerical calibration method. The method allows us to derive optimal contracts for an individual CEO without specifying the cost function $C(e)$ or the production function $P_0(e)$.

Calibration approach. Our strategy for finding optimal contracts adopts the approach pioneered by Grossman and Hart (1983), who divide the solution to the optimal contracting problem into two stages. The first stage solves for the optimal contract by minimizing the expected compensation costs 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 benefits of a particular effort level (here: the value of the firm) with the implementation costs determined at the first stage. We only consider the first stage, i.e. we search for the cheapest contract for a given level of effort.

We work with the null hypothesis that the observed contract is equal to the optimal contract from our model. Effectively, this is a joint hypothesis: it states that our model is the correct model and that shareholders have indeed implemented the optimal contract. One implication of this null hypothesis is that the (unobservable) effort level that is induced by the observed contract is the optimal effort

[8] The Internet Appendix for this article is available online in the "Supplements and Datasets" section at http://www.afajof.org/supplements.asp.

level. In our calibrations, we therefore search for the cheapest contract that implements the effort level induced by the observed contract. If the null hypothesis is correct, the result of this optimization will be equal (or close to) the observed contract. However, if the contract found in the calibrations differs markedly from the observed contract, we can reject the null hypothesis. Then the effort induced by the observed contract can be implemented in a less expensive way.

We now show that this reduced problem can be solved without knowledge of the cost function $C(e)$ or the production function $P_0(e)$. We first rewrite the first order condition (7) as:

$$PPS\left(w\left(P_T\right)\right) = \frac{dC(e)}{de} / \frac{dP_0(e)}{de}, \tag{10}$$

$$\text{where: } PPS\left(w\left(P_T\right)\right) = \frac{d}{dP_0} E\left[V\left(w\left(P_T\right)\right)\right]. \tag{11}$$

Here, PPS denotes the pay-for-performance sensitivity of the contract, which is adjusted for the preferences of the CEO and reduces to the standard (risk-neutral) pay-for-performance sensitivity for $V\left(w\left(P_T\right)\right) = w\left(P_T\right)$.

Let $w^d\left(P_T\right)$ be the observed contract. Here and in the following we use the superscript 'd' in order to refer to observed values or 'data.' Our null hypothesis is that $w^d\left(P_T\right)$ is an optimal contract, which satisfies condition (10). We can therefore rewrite the first-order condition (10) again as:

$$PPS\left(w\left(P_T\right)\right) = PPS(w^d\left(P_T\right)). \tag{12}$$

Using the same reasoning, we can also rewrite the participation constraint (5) as:

$$E\left[V\left(w\left(P_T\right)\right)\right] \geq E\left[V\left(w^d\left(P_T\right)\right)\right]. \tag{13}$$

Finally, since we are solving only the first stage of Grossman and Hart (1983), the objective function for the numerical program is:

$$\min_{w(P_T)} \quad E\left[w(P_T)\right]. \tag{14}$$

Observe that both constraints (12) and (13) do not depend on the cost function $C(e)$ or the production function $P_0(e)$. We assume rational expectations and hence the observed market capitalization equals $P_0(\widehat{e})$, where $\widehat{e}$ is the chosen effort level based on the observed contract, which are both anticipated by the market.

Finding optimal contracts as a solution to program (12) to (14) is intuitive. The program searches for a contract $w(P_T)$ (i) that provides the same incentives as the observed contract from (12); it therefore implements the same level of effort and generates the same value of the firm; (ii) that provides at least the same utility to the CEO as the observed contract from (13); (iii) that minimizes the costs to shareholders. If the model is a good approximation to the pay-setting process, then the optimal contract $w(P_T)$ generated by the model should be close (in a sense to be defined) to the observed contract $w^d(P_T)$.

Note that shareholders' objective to reduce the CEO's rents never plays a role in our analysis. Any rent the agent receives in the observed contract (possibly due to limited liability, rigid base salaries, or managerial power) is preserved in our calibrations, because the participation constraint (13) ensures that the agent's utility from the optimal contract is never lower than her utility from the observed contract.[9]

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.[10]

Piecewise linear contracts. The program (12) to (14) is mathematically well-defined but suitable for numerical calibrations only if we restrict the shape of the contract $w(P_T)$ so that the wage function can be represented with a small number of parameters. In the next section, we work with the restriction that the optimal contract is piecewise linear and consists of fixed salary (ϕ), stock (n_S), and options (n_O):

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

The numerical program (12) to (14) then searches for the cheapest contract across the parameters ϕ, n_S, and n_O, that satisfies the participation and incentive compatibility constraints. The remaining

[9] Rents may also arise in the loss-aversion model in equilibrium. Our preference specification (3) implies that the agent's utility is bounded from below, so that she may obtain rents in equilibrium (see Proposition 2 and Assumption A1 in Grossman and Hart (1983)).

[10] 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 (2008). For experimental studies on the preference parameters which yield parameter values in a comparable range see Abdellaoui (2000) and Abdellaoui, Vossmann and Weber (2005).

parameters, the strike price K and the maturity T of the option grant, are estimated from the data as described in the next section.

B. Data

We identify all CEOs in the ExecuComp database who are CEO for the entire fiscal years 2004 and 2005. We delete all CEOs who were executives in more than one company in either 2004 or 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.[11]

Option portfolios. We estimate the option portfolio held by the CEO from 2004 data using the procedure proposed by Core and Guay (2002). The resulting option portfolio typically consists of several option grants with different strike prices and different maturities. As our model has only one period and consequently allows only one maturity, we map this option portfolio into one representative option: We first set 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 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.

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 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

[11] 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.

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 an upper bound for the reference wage.[12] 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:

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

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

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 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. Below we discuss a robustness check, where we 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 their 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

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

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 I here]

Table I, Panel A 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.[13] 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.[14] We therefore base our inference on medians and Wilcoxon tests that are robust to skewness.

Table I, Panel B shows the corresponding statistics for the full sample of all executives who are the CEO of a firm in the ExecuComp universe in 2004 and 2005. We lose 46% of these observations due to our data requirements - mainly because of the required history of at least 5 years in the database. Compared to the full sample, the CEOs in our sample own significantly more options (1.44% compared to 1.25%) and the firms in our sample exhibit lower stock return volatility (43% compared to 49%). As a robustness check, we therefore analyze quintiles formed on the basis of option holdings n_O and (separately) of volatility σ in order to investigate whether this bias affects our results.

III. Optimal contracts with stock and options

We now numerically solve the program (12) to (14) for linear contracts (15) that consist of fixed salary ϕ, stock n_S, and options n_O. In order to ensure that end-of-period wealth is non-negative (limited liability), we need two additional constraints: First, we assume that the base salary is limited by the manager's non-firm wealth ($\phi > -W_0$). Second, while we allow option awards to

[13] 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.

[14] For the value of the contract, the mean even exceeds the 90^{th} percentile. This extreme skewness is due to two outliers (Warren Buffett and Steven Ballmer) whose contract values both exceed \$10bn. If we exclude them from our sample, the average value of contract drops to \$85m, i.e. far below the 90^{th} percentile, and average wealth drops to \$30m. The remaining averages in Table I do not change substantially. We report a robustness check in Section V, which shows that our results are not affected by these outliers.

become negative (i.e. managers can be required to write options), the manager's short position in options is restricted not to exceed her stock holdings n_S, so $n_O > -n_S$.[15] This restricts the wage function to be non-decreasing. We could alternatively assume that option holdings are non-negative, but the fact that in practice (to the best of our knowledge) managers never write options should be a *result* of a good model and not an assumption. We report on robustness checks with non-negative option holdings below.

For each CEO, we want to compare the observed contract with the optimal piecewise linear contract for the LA-model and for the RA-model. For this we define a metric to evaluate to what extent each model predicts the observed composition of the contract between stock and options:

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

where σ_S and σ_O are the cross-sectional standard deviations of observed stock holdings n_S^d and observed option holdings n_O^d, respectively. D measures the distance between the observed contract and the model contract and gives more weight to the parameter that has less cross-sectional dispersion. D is an ad hoc measure that is meaningful only to compare different models; it is not the objective function that is minimized in our estimation. A similar approach is used in Carpenter (1998) and Bettis, Bizjak, and Lemmon (2005).[16]

D does not take into account fixed salaries. We have also developed a metric similar to D that also includes base salaries and obtain similar results (not tabulated).[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 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 (12) ensures that deviations from the observed value for options result in deviations for stock and vice versa, so that the scaling and

[15] This is the constraint if the dividend yield d equals zero. With dividends it becomes $n_O > -n_S \exp(dT)$.

[16] 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.

[17] Formally, deviations in the fixed salary are determined by the deviations in stock and option holdings as the participation constraint (13) always binds. If $n_S = n_S^d$ and $n_O = n_O^d$, then $\phi = \phi^d$ must hold. Therefore, including salary in the metric (17) does not increase the metric's information content.

weighting of any single parameter relative to the other is largely inconsequential.

[Insert Table II here]

Table II, Panel A summarizes the results for the LA-Model for seven different levels of the reference wage as parameterized by θ (see equation (16)). Panel B shows the results for the RA-model for seven values of the coefficient of relative risk-aversion γ.[18] For each model we show the 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)).

The performance of the LA-model is 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 errors are below 0.5 in absolute value for $\theta = 0.1$ and $\theta = 0.2$ for all three contract parameters and the distance metric D is below 1. 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 ($\theta > 0.4$). Then it predicts negative median option holdings and negative median base salaries, with scaled deviations and D-values in excess of 1. According to this metric, we achieve the best fit for the model for $\theta = 0.1$.

The RA-model (see Table II, Panel B) always predicts negative base salaries and negative option holdings, so optimal RA-contracts are concave. The performance of the RA-model as measured by

[18] We report results for 11 values of the reference wage in the Internet Appendix, Table A.II. We do not consider values of γ below 0.1 in Table II 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 her own firm) equal to $1/\gamma$. E.g., $\gamma = 0.1$, the lowest value considered in Table II, 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.

the distance metric D is generally worse than the performance of any parameterization of the LA-model. The RA-model works best if risk aversion is either very low or very high. High risk aversion reduces the incentives from options more than those from stock, so optimal RA-contracts feature fewer additional shares to replace the existing options compared to lower levels of risk aversion. This reduces the gap between optimal and observed contracts. However, the optimal contract always remains concave, even for levels of risk aversion higher than shown in the table. On the other end of the spectrum, if risk-aversion decreases and converges to zero, *any* observed contract is optimal (i.e. cost minimizing), because subjective values are then identical to market values and all contracts that generate the same incentives are equally costly. It is only for these very low levels of risk aversion ($\gamma = 0.1$, 0.2) that the RA-model is more accurate than some LA-specifications with a high reference wage.

An important limitation of the analysis in Table II 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 or, 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. As a next step, we therefore compare the two models based on parameterizations that hold the overall attitude to risk constant in a meaningful way. In particular, we 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

$$CE^{LA}(w^d, \theta) \equiv CE^{RA}(w^d, \gamma_e) \ . \tag{18}$$

We refer to the value of γ_e that satisfies (18) as the equivalent degree of relative risk aversion, because it holds the certainty equivalent constant. A straightforward implication of this step is that we also hold the risk premium paid by shareholders, $E(w^d) - CE(w^d)$, 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 III compares the two models in this way.

[Insert Table III here]

Table III reports the mean and the median difference $D^{RA} - D^{LA}$ of the distance metric D between the two models. The verdict based on the mean and median of D is clear and independent of the overall attitude to risk: The LA-model dominates the RA-model for the entire range of reference wages. However, the RA-model fits the data better than the LA-model according to D for a small number of observations (3% - 18% of the sample), some of which generate extreme deviations for the LA-model. The equivalent γ_e's are generally very low and below the range we regard as plausible (see Footnote 18). They are also non-monotone 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, so that overall risk aversion (which is captured by γ_e) becomes smaller.

Table III also 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 99% of the CEOs in our sample have such a contract. In sum, 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 standard RA-model is clearly inferior on this dimension.

IV. Optimal nonlinear contracts

Our calibrations show that a suitably parameterized loss-aversion model can explain observed compensation practice better than the standard risk-aversion model. So far, however, our analysis does not provide an intuition *why* this is the case. In this section we derive and analyze the shape of the optimal unrestricted contract. This analysis provides us with a better understanding of when and why the LA-model dominates the RA-model for piecewise linear contracts.

A. The shape of the optimal contract

In this section we drop the restriction that $w\left(P_T\right)$ is piecewise linear. The following proposition describes the shape of the optimal nonlinear contract.

PROPOSITION **1.** *Under the assumptions (i) that the agent's preferences are as described in (2) and (3), (ii) that the stock price P_T is lognormally distributed as described in (1), and (iii) that the condition for the first-order approach (8) holds, the optimal contract $w^*\left(P_T\right)$ for the principal agent problem (4) to (6), is:*

$$w^*\left(P_T\right) = \begin{cases} w^R + \left(\gamma_0 + \gamma_1 \ln P_T\right)^{\frac{1}{1-\alpha}} & if \quad P_T > \widehat{P} \\ \underline{w} & if \quad P_T \leq \widehat{P} \end{cases}, \tag{19}$$

where $\widehat{P}$ is a uniquely defined, strictly positive cut-off value and γ_0 and γ_1 are two parameters that depend on the parameters of the problem and the Lagrange multipliers associated with the two constraints.

The details of the proof and an implicit definition of $\widehat{P}$ are deferred to Appendix A. In the proof we derive a more general result for more general distributional assumptions. We state only the result for the lognormal distribution here because this is what we need for our calibrations. 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 $w^*\left(P_T\right)$ cannot lie strictly between $\underline{w}$ and w^R. The reason is that the agent is risk loving in the loss space, so we can improve on any payment in the loss space 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, because 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 $w\left(P_T\right)$, which gives (19).

Equation (19) shows that for the gain space, where $P_T > \widehat{P}$, we obtain a result very similar to the familiar Holmström condition (Holmström, 1979, equation (7)) for optimal contracts in the standard concave utility model (see also equation (A.10) in the Appendix). This is intuitive, since the problem in the gain space, where preferences are concave, is structurally similar to a standard utility-maximizing framework.

COROLLARY **1.** *In the optimal contract (19), the size of the loss space decreases with loss aversion:* $\frac{d\widehat{P}}{d\lambda} < 0$*. Moreover, there exists an inflection point* P_T^I *so that whenever* $P_T^I \geq \widehat{P}$*, the optimal contract is convex for all terminal stock prices* $P_T \in \left[\widehat{P}, P_T^I\right]$ *and concave above the inflection point.*

Corollary 1 states that the optimal contract will make less use of "sticks," where the manager receives only the minimum wage $\underline{w}$, if she is more loss averse.[19] The optimal nonlinear contract is convex over a certain region, but above this region it becomes concave. Whether the convex region exists and whether it is relevant is an empirical question, which we address in the following using our calibration method from Section II.

B. Calibration of the optimal contract

Proposition 1 provides us with the functional form of the general optimal contract that only depends on the two quantities γ_0 and γ_1. Once these two quantities are fixed, the cut-off value $\widehat{P}$ can be calculated from equation (A.17) in the Appendix. We can then find the solution to the program in equations (12) to (14) for the optimal nonlinear contract just as we did with the piecewise linear contract in Section III.[20]

Figure 1 shows the result of such a calibration for a representative CEO and two alternative reference wages ($\theta = 0.2,\ 0.8$); it also displays the corresponding RA-contract with $\gamma = 2$ and the observed contract.

[Insert Figure 1 here]

The figure illustrates that, for $P_T > \widehat{P}$, the optimal LA-contract pays off only in the gain space. Here, the contract is continuous, increasing and convex. The eventual concavity for high P_T is empirically not relevant for this CEO. If P_T drops below $\widehat{P}$, the optimal contract discontinuously drops to the lowest possible wage $\underline{w}$ that is paid out for all $P_T \leq \widehat{P}$.[21] If the reference point increases,

[19] The proof of Corollary 1 can be found in the Internet Appendix that is available online in the "Supplements and Datasets" section at http://www.afajof.org/supplements.asp.

[20] In contrast to the piecewise linear contract, the general contract depends on only two parameters. As the numerical optimization problem (12) to (14) features two constraints that are binding at the optimum, these two constraints uniquely define the optimal contract, which can therefore be calculated as the solution to a system of two equations (12) and (13) in the two unknowns γ_0 and γ_1.

[21] 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 Tversky and Kahneman (1992) value function (3), 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.

the threshold $\widehat{P}$ increases and the contract becomes flatter for $P_T > \widehat{P}$. Then more incentives are provided by the discontinuous drop (which is reminiscent of firing the manager) compared to the slope for stock prices above $\widehat{P}$. If the reference wage is small, the discontinuity is less important and the contract shape is dominated by its convexity for $P_T > \widehat{P}$ that approximates the observed contract very well. (In fact, the LA-contract for the even lower reference wage $\theta = 0.1$ fits the observed contract best. We do not show this in Figure 1 for visual clarity.) In contrast to the LA-contract, the RA-contract is always concave for a coefficient of relative risk aversion greater than 1 and is thus strikingly at odds with observed contracts for conventional levels of risk aversion (see also Dittmann and Maug (2007)).

A prominent feature of the optimal contract (19) is that it involves disciplining dismissals, i.e. minimum payouts for all realizations of the stock price P_T below the threshold $\widehat{P}$. However, our model does not fully capture the true nature of dismissals. What we capture is the loss of income and human capital of the dismissed CEO. A second, and arguably more important effect is that the old CEO is replaced with a new CEO who is expected to perform better in the future. This second effect cannot be modeled in a one-period set-up and is therefore entirely absent from our model. Also, firm performance explains only a fraction of the variation in CEO turnover (Kaplan (1994), Jenter and Kanaan (2006)) and most dismissed CEOs receive generous severance payments that should not be paid if dismissals were purely disciplinary. We therefore do not include dismissals when we estimate empirical contracts in Section II.B.[22] For brevity, however, we continue to use the word "dismissal" to describe the drop in compensation to the lowest feasible level for low stock price realizations. In the calibrations, we also calculate the probability of such a dismissal as

$$p(\hat{P}) \equiv \Pr(P_T \leq \widehat{P}) = \int_0^{\hat{P}} f(P_T)\, dP_T. \tag{20}$$

To put this number into perspective, we compare the mean dismissal probability generated by the model with the mean in the data. 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

[22] Even if we would like to take into account dismissals when estimating the observed contract we would hardly be able to do so for two reasons. First, only cross-sectional results are available for the probability of dismissals, so we cannot estimate them individually for each CEO from this CEO's data only. Second, we do not have any information on the loss in human capital and the severance pay the CEO expects to receive in case of a dismissal.

and 2004 and obtain an average dismissal probability of 7.4%. For the reasons given above, only some of these dismissals are likely to be disciplinary in nature, so our model should generate a much smaller probability if it fits the data well. We also report the incentives from dismissals, which is the fraction of the total incentives PPS from (11) that is due to the discontinuous drop in compensation at the point $P_T = \widehat{P}$.

Table IV displays our results for the calibrations of the general loss-aversion contract (19) to each CEO in our sample – again for different assumptions on the CEOs' reference wage. We do not tabulate the averages of the two parameters γ_0 and γ_1 as they cannot be interpreted independently from one another. Instead, the table shows the averages of the dismissal probability, the incentives from dismissal, the inflection point where the shape of the contract changes from convex to concave, and the change in CEO wealth if the stock price decreases or increases by -50%, -30%, $+30\%$, or $+50\%$. These numbers serve as an indication of the slope and the convexity of the optimal contract in the gain space.

[Insert Table IV here]

The contracts predicted by the LA-model are on average convex for both measures of convexity. The upward change of wealth for an increase in the stock price is always bigger than the corresponding downward change, e.g., for $\theta = 0.1$, the median change in compensation is 41.8% if the stock price increases by 30%, but compensation declines by only 35.9% if the stock price falls by the same amount. Also, virtually all of the probability mass for this contract lies to the left of the inflection point, rendering the concave part of the contract irrelevant. For low reference wages where $\theta \leq 0.2$, dismissals are unlikely and provide little incentives for the CEO. Here, almost all incentives are generated by the increasing and convex payoff in the gain space. This picture changes markedly for higher reference wages. The probability of dismissals rises quickly with θ to as much as 20% for $\theta = 1$ while the payoff in the gain space becomes flatter, e.g., for a stock price increase of 30% the median increase in compensation falls from 41.8% ($\theta = 0.1$) to 7.5% ($\theta = 1$). So, dismissals provide most of the incentives for high reference wages. Intuitively, CEOs with a higher reference wage demand more 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 and to a larger extent through the threat of dismissals. Given that the empirical dismissal probability is 7.4% and that a (potentially large) fraction of these empirical dismissals are not disciplinary in nature, the model fits the data best for low levels of the reference wage.

C. Why can the loss-aversion model explain convex contracts?

The LA-contract has two prominent features: (1) the discontinuous drop if the stock price P_T falls below the threshold $\widehat{P}$ and (2) the convexity for stock prices above this threshold. It is the second feature that makes the LA-model a good approximation for small reference wages in Figure 1. By contrast, the discontinuous drop does not improve the fit and is the reason why the LA-model is less successful for higher reference wages.

An intuition for the discontinuous drop can be developed rather easily: The agent is risk-loving in the loss space, so any payout within the loss space is inefficient; the agent prefers a fair lottery over $\underline{w}$ and some point above the reference wage w^R to any payout within the loss space. The incentive effect of these extreme payouts is then maximized if all the probability mass where $\underline{w}$ is paid out is concentrated at the lowest stock price realizations.

What is less apparent is that the LA-contract is convex in the gain space even though the preference specification for gains (equation (3)) is the same as the CRRA utility function (equation (9)) for $\gamma = 1 - \alpha = 0.12$. Still, the RA-contract is concave even for risk aversion as low as $\gamma = 0.1$ (see Table II, Panel B). The difference between the two functions is that the LA-model evaluates the outcome relative to the reference wage w^R, whereas the RA-model evaluates the absolute outcome, i.e. the outcome relative to the worst case $-W_0$.

There are two effects that determine the curvature of the optimal contract in both models. The first is the incentive effect that it is optimal to provide incentives where marginal utility $V'(w)$ (or, in the LA-model, marginal value) is high. Marginal utility is decreasing in the wage w in the RA-model and in the gain space of the LA-model, so without any offsetting effect the slope of both contracts should be higher for low stock prices than for high stock prices, i.e. both contracts should be concave. The countervailing effect is risk tolerance $RT(w) = -V'(w)/V''(w)$, which is the inverse of absolute risk aversion in the RA-model and similarly defined in the LA-model. It is cheaper for the principal to provide incentives in regions where risk tolerance is high, because the risk premium the agent requires is then lower. Risk tolerance increases with the wage w in both models: as the payoffs increase the agent becomes less averse to risk. Hence if risk tolerance were the only relevant effect, incentives should be provided at high stock prices and the contract would be convex in both models.[23] In Appendix B we show that the slope of the optimal contract for the RA-model and for the gain space of the LA-model can be written as:

$$\frac{d\,w\,(P_T)}{d\,P_T} = RT\,(w)\,V'\,(w)\,\frac{\alpha\gamma_0}{P_T}. \tag{21}$$

Intuitively, the optimal contract is convex (i.e. the slope (21) increases with P_T) only if risk tolerance $RT\,(w)$ increases faster than marginal utility $V'\,(w)$ (divided by P_T) declines.[24]

The strength of the risk tolerance effect differs greatly between the two models, which can be observed from inspecting the expressions for risk tolerance directly:

$$RT^{LA}\,(w) = \frac{w\,(P_T) - w^R}{1-\alpha}\ ; \qquad RT^{RA}\,(w) = \frac{w\,(P_T) + W_0}{\gamma}. \tag{22}$$

For the LA-model, risk-tolerance is close to zero near the reference point and then increases quickly. This reflects the fact that in the LA-model, risk-aversion is concentrated around the reference point and becomes negligible far away from the reference point. Likewise, in the RA-model risk-tolerance is close to zero near the minimum payout $-W_0$, but relevant and likely payouts are far away from this point, and there risk-tolerance is high and does not vary much with compensation and therefore with the stock price. As a consequence, the risk-tolerance effect is much stronger in the LA-model than in the RA-model. This result cannot be shown theoretically.[25] It is a strength of our calibration method that it allows us to establish this result empirically.

De Meza and Webb (2007) can also rationalize the presence of options in loss-aversion contracts, but their argument is different from ours. They obtain option-type payoffs if they assume that the agent is risk-averse in the loss space (see in particular their Figure 1(b) and their Proposition 1 (iii)), which differs from the Kahneman and Tversky (1979, 1992) framework employed here. By contrast, we obtain option-type contracts from the convexity of the optimal contract in the gain space for the reasons explained above.

[23] If the manager has constant absolute risk aversion, then risk tolerance RT is constant across stock prices and the risk-tolerance effect does not induce any convexity.

[24] This argument is not limited to our context and extends to any utility (or value) function and can easily be extended to other distributional assumptions.

[25] For the RA-model, it can be shown that the incentive effect is always stronger than the risk tolerance effect if $\gamma > 1$, i.e. if the CEO is more risk averse than implied by log utility.

D. Comparison of linear and nonlinear contracts

With these insights about the optimal contract in the loss-aversion model, we can now develop an intuition for the results for the piecewise linear contract in Section III. The piecewise linear contract is an approximation to the general contract shown in Figure 1. A contract with stock and one option grant can only capture *either* the discontinuous drop in compensation for low payouts *or* the convexity for high payouts. For low reference wages most incentives come from the increasing convex payout for higher stock prices. This shape is best approximated by a contract with stock and options, which we find for low reference wages. For higher reference wages the discontinuous drop becomes more important. This discontinuity can best be approximated by a contract with low and even negative base salaries, large stock holdings, and negative option holdings, which we find for higher reference wages.

In principle, the optimal nonlinear contract (19) 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 by fixed salary, stock and one option grant. In Table V we therefore compare the optimal nonlinear contract (19) with the optimal piecewise linear contract (15).[26] For both contracts, the table shows the median change in wealth if the stock price increases or decreases by 30%. In addition, the table 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

$$Savings = \frac{E\left(w^{d}\left(P_{T}\right)\right) - E\left(w^{*}\left(P_{T}\right)\right)}{E\left(w^{d}\left(P_{T}\right)\right)} \ , \tag{23}$$

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. Note that by construction, the savings of the optimal general contract must be higher than the savings of the piecewise linear contract.

[26] We report more detailed results for 11 values of the reference wage in the Internet Appendix, Table A.V.

[Insert Table V here]

The table shows that the linear option contract is always steeper than the general nonlinear contract because the linear contract does not feature dismissals, so that all incentives are provided through the wage function. The difference in slopes increases with the reference wage as the dismissal probability for the nonlinear contract increases. As the reference wage increases, the linear contract becomes less convex and becomes concave for $\theta \geq 0.6$. In contrast, the nonlinear contract is convex for all levels of θ.

The savings are not substantial for either version of the contract. This is important, because it shows that even where the distance between observed contracts and predicted contracts appears large in terms of the metric 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: For $\theta = 0$, it is 0.3% (= 0.5% − 0.2%) of total compensation costs, or \$0.09 million for the median CEO with a pay package worth \$29.8 million. Even if $\theta = 1$, this difference is only 9.4% (15.0% − 5.6%) or \$2.80 million, which is about 0.12% of the value of the median company. These savings have to be related to the governance costs of writing and enforcing such a general contract. We suspect that for most companies, the benefits of incentive provision through CEO dismissals rather than through high-powered wage functions will be negligible.

V. Robustness checks

We perform seven robustness checks. To conserve space, we only include the results for the first check here, all other tables are contained in the Internet Appendix (see http://www.afajof.org/supplements.asp).

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 these parameters.

[Insert Table VI here]

Table VI reports the results of a comparative static analysis for the linear model 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. $\theta = 0.1$). From the metric D 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 in the discussion of Table II. 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 VI therefore underline 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). In particular, higher loss aversion (λ) reduces the incidence of dismissals, which in turn makes for better approximation to observed contracts (see Corollary 1). 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 86% of the probability mass lies for the base scenario in Table VI (see Table II 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 second robustness check we try to identify those observations where the LA-model performs 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. See Table B.III in the Internet Appendix for complete results.

According to our metric D, the LA-model performs much worse for the owner-managers than for the non-owner managers. For $\theta = 0.1$, the median distance is 0.13 for non-owner managers compared to 2.24 for owner-managers. Still, the LA-model performs better than the RA-model in both subsamples. Closer inspection of the data shows that these results are driven by those owner-manager CEOs who have no options (one example in our data set is Warren Buffett). We conclude from this discussion 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.[27]

Restrict salaries and option holdings to be non-negative. 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 $n_O \geq 0$, which also rules out concave contracts. The results in Table B.IV in the Internet Appendix show that ruling out concave contracts and negative base salaries improves the performance of the RA-model significantly. However, the RA-model is still not able to generate positive salaries and positive option holdings simultaneously and one of the two new constraints always binds. 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.

Remove outliers. We remove two extreme observations from our data set (Warren Buffett and Steven Ballmer) who have a contract value that exceeds \$10 billion and causes the mean contract value in Table I to exceed the 90% quantile. When we recompute the main Tables III and IV, the numerical changes in the results are miniscule. See Tables B.V to B.VII in the Internet Appendix for complete results. We can safely conclude that these two outliers do not affect our results.

Biases in our sample. The firms in our sample have a significantly lower stock return volatility and grant significantly more options to their CEOs compared to the full ExecuComp CEO sample (see Table I). We therefore split our data set into quintiles, once according to volatility σ, and once according to option holdings n_O^d. We then recompute Tables III and IV separately for the quintiles (see Tables B.VIII to B.XI in the Internet Appendix) and find that the fit of the LA-model is clearly worse for high levels of volatility and somewhat worse for high levels of option holdings. Nevertheless, the LA-model continues to dominate the RA-model in more than 94% of all cases. The bias in our sample towards higher option holdings therefore works against the LA-model, whereas the bias towards lower volatility makes the LA-model look slightly better. Our qualitative conclusions are not affected by these biases.

[27] 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.

Analysis for 1997. We repeat our analysis for 1997, the first year for which we can calculate our wealth measure from ExecuComp. The results are shown in Tables B.XII and B.XIII in the Internet Appendix. In 1997, option holdings were much lower than in the 2005 sample analyzed above. Also, options were less in the money, and the stock price volatility was much lower than in 2005. Consistent with the comparative statics findings from the previous robustness check, we find that the LA-model performs better for 1997 than for 2005.

Wealth robustness check. Our 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 recalculate Table III twice: once after reducing the estimate of wealth by 50% and once after increasing it by 100%. The results are shown in Table B.XIV in the Internet Appendix. We find that none of our results changes markedly. Wealth plays a much more prominent role in the RA-model than in the LA-model. For an agent with CRRA utility, an increase in wealth is similar to a decrease in the risk-aversion parameter γ. Correspondingly, the equivalent γ from (18) increases with increasing wealth and offsets this effect. Overall, none of our results seems to be affected by measurement errors of CEO wealth.

VI. Conclusion

We analyze stylized contracts to compensate CEOs that consist of stock, options, and fixed compensation in a model where the CEO is loss-averse. We show that the loss-aversion model generates optimal contracts that are similar to observed contracts for a range of parameterizations where the reference wage of the CEO is low. By contrast, for higher reference wages the model generates contracts that are concave and different from contracts observed in practice. We benchmark our results against those for the standard risk-aversion model that is conventionally used for calibrations in the compensation literature and find that the loss-aversion model generates better approximations for the large majority of CEOs in our sample, but generates poor approximations for owner-managers with large stock holdings.

The loss-aversion model generates two important cross-sectional findings that link CEO characteristics to features of their compensation contracts or their employers (see Graham, Harvey, and

Puri (2008) for such evidence). First, the more averse the agent is to losses, the more options are used in the optimal contract, because the risk-tolerance effect increases with the extent of loss-aversion. This is in stark contrast to the risk-aversion model, where a more risk-averse agent receives *fewer* options.[28] Second, higher reference wages lead to a higher dismissal probability and to a lower pay-for-performance sensitivity (PPS) on the job. So any variation in reference wages across CEOs should result in a negative correlation between on-the-job PPS and the dismissal probability. As performance-related dismissals require good corporate governance mechanisms, this finding is in line with Fahlenbrach's (2009) result that good corporate governance and high PPS are substitutes. The model then also suggests that firms with good corporate governance are more attractive for some CEOs (those with high reference wage) than for others or, inversely, that corporate governance is influenced by CEO characteristics.

We conclude that the loss-aversion model is a promising candidate for analyzing executive compensation contracts. However, the results are sensitive to the assumptions on the reference wage. Little research is available on the way individuals establish reference points, and we are not aware of any research that addresses reference wages in a compensation context. Also, our application infers parameter values from the experimental literature. This literature mostly reports results on non-strategic experiments of individual decision-making. Experiments of strategic contexts like ours would enhance our understanding of how individuals set reference points. More research in this direction is needed before a final verdict on the suitability of the loss-aversion model can be reached. In the meantime, the loss-aversion model - parametrized with a low reference wage - may still serve as a suitable workhorse to address normative questions. Its main advantage is that it avoids asking questions on option design from a model that has difficulty accommodating the existence of options.

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 corporation extends to its officers, largely because we do not have data on these items. These

[28] In the simple risk-aversion model that we consider in this paper, the CEO never receives any (long position) in options. However, there are a few extensions of this model (e.g., models with sticky salaries) that can explain some option holdings (see footnote 2). In these models, option holdings decrease in the agent's risk-aversion.

compensation items are not related to stock price performance, so they only bias our estimate of fixed compensation downward. Another unrealistic aspect (not only of the LA-model but also of the RA-model) is the fact that they are both static, whereas shareholders and CEOs typically revise their contracts repeatedly over a number of periods. The development of tractable dynamic models will be an important task for future research.

Appendix

A. 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. In 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. The agent is risk-averse in the gain space. Lotteries over payouts in the gain space are therefore never optimal. Lemma 2 shows that the optimal contract never features lotteries and 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 set up the Lagrangian for the simplified problem, derive the first-order condition, and show that the first order conditions are sufficient. To conserve space, we defer the more technical proofs of Lemma 1 and Lemma 2 to the Internet Appendix that is available online in the "Supplements and Datasets" section at http://www.afajof.org/supplements.asp.

Throughout we assume that the principal wishes to implement some given effort level $\widehat{e}$. We develop the argument first for a generic setup because the argument is then more compact. Hence, we assume that $f(P_T|\widehat{e})$ is a continuous density that satisfies the monotone likelihood ratio property and later assume that f is the lognormal.

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 the reference wage w^R with probability l and the minimum wage $\underline{w}$ with the remaining probability $1-l$.*

See the Internet Appendix for a proof of Lemma 1.

LEMMA 2. ***(Shape of the loss space):*** *There exists a 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: $w^*(P_T|P_T \leq \widehat{P}) = \underline{w}$.*

See the Internet Appendix for a proof of Lemma 2.

The optimal contract in the gain space. Lemma 2 allows us to rewrite the principal's program (4), (5), and (7) as follows:

$$\min_{\widehat{P},w(P_T)\geq w^R} \int_{\widehat{P}}^{\infty} w\left(P_T\right) f(P_T|\widehat{e})dP_T + \underline{w}F(\widehat{P}|\widehat{e}) \tag{A.1}$$

$$s.t. \int_{\widehat{P}}^{\infty} V\left(w\left(P_T\right)\right) f(P_T|\widehat{e})dP_T + V(\underline{w})F(\widehat{P}|\widehat{e}) \geq \underline{V} + C\left(\widehat{e}\right) \quad , \tag{A.2}$$

$$\int_{\widehat{P}}^{\infty} V\left(w\left(P_T\right)\right) f_e(P_T|\widehat{e})dP_T + V(\underline{w})F_e(\widehat{P}|\widehat{e}) \geq C' \ . \tag{A.3}$$

The derivative of the Lagrangian with respect to $w\left(P_T\right)$ at each point $P_T \geq \widehat{P}$ is:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial w(P_T)} &= f(P_T|\widehat{e}) - \mu_{PC}V'\left(w\left(P_T\right)\right) f(P_T|\widehat{e}) - \mu_{IC}V'\left(w\left(P_T\right)\right) f_e(P_T|\widehat{e}) \\ &= f(P_T|\widehat{e}) \left[1 - \mu_{PC}V'\left(w\left(P_T\right)\right) - \mu_{IC}V'\left(w\left(P_T\right)\right) LR\left(P_T\right)\right] \ . \end{aligned} \tag{A.4}$$

Setting (A.4) to zero and solving gives the optimal contract in the gain space as:

$$V'\left(w\left(P_T\right)\right) = \left[\mu_{PC} + \mu_{IC}LR\left(P_T\right)\right]^{-1}, \tag{A.5}$$

where $LR\left(P_T\right)$ denotes again the likelihood ratio.

Uniqueness and definition of $\widehat{P}$. To find $\widehat{P}$ we take the derivative of the Lagrangian with respect to $\widehat{P}$:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \widehat{P}} &= \left(\underline{w} - w(\widehat{P})\right) f(\widehat{P}|\widehat{e}) + \mu_{PC}\left(V(w(\widehat{P})) - V(\underline{w})\right) f(\widehat{P}|\widehat{e}) \\ &\quad + \mu_{IC}\left(V(w(\widehat{P})) - V(\underline{w})\right) f_e(\widehat{P}|\widehat{e}) \qquad (A.6) \\ &= -\left(V(w(\widehat{P})) - V(\underline{w})\right) f(\widehat{P}|\widehat{e}) \times \left[\frac{w(\widehat{P}) - \underline{w}}{V(w(\widehat{P})) - V(\underline{w})} - \mu_{PC} - \mu_{IC}LR\left(\widehat{P}|\widehat{e}\right)\right]. \qquad (A.7)\end{aligned}$$

This derivative of the Lagrangian is zero if the term in squared brackets in (A.7) is zero. Substituting equation (A.5) and rearranging yields:

$$\frac{\partial \mathcal{L}}{\partial \widehat{P}} = 0 \Leftrightarrow V(w(\widehat{P})) - V(\underline{w}) - V'\left(w\left(P_T\right)\right)\left(w(\widehat{P}) - \underline{w}\right) = 0. \tag{A.8}$$

The derivative of the left hand side of (A.8) with respect to P_T is:

$$-V''\left(w\left(P_T\right)\right)w'\left(P_T\right)\left(w(\widehat{P})-\underline{w}\right)>0, \tag{A.9}$$

where w' and V'' denote first and second derivatives, respectively. The expression in (A.9) is positive because $V''<0$ from the concavity of V in the gain space and $w'>0$ since the wage function (A.5) is monotonically increasing because of the monotone likelihood ratio property. Hence, $\widehat{P}$ is unique.

Finally, we note that a contract that pays out only in the loss space cannot be optimal as it does not provide any incentives, so $\widehat{P}<\infty$.

Sufficiency. In the Internet Appendix, we show that equation (19) is also a sufficient condition for the optimal contract.

Derivation of the optimal contract for the lognormal distribution. For the Tversky and Kahneman (1992) preferences (3) we can rewrite (A.5) as:

$$w\left(P_T\right)=w^R+\left[\alpha\left(\mu_{PC}+\mu_{IC}LR\left(P_T|\widehat{e}\right)\right)\right]^{\frac{1}{1-\alpha}}, \tag{A.10}$$

and equation (A.8) becomes:

$$\alpha\left(w(\widehat{P})-\underline{w}\right)-\lambda\left(w^R-\underline{w}\right)^{\beta}\left(w(\widehat{P})-w^R\right)^{1-\alpha}-\left(w(\widehat{P})-w^R\right)=0. \tag{A.11}$$

Equation (A.9) also implies that $w(\widehat{P})>w^R$.

We derive equation (19) by substituting the relevant expressions for the lognormal distribution. From (1), $\ln\left(P_T\right)$ is distributed normal with mean $m\left(e\right)=\ln\left(P_0\left(e\right)\right)+\left(r_f-\frac{\sigma^2}{2}\right)T$ and standard deviation $\sigma\sqrt{T}$. The density $f\left(P_T\left|e\right.\right)$ of the lognormal distribution is then:

$$f\left(P_T\left|\widehat{e}\right.\right)=\frac{1}{P_T\sqrt{2\pi T}\sigma}\exp\left\{-\frac{\left[\ln P_T-m\left(\widehat{e}\right)\right]^2}{2\sigma^2T}\right\}, \tag{A.12}$$

and the likelihood ratio is

$$LR\left(P_T\left|e\right.\right)=\frac{P_0'\left(\widehat{e}\right)}{P_0\left(\widehat{e}\right)}\frac{\ln P_T-m\left(\widehat{e}\right)}{\sigma^2T}. \tag{A.13}$$

We use the following definitions:

$$\gamma_1 = \alpha \mu_{IC} \frac{P_0'(\widehat{e})}{P_0(\widehat{e})\,\sigma^2 T} \ , \tag{A.14}$$

$$\gamma_0 = \alpha \left(\mu_{PC} - \mu_{IC} \frac{P_0'(e)}{P_0(e)} \frac{m(\widehat{e})}{\sigma^2 T} \right) = \alpha \mu_{PC} - \gamma_1 m(\widehat{e}) \ . \tag{A.15}$$

We now substitute (A.13) into (A.10) and use the definitions (A.14) and (A.15) to write:

$$\alpha \left(\mu_{PC} + \mu_{IC} LR\left(P_T | \widehat{e}\right) \right) = \gamma_0 + \gamma_1 \ln P_T \ . \tag{A.16}$$

Equation (19) follows immediately from equation (A.16) together with the fact that the contract pays $\underline{w}$ for $P_T < \widehat{P}$. With (A.14) and (A.15) equation (A.11) becomes:

$$\alpha \left(w^R - \underline{w} \right) = \left(\gamma_0 + \gamma_1 \ln \widehat{P} \right) \lambda \left(w^R - \underline{w} \right)^{\beta} + (1-\alpha) \left(\gamma_0 + \gamma_1 \ln \widehat{P} \right)^{\frac{1}{1-\alpha}} . \tag{A.17}$$

This equation defines the threshold $\widehat{P}$. ■

B. Derivation of equation (21)

Equation (A.5) provides the optimality condition for the gain space of the LA-model. The same expression also holds for the RA-model, where V then represents the utility function with risk aversion. Total differentiation of (A.5) yields:

$$V''(w)\,dw = -\left[\mu_{PC} + \mu_{IC} LR\left(P_T\right)\right]^{-2} \mu_{IC} LR'\left(P_T\right) dP_T = -V'(w)^2 \mu_{IC} LR'\left(P_T\right) dP_T.$$

Rearranging and using the definition of $RT(w)$ gives equation (21).

References

Abdellaoui, Mohammed, 2000, Parameter-Free Elicitation of Utility and Probability Weighting Functions, *Management Science* 46, 1497-1512.

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, 1384-1399.

Barberis, Nicholas, and Ming Huang, 2001, Mental Accounting, Loss Aversion, and Individual Stock Returns, *The Journal of Finance* 56, 1247-1292.

Barberis, Nicholas, and Ming Huang, 2008, Stocks and Lotteries: The Implications of Probability Weighting for Security Prices, *The American Economic Review* 98, 2066-2100.

Barberis, Nicholas, Ming Huang, and Tano Santos 2001, Prospect Theory and Asset Prices, *The Quarterly Journal of Economics* 116, 1-53.

Benartzi, Shlomo, and Richard H. Thaler, 1995, Myopic Loss Aversion and the Equity Premium Puzzle , *The Quarterly Journal of Economics* 110, 73-92.

———, 1999, Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments , *Management Science* 45, 364-381.

Berkelaar, Arjan, Roy Kouwenberg, and Thierry Post, 2004, Optimal Portfolio Choice Under Loss Aversion, *The Review of Economics and Statistics* 86, 973-987.

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

Carpenter, Jennifer N., 1998, The Exercise and Valuation of Executive Stock Options, *Journal of Financial Economics* 48, 127-158.

Chen, M. Keith, Venkat Lakshminarayanan, and Laurie R. Santos, 2006, How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior, *Journal of Political Economy* 114, 517-537.

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, 613-640.

Coval, Joshua D., and Tyler Shumway, 2005, Do Behavioral Biases Affect Prices?, *The Journal of Finance* 60, 1-34.

de Meza, David, and David C. Webb, 2007, Incentive Design Under Loss Aversion, *Journal of the European Economic Association* 5, 66-92.

Dittmann, Ingolf, and Ernst Maug, 2007, Lower Salaries and No Options? On the Optimal Structure of Executive Pay, *The Journal of Finance* 62, 303-343.

Dittmann, Ingolf, and Ko-Chia Yu, 2009, How Important Are Risk-Taking Incentives in Executive Compensation?, discussion paper, Erasmus University.

Edmans, Alex, Xavier Gabaix, and Augustin Landier, 2008, A Calibratable Model of Optimal CEO Incentives in Market Equilibrium, *Review of Financial Studies*, forthcoming..

Fahlenbrach, Rüdiger, 2009, Shareholder Rights, Boards, and CEO Compensation, *Review of Finance* 13, 81-113.

Feltham, Gerald A., and Martin G. H. Wu, 2001, Incentive Efficiency of Stock Versus Options, *Review of Accounting Studies* 6, 7-28.

Genesove, David, and Christopher Mayer, 2001, Loss Aversion and Seller Behavior: Evidence from the Housing Market, *The Quarterly Journal of Economics* 116, 1233-1260.

Gneezy, Uri, Arie Kapteyn, and Jan Potters, 2003, Evaluation periods and asset prices in a market experiment, *The Journal of Finance* 58, 821–838.

Gomes, Francisco J., 2005, Portfolio Choice and Trading Volume With Loss Averse Investors, *Journal of Business* 78, 675-706.

Graham, John R., Campbell R. Harvey, and Manju Puri, 2008, Managerial Attitudes and Corporate Actions, discussion paper, Duke University.

Grossman, Sanford J., and Oliver D. Hart, 1983, An Analysis of the Principal-Agent Problem, *Econometrica* 51, 7-45.

Haigh, Michael S., and John A. List, 2005, Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis, *The Journal of Finance* 49, 523-534.

Hall, Brian J., and Thomas A. Knox, 2004, Underwater Options and the Dynamics of Executive Pay-to-Performance Sensitivities, *Journal of Accounting Research* 42, 365-412.

Hall, Brian J., and Kevin J. Murphy, 2000, Optimal Exercise Prices for Executive Stock Options, *The American Economic Review* 90, 209-214.

———, 2002, Stock Options for Undiversified Executives, *Journal of Accounting and Economics* 33, 3-42.

Haubrich, Joseph G., 1994, Risk Aversion, Performance Pay, and the Principal-Agent-Problem, *Journal of Political Economy* 102, 258-275.

Haubrich, Joseph G., and Ivilina Popova, 1998, Executive Compensation: A Calibration Approach, *Economic Theory* 12, 561-581.

Hemmer, Thomas, Oliver Kim, and Robert E. Verrecchia, 1999, Introducing Convexity into Optimal Compensation Contracts, *Journal of Accounting and Economics* 28, 307-327.

Holmström, Bengt, 1979, Moral Hazard and Observability, *Bell Journal of Economics* 10, 74-91.

Huddart, Steven J., 1994, Employee Stock Options, *Journal of Accounting and Economics* 18, 207-231.

Huddart, Steven, and Mark Lang, 1996, Employee stock option exercises: An empirical analysis, *Journal of Accounting and Economics* 21, 5-43.

Inderst, Roman, and Holger M. Müller, 2005, A Theory of Broad-Based Option Pay, CEPR Discussion Paper no. 4878.

Ingersoll, Jonathan E. Jr., 2006, The Subjective and Objective Evaluation of Incentive Stock Options, *Journal of Business* 79, 453-487.

Jenter, Dirk and Fadi Kanaan, 2006, CEO Turnover and Relative Performance Evaluation, Mimeo, MIT Sloan School of Management.

Johnson, Shane A., and Yisong S. Tian, 2000a, Indexed Executive Stock Options, *Journal of Financial Economics* 57, 35-64

Johnson, Shane A., and Yisong S. Tian, 2000b, The Value and Incentive Effects of Nontraditional Executive Stock Option Plans, *Journal of Financial Economics* 57, 3-34

Kahneman, Daniel, and Amos Tversky, 1979, Prospect Theory: An Analysis of Decision Under Risk, *Econometrica* 47, 263-292.

———, 2000, Choices, Values, and Frames, Cambridge University Press.

Kaplan, Steven N, 1994, Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States, *Journal of Political Economy* 102, 510-545.

Kouwenberg, Roy, and William T. Ziemba, 2007, Incentives and Risk Taking in Hedge Funds, *Journal of Banking and Finance* 31, 3291-3310.

Lambert, Richard A., and David F. Larcker, 2004, Stock Options, Restricted Stock, and Incentives, Mimeo, University of Pennsylvania.

Lambert, Richard A., David F. Larcker, and Robert Verrecchia, 1991, Portfolio Considerations in Valuing Executive Compensation, *Journal of Accounting Research* 29, 129-149.

Langer, T., and Martin Weber, 2001, Prospect Theory, Mental Accounting, and Differences in Aggregated and Segregated Evaluation of Lottery Portfolios, *Management Science* 47, 716-733.

Ljungqvist, Alexander, and William J. Wilhelm, 2005, Does Prospect Theory Explain IPO Market Behavior?, *The Journal of Finance* 49, 1759-1790.

Loughran, Tim, and Jay R. Ritter, 2002, Why Don't Issuers Get Upset About Leaving Money on the Table in IPOs?, *Review of Financial Studies* 15, 413-443.

Margiotta, M.-M., and R. M. Miller, 2000, Managerial Compensation and the Cost of Moral Hazard, *International Economic Review* 41, 669-719.

Massa, Massimo, and Andrei Simonov, 2005, Behavioral Biases and Investment, *Review of Finance* 9, 483-507.

Myagkov, Mikhail, and Charles R. Plott, 1997, Exchange Economies and Loss Exposure: Experiments Exploring Prospect Theory and Competitive Equilibria in Market Environments, *The American Economic Review* 87, 801-828.

Oyer, Paul, 2004, Why Do Firms Use Incentives That Have No Incentive Effects?, *The Journal of Finance* 59, 1619-1650.

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, *The American Economic Review* 95, 530-545.

Post, Thierry, Marijn van den Assem, Guido Baltussen, and Richard Thaler, 2008, Deal or No Deal? Decision making under risk in a large-payoff game show, *The American Economic Review* 98, 38-71.

Rabin, Matthew, 2000, Risk Aversion and Expected-Utility Theory: A Calibration Theorem , *Econometrica* 68, 1281-1291.

Thaler, Richard H., 1999, Mental Accounting Matters, *Journal of Behavioral Decision Making* 12, 183-206.

Tversky, Amos, and Daniel Kahneman, 1991, Loss Aversion in Riskless Choice: A Reference-Dependent Model, *The Quarterly Journal of Economics* 106, 1039-1061.

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

Table I: Description of the data set

This table displays mean, standard deviation, and the 10%, 50% and 90% quantiles of the variables in our dataset. Panel A describes our sample of 595 CEOs. Panel B describes all 1,103 executives in the ExecuComp database who were CEO in 2004 and 2005. Panel B also contains the statistic of the two-sample t-test for equal means (allowing for different variances) and the Wilcoxon signed rank test for equal locations. These tests compare the samples in Panel A and B. Before calculating this statistic, we removed all observations from the sample in Panel B that are also contained in the sample in Panel A. "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 in millions.

Panel A: Data set with 595 U.S. CEOs

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	2.50	3.11	0.59	1.67	4.69
Value of contract	π	179.0	1,887.7	5.5	29.8	158.0
Non-firm wealth	W_0	33.3	113.2	2.3	10.3	60.9
Firm value	P_0	10,651	30,260	342	2,275	19,810
Strike price	K	8,243	26,213	242	1,480	13,915
Moneyness	K/P_0	70.1%	20.5%	40.3%	70.8%	98.9%
Maturity	T	4.6	1.3	3.4	4.4	6.0
Stock volatility	σ	42.8%	21.4%	22.9%	36.1%	75.1%
Dividend rate	d	1.24%	2.70%	0.00%	0.61%	3.28%
Age		56.6	6.6	48	57	64

Panel B: All 1,103 ExecuComp CEO's from 2005

Variable		Mean	Std. dev.	10% Quantile	Median	90% Quantile	P-Value of	
							T-test	Wilcoxon
Stock	n_S	2.25%	6.08%	0.03%	0.32%	6.25%	0.267	0.767
Options	n_O	1.25%	1.60%	0.12%	0.76%	2.81%	0.038	0.001
Fixed Salary	ϕ	2.24	3.25	0.53	1.47	4.16	0.176	0.041
Firm Value	P_0	7,454	23,149	433	2,055	16,262	0.047	0.229
Stock Volatility	σ	49.4%	37.2%	22.7%	39.1%	86.1%	0.000	0.019
Dividend Rate	d	1.14%	1.53%	0.00%	0.49%	3.26%	0.439	0.410
Age		56.0	7.6	47	56	65	0.185	0.045

Table II: Optimal piecewise linear contracts

This table describes the optimal piecewise linear contract. It shows the median of the three parameters of the optimal contract, namely base salary ϕ^*, stock holdings n_S^*, and option holdings n_O^*. It also shows the mean of the scaled errors: $error(\phi)=(\phi^*- \phi^d)/\sigma_\phi$, $error(n_S)=(n_S^*- n_S^d)/\sigma_S$, and $error(n_O)=(n_O^*- n_O^d)/\sigma_O$, where σ_ϕ, σ_S, and σ_O denote the cross-sectional standard deviations of base salaries, stock holdings, and option holdings, respectively, and where superscript 'd' denotes parameter values from the observed contract. The table also shows the mean and median of the distance metric D from equation (17), 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 from equation (16). Panel B shows the results for the risk aversion model for seven levels of the CRRA risk aversion parameter γ. Some observations are lost because of numerical problems. 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)		Stock (n_S)		Options (n_O)		Distance D	
			Median	Mean Error	Median	Mean Error	Median	Mean Error	Mean	Median
0.0	594	4.1%	0.29	-1.594	0.005	0.103	0.007	-0.429	0.54	0.16
0.1	578	13.6%	1.47	0.346	0.005	0.015	0.009	-0.022	0.71	0.15
0.2	571	20.1%	1.29	-0.049	0.006	0.050	0.007	-0.135	1.44	0.40
0.4	585	31.3%	-2.89	3.027	0.011	0.285	0.001	-1.136	2.40	0.87
0.6	586	41.1%	-6.74	2.337	0.017	0.526	-0.002	-2.271	3.07	1.13
0.8	585	51.0%	-8.26	-3.294	0.018	0.647	-0.003	-2.921	3.32	1.23
1.0	582	58.3%	-8.89	-10.729	0.019	0.708	-0.005	-3.292	3.47	1.28
Data	595	N/A	1.67	N/A	0.003	N/A	0.010	N/A	N/A	N/A

Panel B: Risk-aversion model

γ	Obs.	Salary (ϕ)		Stock (n_S)		Options (n_O)		Distance D	
		Median	Mean Error	Median	Mean Error	Median	Mean Error	Mean	Median
0.1	595	-8.68	-9.738	0.018	0.699	-0.004	-3.083	3.16	1.09
0.2	593	-8.83	-9.948	0.018	0.724	-0.005	-3.265	3.35	1.17
0.5	595	-8.84	-10.150	0.020	0.749	-0.007	-3.625	3.71	1.40
1	593	-8.27	-9.770	0.021	0.702	-0.011	-3.827	3.89	1.68
3	594	-5.17	-6.682	0.016	0.494	-0.014	-3.953	3.99	2.01
6	585	-1.23	-3.582	0.010	0.260	-0.011	-3.384	3.40	1.83
20	487	0.97	-0.418	0.005	0.050	-0.006	-2.779	2.81	1.54

Table III: Comparison of loss-aversion model with matched risk-aversion model

This table compares the optimal loss-aversion contract with the equivalent optimal 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 (18)). Contracts are piecewise linear. The table shows the average equivalent γ, the mean and median of the difference between the metric D between the RA-model and the LA-model (see equation (17)), and the frequency of this difference being positive. The table also 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 from equation (16). 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.

θ	Obs.	Average equivalent γ	$D_{RA} - D_{LA}$			Percent with positive option holdings		Percent with positive fixed salary		Percent with positive options and salary	
			Percent > 0	Mean	Median	RA	LA	RA	LA	RA	LA
0.0	594	0.21	96.6%	2.75***	0.92***	30.8%	83.3%	1.7%	59.6%	0.3%	52.5%
0.1	578	0.28	97.2%	2.64***	0.87***	30.1%	91.0%	1.6%	77.5%	0.0%	74.2%
0.2	571	0.41	91.8%	2.04***	0.63***	28.2%	82.0%	1.9%	62.7%	0.4%	60.2%
0.3	577	0.52	87.7%	1.54***	0.44***	28.1%	68.3%	1.6%	46.8%	0.3%	44.0%
0.4	585	0.68	89.7%	1.20***	0.30***	25.8%	56.9%	1.4%	32.6%	0.0%	30.6%
0.5	586	0.83	91.0%	0.95***	0.27***	25.6%	48.3%	1.7%	20.6%	0.3%	19.3%
0.6	586	0.95	90.3%	0.72***	0.25***	22.5%	41.3%	1.5%	12.8%	0.0%	11.1%
0.7	582	1.04	88.7%	0.59***	0.24***	20.8%	36.6%	2.1%	8.6%	0.0%	6.4%
0.8	582	1.09	86.4%	0.59***	0.22***	21.0%	33.7%	2.1%	6.5%	0.0%	4.1%
0.9	579	1.06	84.1%	0.50***	0.18***	21.2%	32.5%	2.2%	4.1%	0.2%	2.6%
1.0	581	0.98	82.1%	0.38***	0.14***	22.2%	31.5%	2.1%	3.3%	0.0%	1.9%

Table IV: Optimal nonlinear loss-aversion contracts

This table describes the optimal nonlinear loss-aversion contract. The table shows the median change in wealth if the stock price changes by -50%, -30%, +30%, or +50%. In addition, the table shows the average dismissal probability, defined as the probability with which the contract pays the minimum wage $\underline{w}$ (from equation (20)), the incentives from dismissals that are generated by the drop to the minimum wage $\underline{w}$, and the mean inflection quantile, which is the quantile at which the curvature of the optimal wage function changes from convex to concave. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

θ	Obs.	Mean dismissal probability	Incentives from dismissals	Mean inflection quantile	Median change in wealth if stock price changes by			
					-50%	-30%	+30%	+50%
0.0	571	0.00%	0.01%	98.9%	-59.7%	-37.8%	41.2%	70.1%
0.1	571	0.05%	0.30%	99.9%	-55.5%	-35.9%	41.8%	71.8%
0.2	570	0.56%	2.70%	100.0%	-49.4%	-32.6%	39.9%	70.4%
0.3	574	1.83%	8.79%	100.0%	-40.8%	-27.8%	36.9%	65.8%
0.4	572	4.13%	17.00%	100.0%	-31.4%	-22.3%	32.2%	58.4%
0.5	573	6.53%	24.54%	100.0%	-23.0%	-16.8%	26.5%	49.6%
0.6	573	9.30%	32.80%	100.0%	-16.8%	-12.1%	20.4%	39.1%
0.7	574	12.12%	40.19%	100.0%	-12.6%	-8.9%	16.4%	31.0%
0.8	569	14.82%	47.21%	100.0%	-9.8%	-6.4%	12.7%	24.4%
0.9	563	17.30%	53.57%	100.0%	-8.4%	-4.7%	9.8%	19.5%
1.0	547	19.89%	59.32%	100.0%	-8.2%	-3.5%	7.5%	15.3%

Table V: 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 both models, the table shows the median change in wealth if the stock price changes by -30% or +30%. 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 seven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

θ	Obs.	Linear option contract			General nonlinear contract		
		Median change in wealth if stock price changes by		Mean savings	Median change in wealth if stock price changes by		Mean savings
		-30%	+30%		-30%	+30%	
0.0	570	-39.0%	47.1%	0.2%	-37.9%	41.3%	0.5%
0.1	557	-39.5%	49.8%	0.4%	-35.9%	41.8%	1.5%
0.2	547	-38.6%	47.7%	1.0%	-32.5%	40.1%	3.3%
0.4	567	-34.3%	37.4%	2.3%	-22.2%	32.2%	6.9%
0.6	570	-32.9%	32.5%	3.7%	-12.1%	20.4%	10.1%
0.8	569	-33.7%	30.6%	4.9%	-6.4%	12.7%	13.0%
1.0	546	-34.8%	30.7%	5.6%	-3.5%	7.5%	15.0%

Table VI: 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 (equation (3)). 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 (16). Panel A shows the results for the parameter λ, Panel B for α, and Panel C for β. The table shows the mean and median of the three parameters of the optimal contract, namely, base salary ϕ^*, stock holdings n_S^*, and option holdings n_O^*. In addition, it displays the mean and median of the distance metric D from equation (17). 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	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
1.00	413	-2.28	0.26	0.040	0.007	-0.013	0.004	2.05	0.17
1.50	452	1.53	1.34	0.026	0.006	0.007	0.008	0.76	0.13
2.00	459	2.36	1.58	0.022	0.005	0.013	0.009	0.55	0.13
2.25	578	3.60	1.47	0.019	0.005	0.014	0.009	0.71	0.15
2.50	471	2.61	1.63	0.021	0.006	0.016	0.010	0.62	0.12
3.00	465	2.82	1.68	0.020	0.005	0.015	0.010	0.63	0.13
4.00	466	2.97	1.77	0.020	0.005	0.017	0.010	0.68	0.13

Panel B: Gain space curvature α

α	Obs.	Salary (ϕ)		Stock (n_S)		Options (n_O)		D	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.60	312	-1.30	0.51	0.024	0.006	-0.004	0.002	1.24	0.28
0.70	362	-0.98	0.39	0.023	0.007	-0.003	0.003	1.39	0.26
0.80	394	0.09	0.81	0.028	0.007	0.003	0.006	1.05	0.18
0.88	578	3.60	1.47	0.019	0.005	0.014	0.009	0.71	0.15
0.95	546	2.97	1.93	0.019	0.003	0.016	0.011	0.76	0.11

Panel C: Loss space curvature β

β	Obs.	Salary (ϕ)		Stock (n_S)		Options (n_O)		D	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.60	267	-4.59	-1.33	0.022	0.007	-0.002	0.002	1.17	0.29
0.70	349	-6.24	-1.40	0.034	0.009	-0.014	0.001	2.36	0.39
0.80	388	0.16	0.94	0.036	0.006	-0.005	0.007	1.64	0.14
0.88	578	3.60	1.47	0.019	0.005	0.014	0.009	0.71	0.15
0.95	508	2.58	1.64	0.020	0.005	0.015	0.010	0.63	0.13

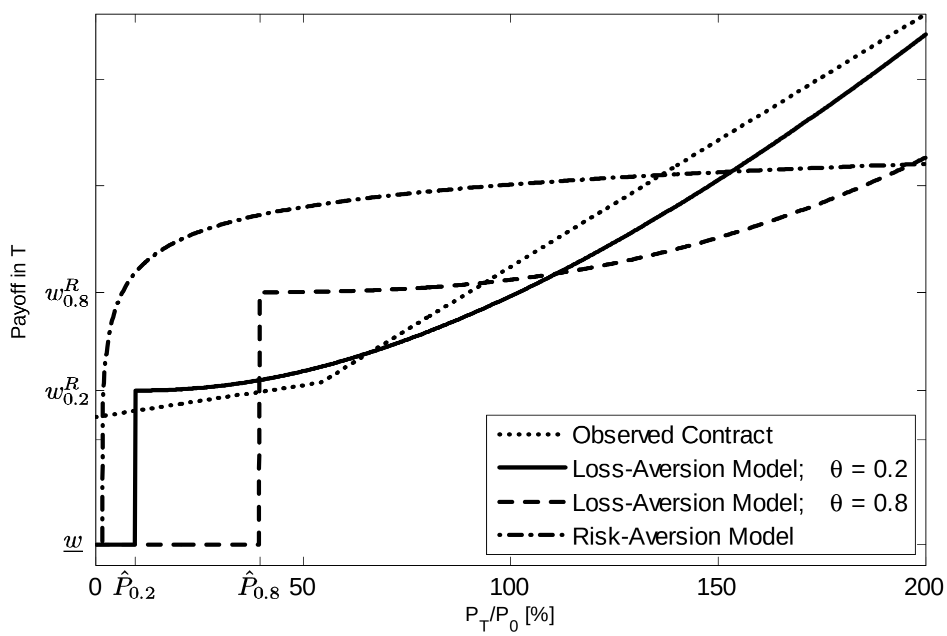

Figure 1. Optimal and observed contracts for a representative CEO. The figure plots the end-of-period wealth for the loss-aversion-contract (equation (19)) for two reference wages ($\theta = 0.2$ and $\theta = 0.8$), the risk-aversion-contract for $\gamma = 2$, and the observed contract for a representative CEO in our sample. The parameters of the observed contract are $\phi = \$2.3$m, $n_S = 0.30\%$, $n_O = 0.95\%$. Initial nonfirm wealth is $W_0 = \$10.5$m, P_0 is \$2.3bn, K/P_0 is 54%, $T = 4.2$ years, $r_f = 3.6\%$, and $d = 2.0\%$.