Oliver Spalt

Stock Options as a Compensation Device: A Behavioral Approach

Inauguraldissertation

zur Erlangung des akademischen Grades

eines Doktors der Wirtschaftswissenschaften

der Universität Mannheim.

Vorgelegt im Frühjahrsemester 2008

Dekan:	Professor Dr. Hans H. Bauer
Referent:	Professor Ernst Maug, Ph.D.
Korreferent:	Professor Dr. Dr. h.c. Martin Weber
Tag der Disputation:	26. Juni 2008

Contents

I Introduction: Stock Option Compensation and Prospect Theory **1**

1 Overview **1**

2 Stock options as a compensation device for executive and non-executive employees **3**

3 Prospect theory as a description of employee preferences **11**

4 Outline of the dissertation and main results **16**

II Probability Weighting and Employee Stock Options **18**

1 Introduction **18**

2 Related literature on stock options for non-executive employees **22**

3 The model **25**

3.1 Model set-up 25

3.2 Theoretical results 31

4 Calibration of the model **34**

4.1 Parameterizing the model 34

4.2 Calibration results 36

4.3 Hypotheses 39

5 Empirical tests of the model **40**

5.1 Dataset 40

5.2 Empirical results 44

5.3 Robustness checks 52

6 The relation between stock option grants and exercises **54**

7 Are firms in a special position to exploit the bias? **59**

8 Conclusion **60**

Appendix **63**

A Proof of Proposition 1 **63**

III Sticks or Carrots? Optimal CEO Compensation when Managers are Loss Averse **68**

1 Introduction **68**

2 The model **73**

3 Analysis **77**

3.1 Discrete effort 77

3.2 Continuous effort 79

4 Implementation and data **81**

4.1 Implementation 81

4.2 Data 87

5 Contracts with restricted stock and options **91**

6 General non-linear loss-aversion contracts **101**

7 Robustness checks **106**

8 Conclusion **115**

Appendix **118**

A Proofs **118**

B The optimal contract when P_T is lognormal and effort is continuous **125**

IV The Optimal Design of Stock Options: A Loss-Aversion Approach **127**

1 Introduction **127**

2 The analytic approach **131**

2.1 The model 131

2.2 Analyzing optimal contracts 134

3 Data **137**

3.1 Observed contracts 137

3.2 Preference parameters 139

3.3 Measures of agency costs and managerial discretion 139

3.4 Descriptive statistics 141

4 Results **143**

4.1 Are premium options optimal? 143

4.2 Comparative static analysis 145

4.3 Which companies should use premium options? 152
4.4 Governance implications 155

5 Robustness checks **157**

6 Conclusion **164**

References **167**

List of Figures

1.1 Average annual CEO compensation in S&P 500 firms from 1992 to 2005 . 3

1.2 Average annual compensation to CEOs, other top 5 executives, and other employees in S&P 500 firms from 1992 to 2005 10

1.3 The prospect theory value function 12

2.1 The probability weighting function as proposed by Tversky and Kahneman (1992) for different probability weighting parameters δ. 27

3.1 Optimal non-linear contracts under the loss-aversion model and risk-aversion model. 79

3.2 Optimal linear contracts under the loss-aversion model and risk-aversion model. 100

4.1 Histogram of optimal strike prices 143

4.2 Comparative static results for the representative CEO 147

4.3 Optimal non-linear and optimal piecewise-linear contracts for the representative CEO 149

4.4 Robustness with respect to the reference point 159

List of Tables

2.1 Estimates of the parameter δ in the probability weighting function. 36

2.2 Calibration results 37

2.3 Descriptive statistics 43

2.4 Employee stock option grants sorted by firm volatility 45

2.5 Regressions of the log of the number of employee stock options on firm volatility and control variables 47

2.6 Regressions of the log of the number of employee stock options on firm volatility and enlarged set of control variables 49

2.7 Regressions of an indicator variable for the existence of a broad-based employee stock option plan on firm volatility and control variables 51

2.8 Regressions of an indicator variable for the existence of a broad-based employee stock option plan on firm volatility and enlarged set of control variables 53

2.9 Robustness checks 55

2.10 The influence of the evaluation horizon and the moneyness of options on exercise decisions 58

3.1 Description of the dataset 89

3.2 Optimal piecewise-linear contracts. 92

3.3 Comparison of loss-aversion model with matched risk-aversion model . . . 98

3.4 Optimal non-linear loss-aversion contract 103

3.5 Comparison of linear and non-linear loss-aversion models 106

3.6 Wealth robustness check 107

3.7 Restricted models with positive salaries and positive option holdings 109

3.8 Results for the 1997 sample 111

3.9 Comparative statics for the parameters of the value function 113

3.10 Ownership robustness check 114
4.1 Descriptive statistics 142
4.2 Optimal strike prices and savings 144
4.3 Representative CEO 146
4.4 Cross-sectional comparative static results 151
4.5 Regression of predicted premium on observed contract parameters 153
4.6 Cross-sectional variation of contracts 154
4.7 Agency costs and contractual inefficiency 156
4.8 Optimal strike prices and savings for the model with restricted base salary 158
4.9 Comparative static results for preference parameters 160
4.10 Descriptive statistics for the 1997 sample 162
4.11 Optimal strike prices and savings for the 1997 sample 163
4.12 Agency costs and contractual inefficiency for the 1997 sample 164

Chapter I

Introduction: Stock Option Compensation and Prospect Theory

1 Overview

This dissertation analyzes existing compensation schemes for executive and non-executive employees in the light of recent advances in decision sciences and behavioral corporate finance. The primary object of study are stock options, which give the holder the right to buy a share of stock during a pre-specified future period of time, at a pre-specified price (called the "strike price"). Stock options are the predominant device for compensating CEOs and other executive employees. They are also frequently used to remunerate non-executive employees. The importance of stock options as an economic reality in compensating employees is undisputed.[1] Yet, while financial researchers have spent a great amount of time and effort on the subject, there is still considerable uncertainty with regard to many of the most basic questions, including why executive employees are compensated with stock options instead of, for example, restricted stock, why important terms of option grants, such as strike prices, are so uniform across companies, and why stock options are used to compensate non-executive employees.

The contribution of this dissertation is to show that many of these questions can be answered by using prospect theory preferences for employees in otherwise standard

[1] See for example Figures 1.1 and 1.2 in the next section.

economic models. For non-executive employees, I show in Chapter II that probability weighting, which is a main component of prospect theory, can explain why a large number of firms have broad-based stock option plans, and why firms with such a plan tend to have more volatile stock prices. For CEOs, I find in Chapter III that loss aversion, which is another main component of prospect theory, can explain why we see stock options in almost all executive pay contracts. Chapter IV shows that loss aversion can also explain why strike prices are so homogenous across firms and why strike prices are usually set equal to the grant date stock price.

The work presented here is conducted on the basis of Standard and Poor's ExecuComp database, which collects detailed annual compensation data for the top executive officers of more than 2,500 U.S. companies beginning in the year 1992. Apart from data availability, focusing on the U.S. has the additional benefit of making my results comparable to most prior research on stock option compensation.[2]

In this introductory chapter I will outline and present the main ideas of this dissertation, as well as some institutional background on stock option compensation in the U.S. A brief introduction to prospect theory will be followed by a synopsis of the main results. Detailed discussion of the models and the relevant literature is relegated to the respective Chapters II to IV.

[2] A comprehensive survey of the vast literature on stock option compensation and executive compensation is far beyond the scope of this chapter. Survey articles on various aspects are Abowd and Kaplan (1999), Murphy (1999), Prendergast (1999), Core, Guay, and Verrecchia (2003), Hall and Murphy (2003) and Jensen and Murphy (2004). For studies on international comparisons of pay arrangements see Abowd and Bognanno (1995), Abowd and Kaplan (1999), and Conyon and Schwalbach (2000).

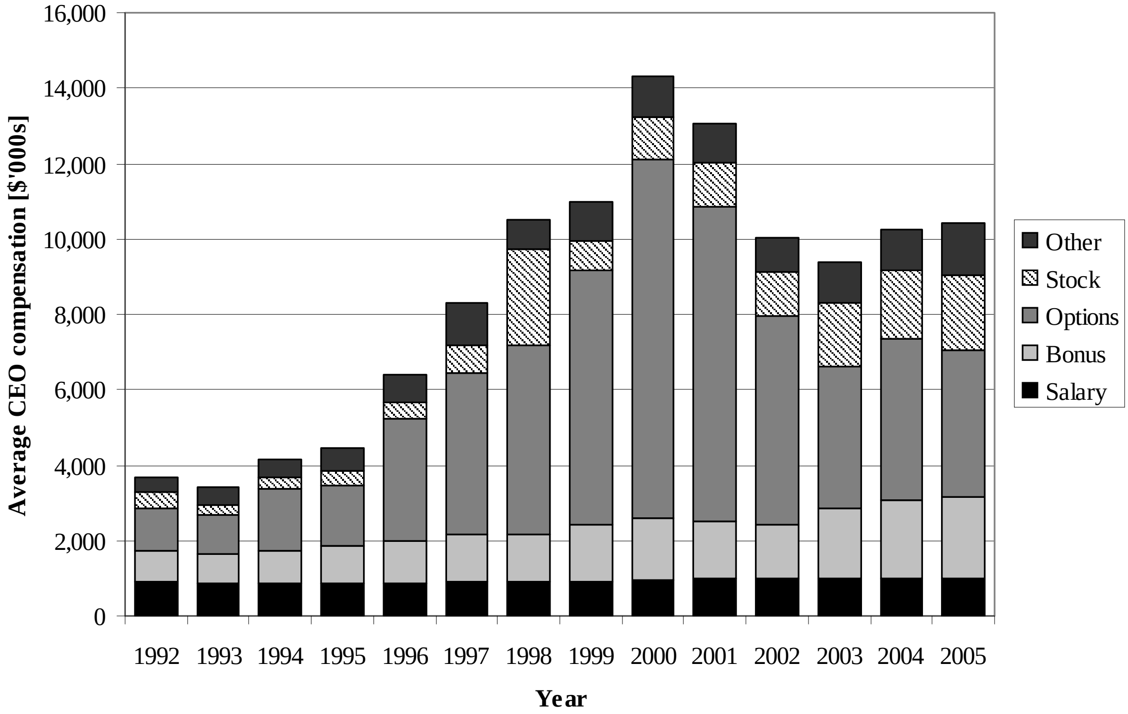

Figure 1.1: This figure shows the average level and structure of CEO compensation in S&P 500 firms from 1992 to 2005. Salary, bonus, the value of stock option grants (valued using the Black-Scholes methodology), and the value of restricted stock grants are taken directly from ExecuComp. Other compensation includes ExecuComp items: "other annual", "all other total" and payouts from long-term incentive plans. All values are given in 2005 constant dollars using the Bureau of Economic Analysis GDP deflator.

2 Stock options as a compensation device for executive and non-executive employees

This section presents an overview of the use of stock options as a compensation device in the U.S., reviews the main economic arguments for their use, and extracts three main questions that this thesis addresses.

I start by looking at executive employees. As a first remarkable empirical fact, almost all CEOs of the largest U.S. companies receive stock options as part of their compensation package. In 2005, about 96% of S&P 500 companies granted options to their CEO.[3] Figure 1.1 shows the average annual CEO compensation in S&P 500 firms, derived on the basis of the ExecuComp database, over the years 1992 to 2005.[4] CEO pay has increased enormously over this period and the average annual pay for an S&P 500 CEO grew from \$3.6 million in 1992 to over \$14 million in 2000.[5,6] Although, by 2005, CEO compensation has decreased to the 1998 level at about \$10.4 million, it still surpasses the 1992 level by factor 2.9, which implies growth of 8.5% per annum in real terms over 13 years. Much of the growth in CEO pay is attributable to growth in stock options grants. Stock options are also the single most important component in the structure of CEO pay. In 2005, 33% of the fair value of the pay package awarded to the average S&P 500 CEO came from stock options, 23% from bonus payments and 16% from fixed salaries. The relative importance of restricted stock has increased from about 7% of pay in 1997 to 15.8% in 2005.[7]

Figure 1.1 shows that through the use of stock options and restricted stock, a substantial part of CEO pay is tied directly to the performance of the stock price of the company. The relation between pay and performance is even stronger when one takes into account that, in addition to the most current grant of stock and options, the CEO usually holds a portfolio of stock options and restricted stock from previous grants. The value of these holdings varies systematically with performance, thus amplifying the sensitivity of CEO

[3] This number is based on the number of S&P 500 CEOs who are recorded in ExecuComp in the year 2005.

[4] See Jensen and Murphy (2004) for a similar presentation of the S&P 500 compensation data over the years 1992 to 2002.

[5] All dollar values here are in 2005 constant dollars, which were computed using the GDP deflator of the Bureau of Economic Analysis (www.bea.gov).

[6] I do not focus on the level of pay in this dissertation. For important contributions to the debate on the level of pay, see Holmström and Kaplan (2001), who argue that large stock option plans helped to overcome managerial resistance to efficiency-increasing corporate restructurings in the 1980s and 1990s. Bebchuk and Fried (2004) and Bebchuk and Grinstein (2005) argue strongly for a managerial-power view, where managers use stock options as a means to extract rents from their companies. Gabaix and Landier (2008) show that a model of CEO talent, where the most talented CEOs are matched to the largest firms is consistent with the observed rise in CEO pay.

[7] The remaining 12% are miscellaneous items such as payouts from long-term incentive plans, perquisites, signing bonuses, and 401K contributions.

wealth with respect to corporate performance (Hall and Liebman, 1998).

Why do firms tie pay to performance? The standard economic argument for granting performance-related pay is the need to mitigate agency problems which stem from the separation of ownership and control in modern corporations (Berle and Means, 1932, and Jensen and Meckling, 1976).[8] The basic idea is that shareholders (the "principal") cannot perfectly monitor managers (the "agent"), and that in the absence of monitoring, managers maximize their own utility rather than the utility of the shareholders. As a result, CEOs may, for example, not work hard enough, they may spend money on pet projects, they may buy a new corporate jet, they may engage in empire building, or they may pay above market-level wages to employees.

The seminal theoretical literature on the principal-agent problem has condensed this myriad of ways to waste corporate resources into a single variable called "effort." Holmström (1979), for example, considers a static model in which a risk-neutral principal can make a take-it-or-leave-it offer of a pay contract w to a risk-averse manager.[9] The manager will accept the contract if it provides at least the same expected utility as an exogenous outside option. If the manager accepts, she can exert costly effort, denoted by e. The stock price at the end of the period, denoted by $P(e, u)$, is assumed to depend on effort and a random state of nature u. Moreover, by assumption, if the manager exerts more effort, it is more likely that the end-of-period stock price is high. The central assumption is that effort is not observable by the principal. Hence, given the pay contract, the manager

[8] Although the principal-agent motivation outlined here is the most common approach, tying pay to performance may also help to attain other objectives such as: providing retention incentives for employees (Oyer, 2004), providing incentives for risk-averse CEOs to choose positive NPV projects even if these are risky (Hemmer, Kim, and Verrecchia, 1999, Feltham and Wu, 2001), attracting a certain type of employees (Oyer and Schaefer, 2005, Lazear, 2005), or, at least up to 2005, providing favorable accounting treatment (Jensen and Murphy, 2004).

[9] This model will also be the basis for the extension to loss-averse CEOs in Chapters III and IV. The original Holmström (1979) model is slightly more general in that it allows for risk-averse principals, and uses "payoffs" instead of stock prices. I adapt the model to the present context in this exposition. For other seminal papers on the principal-agent problem see Ross (1973), Mirrlees (1974, 1976), Holmström (1982), Grossman and Hart (1983), Holmström and Milgrom (1987) and Holmström and Milgrom (1991).

will choose the level of effort that maximizes her expected utility. If w is a constant, that is if pay is not tied to the stock price, the manager will choose the least costly level of effort, and, as a result, the end-of-period stock price will likely be low. The principal can "incentivize" the manager by making pay contingent on the stock price. Exerting at least some effort becomes attractive then, since the manager shares in the benefits of increasing company value. However, CEO risk aversion introduces a countervailing effect: the more pay is tied to performance, the more risk is conferred to the manager (because the stock price is a random variable). Hence, there is a trade-off between incentivizing the manager to increase company value and efficient risk sharing.[10] Holmström (1979) has shown that under some technical assumptions the optimal contract will be a monotonically increasing function of the stock price.[11]

According to the principal-agent paradigm, stock and stock options are granted to executives to align their interests with the interests of the shareholders. An important question is how one can measure incentives from observed pay packages. Following Jensen and Murphy (1990), a large literature has defined incentives as the dollar change in CEO wealth for a dollar change in firm value (called the "pay-performance sensitivity"). Jensen and Murphy (1990) analyze a sample of U.S. firms from 1974 to 1986 and find a pay-performance sensitivity of \$3.25 for a \$1,000 change in firm value, which they regard as too low. Hall and Liebman (1998) show that pay-performance sensitivities have increased

[10] There is an ongoing debate about the trade-off between risk and incentives in the literature. Aggarwal and Samwick (1999) find that pay-performance sensitivities for executives are decreasing in the variance of the firm's performance, thus confirming the predicted trade-off. Prendergast (2002) presents evidence that the trade-off may be more tenuous than commonly assumed.

[11] In particular, he assumes the monotone likelihood ratio property (MLRP)

$$\frac{f_e\left(P|e\right)}{f\left(P|e\right)} > 0,$$

where f is the density function of the stock price conditional on effort, and f_e is its first derivative with respect to effort. Intuitively, MLRP implies that the higher the observed stock price, the more likely it is that more effort was exerted.

significantly over the later period from 1980 to 1994 and that almost all incentives come from the influence of the stock price on the holdings (as opposed to the flow) of stock and stock options. Almost no incentives come from salary and bonus payments.

Another important question is how pay packages should be structured. Specifically: should firms use stock options, or, rather, restricted stock as a compensation device? On the one hand, stock options are an efficient way to provide incentives, since for the same dollar outlay, the number of options that can be granted exceeds the number of shares that can be granted, and total incentives in terms of pay-performance sensitivity are usually higher for options than for shares (Hall, 1998). Hall and Murphy (2000, 2002) show that the advantage of options over stock in terms of incentives is robust to endowing the manager with a constant relative risk aversion (CRRA) utility function and defining incentives as the derivative of the expected utility of the pay contract with respect to the stock price. On the other hand, options are an expensive form of compensating risk-averse executives since risk aversion drives a wedge between the cost to the company and the value to the executive. Since stock options are riskier than shares, the company can save compensation costs and keep expected utility constant by exchanging a given number of options for restricted stock (Lambert, Larcker, and Verrecchia, 1991, Oyer and Schaefer, 2005). Dittmann and Maug (2007) bring together these two perspectives by calibrating the traditional Holmström (1979) model for lognormally distributed stock prices and CEOs with CRRA utility. They find that for almost all CEOs in their sample, stock options are not efficient and that the predicted optimal structure of CEO pay would be contracts that feature restricted stock, negative base salaries and no options.[12]

The results by Dittmann and Maug (2007) are intriguing because, taken at face value, they suggest that either observed pay contracts are not efficient, or that the standard principal-agent model used in the literature is not the right model to use. Chapter

[12] Some of these results are already implicit in Hall and Murphy (2002).

III presents an efficient-contracting model, using prospect theory preferences for CEOs, that can reconcile the principal-agent model in the spirit of Holmström (1979) with the observed structure of CEO pay.

A striking feature of stock option compensation is the near uniformity of stock option design (Hall and Liebman, 1998, Murphy, 1999, Hall and Murphy, 2002). A typical stock option plan contains options with a maturity of 10 years and stipulates that managers cannot exercise the options for a certain time. In most cases, options become exercisable (they "vest") two to three years after the grant date, or they become exercisable in equal annual increments over the first two to four years following the grant. The options cannot be sold and executives are prohibited from diversifying the risk imposed on them by, for example, short-selling company stock.[13]

On of the most prominent regularities is that almost all options have a fixed strike price equal to the stock price at the grant date (such options are called "at the money options"). Hall and Murphy (2002) report that in 1998, 94% of companies in the S&P 500 used at the money options to compensate their executives. This practice has come under severe scrutiny by some academics who argue that granting at the money options provides executives with huge windfall profits since, on average, share prices tend to appreciate, and executives are rewarded for general market movements beyond their control (Rappaport, 1999, Bebchuk and Fried, 2004). These scholars argue that the windfall profits generated by at the money options are indicative of inefficient pay-setting processes and suggest that companies should instead use premium options, which have a fixed strike price set above the current stock price, or indexed options, which reward the executive relative to some benchmark index.[14] At the other extreme, Hall and Murphy (2000, 2002) argue that

[13] Bettis, Bizjak, and Lemmon (2001) provide evidence that by using specifically tailored derivative structures, executives can hedge part of their risk exposure. The small sample size in this study indicates, however, that these practices are not particularly widespread.

[14] A theoretical justification for using indexed options comes from the "informativeness principle" in Holmström (1979, 1982), which holds that contracts should only be written on informative signals of the

setting at the money strike prices is efficient if one explicitly models the trade-off between incentives and cost to the company.

A major drawback for the efficient-contracting view is that there is currently no benchmark efficient-contracting model that can explain stock options in CEO pay contracts and, consequently, no firm basis for investigating stock option design features. Chapter IV of this dissertation extends the model in Chapter III by endogenizing strike prices and shows that at the money options are efficient for sensible assumptions about model parameters.

Figure 1.1 has shown that stock options are important as a compensation device for executives. Figure 1.2 shows that stock option grants to CEOs are just the tip of the iceberg.[15] The vast majority of options are granted to employees below the top 5 executives reported in ExecuComp ("other employees"). In any year from 1992 to 2005, more than 85% of options in S&P 500 companies were on average granted to non-top 5 employees. In 2005, these employees have on average been granted stock options worth $82 million. In 2000 this number has amounted to $227 million. Hall and Murphy (2002) cite a study by the American Compensation Association which reports that about 45% of exempt salaried employees, 12% of non-exempt salaried workers and 10% of hourly employees received stock options in a broad sample of U.S. companies in 1998. The National Center for Employee Ownership estimates that as of early 2008, there are some 3,000 U.S. companies with a broad-based stock option plan, which is a stock option plan

actions of the agent. Since general market movements are not influenced by the CEO, compensation should not be based on them. For a more detailed discussion on (the lack of) relative performance evaluation see Murphy (1999) and Core, Guay, and Verrecchia (2003).

[15] The ExecuComp database contains only the compensation of the top 5 executives for each firm. Also reported is the variable "pcttotopt", which reports for each executive stock option grant the percentage this grant represents of the total number of options granted by the firm in the fiscal year. Hence, it is possible to infer the total number of options to other employees from the reported grants. The Black-Scholes value of option grants to other employees in Figure 1.2 was estimated using the assumption that the ratio of Black-Scholes value of options granted to the top 5 executives to the Black-Scholes value of options granted to other employees is equal to the ratio of the respective number of options.

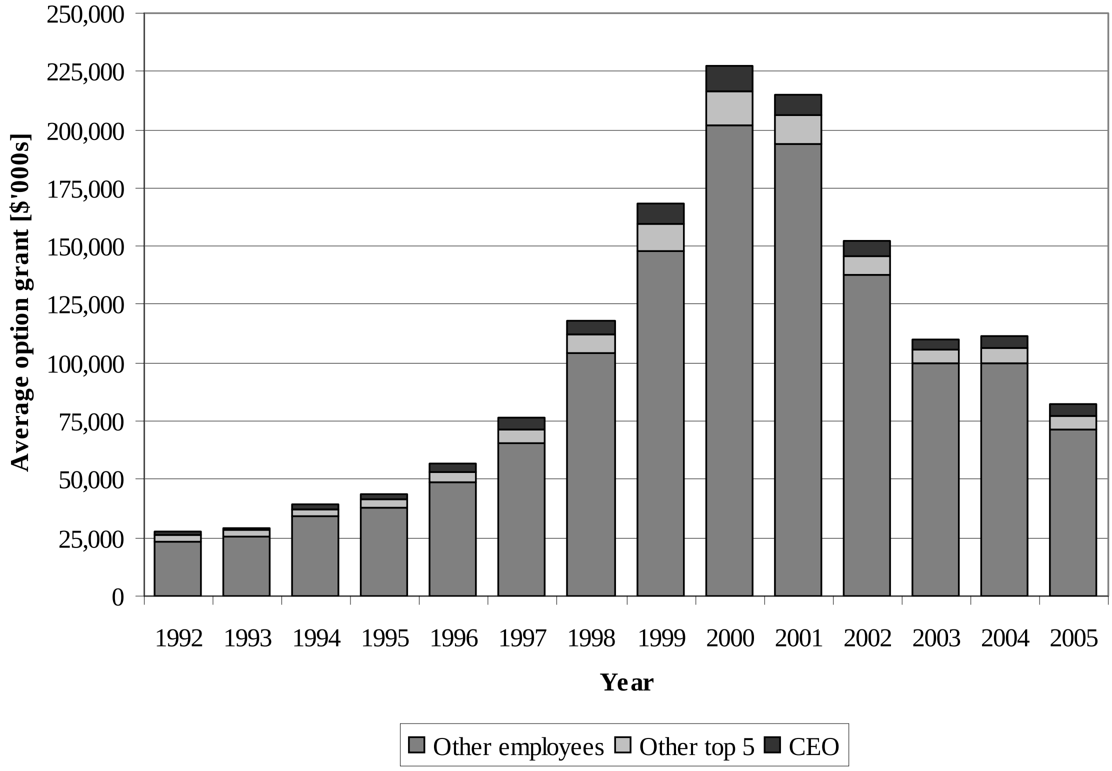

Figure 1.2: This figure shows the average option grant to CEOs, other top 5 executives, and other employees in S&P 500 firms from 1992 to 2005. The value of stock option grants for CEOs and other top 5 executives are taken directly from ExecuComp (valued using the Black-Scholes methodology). The value of grants to other employees is estimated based on the ExecuComp variable "pcttotopt". All values are given in 2005 constant dollars using the Bureau of Economic Analysis GDP deflator.

that grants options not only to executives but also to non-executive employees, and that about 9 million Americans participate in these plans.[16] These numbers make it clear that stock options are of great economic importance not only for CEOs and other executives, but also for non-executive employees.

The fact that stock options are also used to compensate non-executive employees is puzzling in the light of the principal-agent paradigm outlined above because the stock price of a company is a largely uninformative measure of the effort put in by a single

[16] Data available at: www.nceo.org/library/eo_stat.html

non-executive employee. Hence, according to the "informativeness principle" of Holmström (1979, 1982), pay to non-executive employees should not be tied to the stock price. Moreover, compensating lower-level employees based on company performance introduces a substantial free-rider problem, since all employees will have an incentive to cut back on effort and let the other employees do the work (Hall and Murphy, 2003, Bergman and Jenter, 2007). Moreover, Oyer and Schaefer (2005) calibrate a benchmark model to show that incentives from stock options – if they existed – would have to be much larger than what can reasonably be assumed to justify the cost of imposing additional risk on the employees. Chapter II provides a novel explanation, based on employees with prospect theory preferences, for the existence of broad-based stock option plans.

To sum up, stock options are one of the most important and widespread compensation devices. Despite the large literature on compensation, some of the most basic questions are still not satisfactorily answered. From the previous discussion, there emerge three important questions, which this dissertation addresses in the following three chapters:

- Why do some firms grant stock options to non-executive employees?
- Why are stock options part of almost all observed CEO pay contracts?
- Are at the money strike prices efficient?

3 Prospect theory as a description of employee preferences

This dissertation takes a behavioral approach to explaining stock option compensation by assuming that executive and non-executive employees have prospect theory preferences. Prospect theory was developed on the basis of substantial experimental evidence by Kahneman and Tversky (1979) as a descriptive theory of risky choice between gambles with

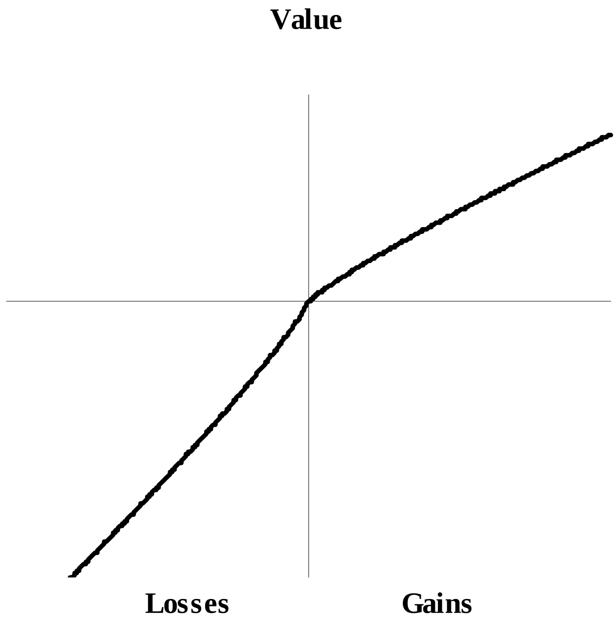

Figure 1.3: The prospect theory value function. The reference point is located at the intersection of the horizontal and vertical axes. Values on the horizontal axis correspond to the distance of an outcome from the reference point.

a small number of outcomes.[17] It was subsequently extended by Tversky and Kahneman (1992) to the domain of uncertainty and to risky gambles involving any number of outcomes. The key features of prospect theory are reference dependence, loss aversion, diminishing sensitivity, and probability weighting. All these features will be mathematically defined in the following chapters so I will only give a brief overview here.

Figure 1.3 shows the "value function", which is the prospect theory equivalent to the utility function. The particular shape of the function is determined by three principles.

Reference dependence. *Individuals evaluate the outcome of a gamble relative to a reference point. Hence, the carriers of value are gains and losses, not final wealth positions.*

[17] See Kahneman and Tversky (2000) for a collection of papers that summarizes much of the available empirical evidence.

Kahneman and Tversky (1979), p. 277, explain:

> "Our perceptual apparatus is attuned to evaluation of changes or differences rather than the evaluation of absolute magnitudes. When we respond to attributes such as brightness, loudness or temperature, the past and present context of experience defines an adaptation level, or reference point, and stimuli are perceived in relation to this reference point. [...] The same principle applies to non-sensory attributes such as health, prestige, and wealth."

Outcomes below the reference point are labeled losses, while outcomes above the reference point are labeled gains. In Figure 1.3, the reference point is located at the intersection of the horizontal and vertical axes.[18]

Loss aversion. *Losses loom larger than gains.*

Loss aversion implies that individuals dislike losses more than they are attracted to equal-sized gains. Loss aversion introduces a kink in the value function. As a result, individuals are infinitely risk averse at the reference point and the value function is steeper for losses than for gains. There is extensive experimental support for loss aversion in a large set of contexts.[19]

Diminishing sensitivity. *The marginal value of both losses and gains decreases with their size.*

[18]Rayo and Becker (2007) have proposed an evolutionary foundation for reference-dependent preferences.

[19]For example, Thaler (1980) and Kahneman and Tversky (1984) find that individuals are more sensitive to out-of-pocket costs than to opportunity costs, and more sensitive to losses than to foregone gains. Dunn (1996) finds evidence for loss aversion by analyzing labor/leisure choices. Loss aversion explains asymmetric consumption responses to expected losses or gains in income (Bowman, Minehart, and Rabin, 1997). Camerer, Babcock, Loewenstein, and Thaler (1997) analyze hours worked by Taxi drivers in New York City and find that the observed pattern of work supply can be readily explained by loss aversion but not by standard models. For additional evidence see Tversky and Kahneman (1991) and the references therein. Ashraf, Camerer, and Loewenstein (2005) argue that loss aversion was already recognized by Adam Smith. Rabin (2000), p. 1288, calls loss aversion "the most firmly established feature of risk preferences."

Since gains and losses are valued relative to the reference point, diminishing sensitivity implies that the value function is concave over gains and convex over losses. The convexity over losses introduces risk-loving behavior in the loss domain.

Another key ingredient of prospect theory is probability weighting. Probability weighting captures the fact that individuals tend to overweight small probabilities, underweight medium to large probabilities, and that individuals are more sensitive to changes in probabilities if these changes are close to certainty or close to impossibility. I present a detailed mathematical exposition of probability weighting in Chapter II.

There is a large and comprehensive body of experimental evidence that shows that individuals frequently violate basic tenets of rationality in making decisions under risk and uncertainty.[20] Serious attempts to incorporate the insights of decision scientists and cognitive psychologists into economics have been made at least since the seminal article on prospect theory from Kahneman and Tversky (1979). Since then, hardly any field has grown so rapidly as the field of behavioral economics and behavioral finance.[21]

The literature on behavioral corporate finance is also growing, but existing research has focused mainly on irrational investors, and the ways in which rational managers can exploit these investors by, for example, issuing overvalued equity.[22] By contrast, the literature on irrational managers is not as well developed and has so far focused mainly on managerial optimism and overconfidence, but not on alternative preference specifications

[20]Rationality is defined here in the sense of the axioms underlying expected utility theory of von Neumann and Morgenstern (1944). One of the earliest criticisms of these axioms comes from Allais (1953) who shows that individuals frequently violate the independence axiom. Another early demonstration of individual biases in decision making can be found in Tversky and Kahneman (1974). For an overview of some of the literature on the limits of, and alternatives to, expected utility theory see for example, Schoemaker (1982), Starmer (2000) and Fox and See (2003).

[21]For detailed surveys of the literature in behavioral economics see Rabin (1998, 2002), Camerer, Loewenstein, and Rabin (2003) and DellaVigna (2007). For a comprehensive overview of research on behavioral finance see Thaler (1993, 2005).

[22]The distinction between the irrational-investor and irrational-manager approaches used here is borrowed from the survey article on behavioral corporate finance by Baker, Ruback, and Wurgler (2006). Detailed references to the respective approaches can be found there.

like prospect theory.[23] This is surprising for two reasons. First, if deviations from rationality are relevant for investors, then there is every reason to believe that they are also relevant for other individuals including managers. Second, limits to arbitrage are likely to be greater inside organizations. Managers can entrench themselves and there is ample evidence that takeover threats, proxy fights and other corporate governance mechanisms frequently leave some discretion with management (Baker, Ruback, and Wurgler, 2006). Hence, one would expect effects of behavioral biases to show up very prominently in decisions of managers. This dissertation contributes to the literature on irrational managers in Chapters III and IV. It contributes to an even smaller behavioral corporate finance literature, which investigates the consequences of employee irrationality, in Chapter II.[24]

Stock option compensation is an ideal field to test the relevance of behavioral theories of decision making for CEO and employee behavior, because CEOs and employees are prohibited from hedging their exposure to the risk of their equity contracts, and because there is no close substitute for the 10 year options typically awarded in stock option plans. Hence, there is some protection from the forces of arbitrage. If preferences matter, they are likely to matter in such an environment. Although there is a large number of behavioral biases and alternatives to expected utility theory, I focus exclusively on prospect theory.[25] By focusing on prospect theory I can build on a firmly established theoretical framework, use existing results on functional forms, and experimental results on values for preference parameters in my numerical work.

[23] Papers that examine the influence of optimism and overconfidence on manager are, for example, Heaton (2002), Malmendier and Tate (2005a, 2005b, 2008) and Hackbarth (2008).

[24] Other papers in this category are Oyer and Schaefer (2005) and Bergman and Jenter (2007).

[25] Examples of alternative theories include Machina (1982), Bell (1982), Loomes and Sugden (1982), Fishburn (1984), Dekel (1986), Gul (1991), and Neilson (1992). See Starmer (2000), Wu, Zhang, and Gonzales (2004), and Fox and See (2003) for surveys.

4 Outline of the dissertation and main results

The questions derived at the end of Section 2 are addressed in the following three chapters of this dissertation.

Chapter II presents a model which can rationalize the puzzling fact that stock options are frequently used as a compensation device for non-executive employees. The model proposes that this can be attributed to the tendency of individuals to overweight small chances of large payoffs. A simple model with employees that have prospect theory preferences including (cumulative) probability weighting is calibrated using standard assumptions on stock price distributions and preference parameters. The calibrated model predicts that employees in risky firms, which are firms with a high stock price volatility, attach a value to their stock options in excess of the cost to the firm. Hence, exchanging base salaries for stock options in contracts for lower-level employees is attractive to risky firms as long as they can cut back base salaries. The model predicts that stock option grants to non-executive employees should be concentrated among the most risky firms. Conversely, the model predicts significantly less stock option grants to non-executive employees in low-risk firms. These predictions are strongly supported by results from regression analyses using a large panel with more than 2,000 firms over the years 1992 to 2005.

Chapter III explains the observed structure of CEO compensation and in particular the presence of stock options in almost all observed compensation contracts. The chapter extends the principal-agent model by Holmström (1979) by assuming prospect theory preferences for the agent. The general optimal contract is established analytically and the model is calibrated to the observed contracts of 595 CEOs, using parameter values for the prospect theory value function from the experimental literature. The model explains the observed structure of executive compensation contracts significantly better than the original model with expected utility preferences for the CEO. This holds especially for the mix between stock and options. The model predicts convex contracts with substantial option holdings that provide a strong upside ("carrots"). By contrast, derived optimal contracts in the standard model with a risk-averse CEO are concave and expose the CEO to significant downside risk ("sticks"). The key insight is that a contract that combines stock options with a higher base salary is attractive to loss-averse CEOs, since such a contract provides valuable downside protection. The results in Chapter III suggest that loss aversion is a better paradigm for analyzing the design features of stock options and for developing preference-based valuation models than the conventional risk-aversion paradigm used in the literature.

Chapter IV extends the model in Chapter III to analyze optimal strike-price design of executive stock options. The results show that the absence of premium options is not as puzzling as it may seem. The extended model is calibrated to a sample of 724 CEOs, and for each CEO optimal strike prices are endogenously and jointly determined with the number of granted options, the number of granted shares, and the optimal level of base salary. For low reference points, at the money options are optimal for the average firm in the sample. For higher reference points, premium options become optimal, but the savings firms could realize by switching from at the money options to premium options would be very small. Hence, the model suggests that optimal strike-price design is a second-order issue in terms of efficiency.

Chapter II

Probability Weighting and Employee Stock Options

In 2004, approximately 99% of the options that Intel granted went to employees other than its top six most highly compensated executive officers; for the period 2000 to 2004, only 1.2% of all options that Intel granted went to its top five most highly compensated executive officers.
Intel Corporation, 2005 proxy statement

1 Introduction

The National Center for Employee Ownership estimates that in early 2008, broad-based stock option plans, which they define as plans that grant options to 50% or more of their full-time employees, are used in about 3,000 U.S. companies.[1] They also estimate that approximately 9 million Americans participate in broad-based option plans.[2] While the economic importance of such plans is obvious, the economic rationale for issuing options to non-executive employees is still a puzzle.

This chapter proposes probability weighting as a novel, preference-based, explanation for why firms use stock options to compensate non-executive employees. Probability

[1] I thank Aurelien Baillon, Roman Inderst, Elu von Thadden, Martin Weber, David Yermack, seminar participants at the University of Mannheim, and especially Ingolf Dittmann, Ernst Maug and Christoph Schneider for helpful comments and discussions. Any errors are my own.

[2] Data available at: http://www.nceo.org/library/eo_stat.html. The estimates are based on various sources including the General Social Survey of the National Opinion Research Center and a 2007 Bureau of Labor Statistics survey.

weighting captures the well-documented fact that individuals tend to overestimate small probabilities and underestimate medium to large probabilities in evaluating risky gambles.[3]

I develop a simple model of efficient pay-setting between a risk-neutral firm and an employee who is subject to probability weighting. I calibrate the model using standard parameter estimates from the experimental literature and show that it predicts that (i) broad-based employee stock option plans are more common among firms with more volatile stock prices, that (ii) the per employee number of granted stock options increases with stock price volatility and that (iii) the per employee number of stock options increases at an increasing rate with the stock price volatility. In the second part of the chapter, these predictions are tested on the universe of ExecuComp firms from 1992 to 2005. I estimate the number of options granted to lower-level employees at the individual firm level and find that all three predictions are strongly supported by the data, even after controlling for a large number of variables previously found important in the literature, and after including industry and firm effects.

Stock options are contracts with an asymmetric payoff profile. There is a substantial chance that they will expire worthless. At the same time, they offer a small chance of very large payoffs to the option holder. Employees subject to probability weighting overweight the small probabilities of large gains inherent in their stock options, which makes them attractive. At the same time, stock options impose risk on employees, which, in the absence of probability weighting, makes them unattractive. This chapter shows that for plausible assumptions about preference parameters and the distribution of future stock

[3] The seminal work on probability weighting is Kahneman and Tversky (1979) and Tversky and Kahneman (1992). Gonzales and Wu (1999), Bleichrodt and Pinto (2000), and Abdellaoui, Vossmann, and Weber (2005) confirm the inverse-S-shape of the probability weighting function, which implies overweighting of small probabilities and underweighting of medium to large probabilities. Applications of overweighting small winning probabilities include Thaler and Ziemba (1988), Cook and Clotfelter (1993), Hausch and Ziemba (1995), and Jullien and Salanie (2000). See also Camerer (2000) for a review of some of the literature.

prices, the certainty equivalent of stock options for employees can exceed the value of the option to an outside investor if employees are subject to probability weighting, but not otherwise. This is in accord with a growing body of empirical and survey evidence which documents that employees frequently value options in excess of the Black-Scholes value.[4] The economic rationale I propose for the use of stock options to compensate non-executive employees is that firms can reduce their personnel cost by granting overvalued stock options to their lower-level employees in lieu for reducing base salaries.

As an extension, I show that the model can provide a unified framework for thinking about both option grants *and* exercises. In the spirit of Benartzi and Thaler (1995), I propose a connection between the horizon over which an individual evaluates the option and the value attached to it. Myopic employees will exercise their option whenever the value of holding it for another period is less than the intrinsic value obtained by immediate exercise. The payoff distribution of options is less skewed for shorter horizons. Since overweighting of small probabilities induces an employee preference for skewness, shorter evaluation horizons increase the relative benefit of exercising early. I show in calibrations that this idea is quantitatively meaningful: employees with short evaluation horizons (one year and less) tend to exercise their options early and the tendency to exercise early increases in the moneyness of the options – consistent with stylized facts of option exercise behavior.

This chapter contributes to the compensation literature by providing a preference-based explanation for employee stock options, and by showing that stock price volatility is an important, previously neglected, driver of employee stock option grants. This paper also contributes to the growing body of research on how firms can profit from individual biases, by suggesting that stock option pay for lower-level employees is used at least in part

[4]See Lambert and Larcker (2001), Hodge, Rajgopal, and Shevlin (2006), Sawers, Wright, and Zamora (2006), Hallock and Olson (2006), and Devers, Wiseman, and Holmes (2007).

to take advantage of biased probability assessments.[5] Firms may be in a particularly good position to exploit biased non-executive employees by issuing stock options, since these employees tend to have no special financial expertise and since reduction in base salaries can be accomplished by reducing wage increases, and thus without exposing employees to perceived losses.

The context of employee stock options is particularly well suited for analyzing the effects of probability weighting. Employee stock options are asymmetric contracts in which small probabilities are associated with large gains. Hence, in contrast to analyzing symmetric payoffs, there is a clear prediction for the effect of overweighting small probabilities. Moreover, employees are prohibited from selling their options, they are only allowed to exercise them a significant period of time after the grant date, and there is effectively no outside market for the 10 year options usually awarded. Hence, in the absence of a market price for the options, preferences are likely to matter. Lastly, probability weighting is particularly relevant in the context of employee stock options because there is no timely, informative feedback on previous choices which would enable employees to "learn their way out" of distorting probabilities.[6]

Probability weighting is a key element of Tversky and Kahneman's (1992) cumulative prospect theory (CPT), and I use the CPT framework in this chapter. CPT is probably the most established alternative to expected utility theory and there is a substantial amount of experimental evidence on preference parameters, which I use in calibrating the model. Fully-fledged prospect theory models – that is: models incorporating also the central feature of probability weighting – are still rarely used in economics and finance. A well known exception is Benartzi and Thaler (1995) who propose a solution to the equity

[5]For other examples of such contractual arrangements see DellaVigna and Malmendier (2006) in the context of gym contracts and Gabaix and Laibson (2006), who show that exploitation can also survive in competitive markets.

[6]For the hypothesis that biases are reduced with experience and timely feedback see Plott (1996). For some evidence that this is relevant for probability weighting see van de Kuilen (2008).

premium puzzle based on prospect theory and myopically loss-averse agents. The few existing papers which explicitly focus on probability weighting include Barberis and Huang (2008) who show that such preferences can be consistent with the CAPM and explain why investors may hold underdiversified portfolios and prefer skewness in individual securities. Polkovnichenko (2005) shows that probability weighting is quantitatively consistent with observed household investment patterns. To the best of my knowledge, my study is one of the first to use a model with prospect theory preferences including probability weighting in the corporate finance literature.[7] It is to my knowledge the first to apply probability weighting in the domain of compensation research.

Section 2 reviews the relevant literature. The model is developed and solved in Section 3. Section 4 presents calibrations and derives testable hypotheses on option grant behavior which are confirmed empirically in a large dataset for U.S. firms in Section 5. The relation between option grants and option exercises is investigated in Section 6. Section 7 presents arguments why firms may be in a particularly good position to exploit behavioral biases of employees. Section 8 concludes.

2 Related literature on stock options for non-executive employees

A standard argument for the use of stock options for top executives is that they align the interests of shareholders and managers by providing an incentive for managers to take actions that maximize shareholder value. In a similar vein, firms might issue options to incentivize non-executive employees (Core and Guay, 2001, Kedia and Mozumdar, 2002). A serious caveat with this agency view of equity-based employee compensation is that

[7] As an illustration, the survey article on behavioral corporate finance by Baker, Ruback, and Wurgler (2006) does not even contain the term "probability weighting".

actions of lower-level employees are unlikely to move stock prices at all, let alone substantially. Stock prices are an uninformative signal of an individual employee's effort and possible free-riding is bound to dwarf any incentive effect of equity-based compensation for non-executives.[8] Incentives will thus play no role in the theoretical model developed in this chapter. The empirical evidence in this chapter, that stock option grants to employees are larger for riskier firms is clearly at odds with standard agency-theoretic models.[9]

Employee stock options could be a way to provide retention incentives, as their value is high in exactly those states of the world where demand for labor is highest (Oyer, 2004).[10] Hence, labor market conditions and industry specific factors may be first-order drivers of grant behavior. Oyer and Schaefer (2005) provide some evidence for this view. However, it is not obvious why risky instruments such as stock options should be the most efficient mechanism to provide retention incentives to risk-averse employees when alternatives such as retention bonuses, pension benefits as a function of years at the firm and the possibility of creating level-of-pay paths that benefit employees that stay with the firm, are available alternatives (see Hall and Murphy, 2003, for a more extensive discussion). I incorporate both industry volatility as a measure of labor market competition and industry fixed effects in my empirical analysis below and show that, even after controlling for these factors, the riskiness of firms still explains a large fraction of the variation of stock option grants across firms.

Stock options may be used in particular by cash-constrained firms (Yermack, 1995, Core and Guay, 2001). I include proxies for cash constraints in the empirical part of this

[8] For related claims on the ineffectiveness of incentive provision as a main driver of non-executive employee compensation, see Hall and Murphy (2003), Oyer and Schaefer (2005) and Bergman and Jenter (2007).

[9] For similar results for employee stock options see Oyer and Schaefer (2005). A standard principal-agent model that predicts less options for riskier firms is, for example, Holmström and Milgrom (1987). Aggarwal and Samwick (1999) provide empirical evidence for this model in the context of CEO stock option compensation.

[10] Inderst and Mueller (2007) use a related idea on firm-specific human capital.

chapter and find mixed evidence, which is consistent with recent results in the literature (Ittner, Lambert, and Larcker, 2003, Bergman and Jenter, 2007).

Hall and Murphy (2003) advance the hypothesis that, prior to the new laws on option expensing which came into effect in 2005, options were perceived by boards as a cheap way to remunerate employees, since they did not carry an accounting charge. Analyzing the costs and potential benefits of option compensation, Oyer and Schaefer (2006) conclude that accounting considerations are "not the sole (or even the main) driver of option grants." Since accounting rules do not differ (much) across firms and industries, the perceived cost hypothesis does not seem to be able to explain why riskier firms grant more employee stock options.

Options may also be used by some firms to attract employees with the "right mind-set," which is usually meant to describe more risk-tolerant and optimistic employees. If optimism leads employees to value their stock options higher than a well-diversified outside investor, companies can benefit from granting stock options by using the funding opportunity implicitly provided by the employees. Bergman and Jenter (2007) present qualitative support by documenting that past stock price performance, which is interpreted as a proxy for investor optimism and sentiment, is positively related to employee stock option grants. On the other hand, Oyer and Schaefer (2005) find at best limited support for the quantitative validity of this approach for typical firms in their sample. Contrary to the model developed in the present chapter and to the empirical findings presented below, existing optimism models predict a negative relation between firm volatility and option grants.[11]

There is a limited number of empirical studies on factors influencing the existence or adoption of broad-based employee stock option plans. For the U.S., using cross-sectional

[11] This is largely driven by modeling agents with concave utility functions. But even if agents were risk-neutral, overestimating the mean of future returns does not in itself generate a preference for riskier firms.

data for 1998, Oyer and Schaefer (2005) find that industry volatility is positively related to broad-based employee stock option grants and that firm volatility is insignificant when industry volatility is also included as a regressor. Ittner, Lambert, and Larcker (2003), using survey data from the years 1999 and 2000, document that new economy firms are particularly likely to grant employee stock options. Krumova and Sesil (2006) find weak evidence that broad-based plans are positively associated with sales volatility. Outside the U.S., Nagaoka (2005) and Jones, Kalmi and Mäkinen (2006) find some support for a positive relation between stock price volatility and employee stock option grants in Japan and Finland, respectively.

The dataset I use in this chapter is considerably larger than the datasets used in previous studies. Moreover, most of the previous papers have worked only on the basis of discrete indicators of broad-based stock option plans or used the total number of options granted to all employees including executives. In this chapter I use firm-specific estimates of the number of granted employee stock options and exclusively focus on non-executive employees to rule out any potentially confounding effects – in particular incentive effects – which are relevant for executives but not for non-executive employees.

3 The model

3.1 Model set-up

This section presents a simple static model in which a risk-neutral firm makes a take-it-or-leave-it offer of a pay contract, denoted by w, to a representative employee. Contract negotiations take place in $t = 0$ and the contract pays off in $t = T$. The contract is a function of the time T stock price of the company, denoted by P_T.

Employee preferences. The employee has preferences according to cumulative prospect theory (Tversky and Kahneman, 1992). For continuous probability distributions CPT preferences imply that an employee evaluates the risky payoffs from her compensation contract according to[12,13]

$$E^{\psi}\left[v\left(w\left(P_T\right)-RP\right)\right]=\int v\left(w\left(P_T\right)-RP\right)d\psi\left(F(P_T)\right). \tag{1}$$

Here, $F\left(P_T\right)$ is the cumulative distribution function of the stock price P_T. The functional in equation (1) combines two separate functions: the "value function" $v\left(\cdot\right)$ and the "probability weighting function" $\psi\left(\cdot\right)$.

The value function is given by

$$v\left(w\left(P_T\right)-RP\right)=\begin{cases}\left(w\left(P_T\right)-RP\right)^{\alpha} & , \ \textit{if} \quad w\left(P_T\right)\geq RP\\ -\lambda\left(-\left(w\left(P_T\right)-RP\right)\right)^{\alpha} & , \ \textit{if} \quad w\left(P_T\right)<RP\end{cases} \tag{2}$$

where $0<\alpha\leq 1$, and $\lambda\geq 1$.[14] It assigns a value to payoffs from the pay contract relative to a reference point RP. If the payoff is greater than the reference point, it is called a "gain", otherwise a "loss." The function is convex over losses and concave over gains, which captures diminishing sensitivity with respect to outcomes further away from the reference point. The loss aversion parameter λ governs the steepness of the function in the loss space. If $\lambda>1$, then employees dislike losses more than they are attracted by equal-sized gains.

The probability weighting function transforms cumulative probabilities into decision

[12] I use the term "value" here instead of the term "utility" to stress the distinction to standard concave von Neumann-Morgenstern utility.

[13] See also Barberis and Huang (2008), who use a similar set-up in an asset-pricing context.

[14] I restrict the slightly more general value function in Tversky and Kahmeman (1992), which uses a curvature parameter β for the loss space, by setting $\alpha=\beta$.

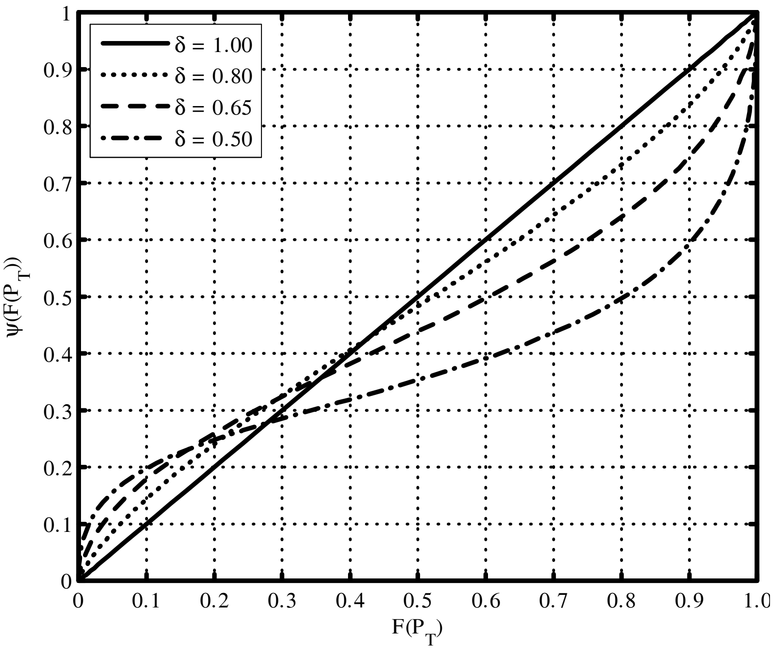

Figure 2.1: The probability weighting function as proposed by Tversky and Kahneman (1992) for different probability weighting parameters δ.

weights via the function

$$\psi\left(F(P_T)\right) = \begin{cases} \frac{-(1-F(P_T))^{\delta}}{\left(F(P_T)^{\delta}+(1-F(P_T))^{\delta}\right)^{\frac{1}{\delta}}} & , \quad if \quad w\left(P_T\right) \geq RP \\ \frac{F(P_T)^{\delta}}{\left(F(P_T)^{\delta}+(1-F(P_T))^{\delta}\right)^{\frac{1}{\delta}}} & , \quad if \quad w\left(P_T\right) < RP \end{cases} \tag{3}$$

where $0.28 < \delta \leq 1$ measures the degree of probability weighting.[15,16] Like the value function, the probability weighting scheme depends on the sign of $w\left(P_T\right) - RP$.

Figure 2.1 shows the probability weighting function over the loss space (lower branch of equation (3)).[17] It is inverse-S-shaped and intersects the 45 degree line from above – a

[15] The lower bound at 0.28 is a technical assumption to keep $\frac{\partial \psi(F(P_T))}{\partial P_T}$ positive. All experimental evidence suggests that δ is substantially above 0.28. For a more detailed discussion see Ingersoll (2008).

[16] I do not use different values of δ for gains and losses to keep the exposition of the model tractable. For a similar approach in the behavioral finance context see Barberis and Huang (2008).

[17] I only show the loss-part of the function. The intuition presented carries over straightforwardly to the gain-part as well.

functional form which has been shown to be empirically relevant in various studies.[18] The lower the weighting parameter δ, the more will small probabilities be overweighted and medium to large probabilities underweighted. For $\delta = 1$, which is the benchmark case of no probability weighting, the weighting function approaches the 45 degree line.[19]

The cost of granting stock options. I assume that there exist externalities from granting stock options such that the total cost to the firm exceeds the price of the option to an outside investor and that these external costs from granting options increase with the size of the grant. There are at least three justifications for this assumption.

First, firms engage in repurchases of shares in the open market to fund option exercises, and to counter the impact of stock option exercises on diluted earnings per share (Kahle, 2002, Bens, Nagar, Skinner, and Wong, 2003). If stock option grants get large, firms have to repurchase a large number of shares, which gets increasingly difficult and costly.

Second, there is an ownership issue associated with large employee stock option grants. For example, Intel's CEO Craig Barrett states in a filing to the SEC that: "Intel stockholders are concerned about their ownership in the company being "reduced" or "diluted" by our stock option program. If we don't take some measured action, the stockholders will not support our option plan."[20] By assuming externalities from granting stock options I make the assumption that these "measured actions" are costly.

[18] For example Tversky and Kahneman (1992), Gonzales and Wu (1999), Bleichrodt and Pinto (2000), and Abdellaoui, Vossmann, and Weber (2005).

[19] Figure 2.1 shows that the point of intersection of the weighting function with the 45 degree line shifts with δ. Thus, strictly speaking, the Tversky and Kahneman (1992) function mixes the logically distinct features of elevation of the weighting function and the curvature of the function. Prelec (1998) has proposed a weighting function which is axiomatically founded and which produces a fixed intersection point at $1/e$ (one-parameter form) and which thus separates the elevation effect from the curvature effect. Empirically, however, for typical parameter values the Tversky and Kahneman (1992) and Prelec (1998) functions are almost indistinguishable. I have recalculated all results presented with the Prelec (1998) function. All results presented continue to hold qualitatively and quantitatively when the Prelec (1998) function is used with the best fit for $\delta = 0.7$ (one-parameter specification).

[20] Available as part of a filing with the SEC at: http://www.secinfo.com/d14D5a.12dJc.htm.

Third, if a large part of employee pay is tied to stock options, underwater options may foster employee discontent and thus impair firm productivity. There is evidence that employees regard large wage reductions brought about by low payoffs from variable pay components like stock options as breach of an implicit agreement of mutual trust between company and employee, and that employees partly blame the company for their losses (Bewley, 1999). Employees may reciprocate in a variety of ways such as not putting in extra effort, spreading rumors, lowering morale of fellow workers, or even committing sabotage (Akerlof and Yellen, 1990). All this lowers firm productivity. Ex post, taking costly measures to keep morale up, for example bailing out underwater options, may be necessary. I assume that the firm will already anticipate this cost ex ante when designing a stock option program.[21]

The problem of the firm. The problem of the firm is to offer a compensation contract w such that the cost to the firm is minimized while providing the employee at least with her reservation value. The pay contract w consists of a fixed salary ϕ and n_o options with maturity T and strike price K on the company stock with random stock price P_T:

$$w\left(P_T\right) = \phi + n_o \max\left(P_T - K, 0\right).$$

The firm wants to minimize compensation costs subject to the standard participation constraint of the employee and thus offers the combination of salary and options to the

[21] This argument is likely to affect larger companies more: effective pay reductions are correlated among the firm's workers and larger firms cannot easily substitute a large number of discontent or disappointed workers in the labor market because labor supply is likely limited in most industries. My empirical results will show that this reasoning is consistent with the data: smaller firms grant more options per employee.

employee that solves:

$$\begin{aligned} \min_{n_o,\phi} \quad & E[\phi + n_o \max(P_T - K, 0)] + c(n_o) \\ s.t. \quad & E^{\psi}[v(w(P_T) - RP)] \geq v(\overline{V} - RP) \\ & n_o \geq 0. \end{aligned} \tag{4}$$

Here E is the standard expectation operator, and E^{ψ} are expectations with respect to the weighted probabilities according to equation (1).[22] $\overline{V}$ denotes the outside opportunity of the employee.[23] I assume that employees cannot write options on the firm and hence $n_o \geq 0$. Externalities from granting stock options are captured by a standard increasing and convex cost function $c(n_o)$ with $c(0) = 0$.

Reference point. We need to make an assumption about the reference point RP of the employee. Unfortunately, prospect theory is largely silent on this parameter, and while the status quo has often been used in simple settings, Kahneman and Tversky (1979) themselves note that "[...] there are situations in which gains and losses are coded relative to an expectation or aspiration level that differs from the status quo." In the absence of clear guidance from previous research, I make the following general assumption:

[22] The interest rate is set to zero in this section to simplify notation but it is included in the numerical work below.

[23] For tractability, $\overline{V}$ is assumed to be independent of the proposed contract. This is defendable if $\overline{V}$ is determined some time before the actual contract negotiations and thus predetermined. Alternatively $\overline{V}$ can be taken to be a pure cash payment. In reality, it seems plausible that employees get an idea about competitive salaries in cash equivalents from statements like: "Typically employees in industry X (and position Y etc.) can expect to get a pay package worth $\overline{V}$ dollars."

Assumption 1. *The reference point RP, over a pay-package of n_o options and a fixed salary of ϕ is linear in n_o and ϕ and has the functional form:*

$$RP = n_o\theta + \phi, \tag{5}$$

where θ is a constant with $\theta \geq 0$.

Assumption 1 is intuitive: θ represents any payoff expectation or aspiration level the employee holds for one option. This could be, for example, the Black-Scholes value, or the intrinsic option value for her best-guess future stock price. Since she gets n_o of these options, $n_o\theta$ represents the expectations on the risky part of the portfolio. Since the fixed wage ϕ is non-random, it is simply added to any expectation the employee holds on the risky part of the pay package. Consider, for example, an employee who receives 10,000 options and \$200,000 base salary over a planning horizon of four years. She anticipates that her options pay off \$5 per option in T. Hence her reference point is $10,000 \times \$5 + \$200,000 = \$250,000$.

There is also empirical support for Assumption 1. Hodge, Rajgopal, and Shevlin (2006) conduct a survey among 77 current mid-level managers and 111 future entry-level managers to analyze how employees value their stock options. Their results provide evidence that employees use simple heuristics, like subtracting the strike price from the best guess of the future stock price, as a basis for attaching a value to options. Such a reference point is a special case of Assumption 1, and I will use it as a candidate reference point in the calibrations in the next section.

3.2 Theoretical results

The model from the previous section admits an intuitive solution, which is summarized in the following proposition. I prove it in the appendix.

Proposition 1. *The optimal contract* (ϕ^*, n_o^*) *from program (4) and a reference point of the form (5) is given by*

$$(i) \quad n_o^* = 0 \quad and \quad \phi^* = \overline{V},$$

if $CE \leq E[\max(P_T - K, 0)]$, *and*

$$(ii) \quad c'(n_o^*) = CE - E[\max(P_T - K, 0)] \quad and \quad \phi^* = \overline{V} - n_o^* CE,$$

if $CE > E[\max(P_T - K, 0)]$. *The certainty equivalent the employee holds for one stock option, denoted by* CE, *is implicitly defined as*

$$E^{\psi}[v(\max(P_T - K, 0) - \theta)] = v(CE - \theta).$$

Proposition 1 states that firms grant stock options to non-executive employees if and only if the certainty equivalent of the employee for options exceeds the value of the options to an outside investor. Under risk-neutral valuation this value is the Black-Scholes value of the option. If the certainty equivalent of an option is lower, the company is better off paying the reservation wage in cash and not issuing options at all (Part (i)). Part (ii) shows that employee stock option plans are driven by employees who, in line with recent survey and empirical evidence, subjectively value options higher than outside investors.[24] Firms can exploit the biased probability assessments by replacing fixed salary worth $n_o^* CE$ by stock options which are worth less, $n_o^* E[\max(P_T - K, 0)]$, to an unbiased investor. Hence, lower-level employees in companies with broad-based employee stock option plans essentially allow firms to lower their overall personnel cost. It follows from Part (ii) that the predicted number of granted stock options increases with the difference in option value to employees and to outside investors, and that firms will grant options

[24] See references in the Introduction.

as long as the benefit from granting options exceeds option-related negative externalities captured by the cost function $c(n_o)$. The cost function thus affects how much options are granted. For the existence of stock option plans, which depends only on the sign of $CE - E[\max(P_T - K, 0)]$, the cost function is irrelevant.

Proposition 1 is stated in terms of one single option only. By virtue of the power form of the value function and the linear specification of the reference point, a scale invariance result applies (shown in the appendix). In analyzing the implications from the model, it is thus sufficient to look at the value of one single stock option for employees and outside investors, respectively. In addition, both, the optimal number of options granted, and the existence of stock option plans, are not dependent on the outside option $\overline{V}$. These properties are extremely convenient for calibrating the model and for numerically developing the predictive content of Proposition 1 in the next section.

A direct implication of the modeling assumptions as reflected in Proposition 1 is that employee stock option grants are only limited by the firm's willingness to supply more options.[25] This is consistent with anecdotal evidence of employees "clamoring" for stock options during the internet boom (Hall and Murphy, 2003) and with Bergman and Jenter (2007) who propose that equity compensation of employees is driven by "exuberant employees who *demand* to be paid in options".[26,27]

[25] Direct empirical evidence on this conjecture is scarce. The closest empirical finding to my knowledge is Sautner and Weber (2005). They analyze the behavior in an employee stock option plan of 70 high-ranking employees of a large German company between 2003 and 2004. They find that the median individual demands 100% of the options she is eligible to – consistent with the conjecture that option grants are limited by firms' supply.

[26] Bergman and Jenter (2007), p.672 (emphasis added).

[27] It is possible to model explicitly the unwillingness of employees to hold a very large number of options or similarly the unwillingness to accept pay cuts by altering the preferences or constraints imposed on the employee. Examples in the literature include augmenting the value function by a term which makes marginal utility decline as employee wealth gets sufficiently small (Gomes, 2005) or augmenting a standard concave utility function with a loss-aversion term (Barberis, Huang, and Santos, 2001). However, this would make the model less tractable and less focused, without altering anything substantial.

4 Calibration of the model

Proposition 1 predicts stock options for non-executive employees whenever the certainty equivalent for employees exceeds the value of the options to an outside investor. An advantage of modeling the underlying preferences explicitly is that it allows me to calibrate the model and investigate whether it predicts options under reasonable assumptions about preference parameters and firm characteristics. Because the experimental literature provides some guidance on values to parameterize CPT preferences, the calibration exercise below is a comparatively strict test of the validity of the model. Moreover, I can derive hypotheses on which firms are likely to grant options to non-executive employees.

4.1 Parameterizing the model

I calibrate the model by calculating, for different combinations of firm volatility and probability weighting, the ratio of the certainty equivalent of one option for an employee, to the value of the option for an outside investor. For the stock price P_T, I assume a lognormal distribution, which depends on the risk-free rate of interest r, the length of the period T, firm volatility σ and a standard normally distributed random variable u:[28,29]

$$P_T = P_0 \exp\left\{\left(r - \frac{\sigma^2}{2}\right)T + u\sigma\sqrt{T}\right\}.$$

I set r to 5% and T to 4 years.[30] Setting T to 4 years is motivated by the observation that most employees exercise most of their options shortly after they become exercisable

[28] The median dividend yield in the sample of companies analyzed below is 0.25%. I thus set dividend yields to zero in what follows. Incorporating sensible dividend yields would be straightforward and does not alter the main results.

[29] Following Dittmann and Maug (2007), I assume risk-neutral pricing throughout. This ensures that if the preferences of the employee approach risk-neutrality ($\alpha = 1$, $RP = 0$ or $\alpha = \lambda = 1$ and $\delta = 1$) the certainty equivalent of one option approaches the Black-Scholes value. This implies that all risk in the model is firm specific.

[30] These values for r and T are also used by Oyer and Schaefer (2005, 2006). Setting $T = 7$ does not alter the main results presented here.

(Huddard and Lang, 1996). In addition, Benartzi and Thaler (1995) argue that in evaluating equity portfolios, individuals routinely use a one year horizon. Since employee stock options cannot be exercised in the vesting period, it seems natural to assume that the evaluation period is extended until options become exercisable.[31] The strike price of the option, K, is equal to the grant date stock price, P_0. In this set-up, the value of one option to an outside investor is equal to the Black-Scholes value.

To parameterize the value function, I set the curvature parameter α and the coefficient of loss aversion λ to the standard values of 0.88 and 2.25, respectively. As indicated above, there is to date little research on how people set reference points for complex distributions like payoffs from stock options. I propose two candidate reference points, which are special cases of Assumption 1. The first assumes a simplified intrinsic option value calculation and is based on interview evidence reported by Hodge, Rajogopal, and Shevlin (2006). This approach suggests the reference point to be the expected future stock price less the strike price of the option. To focus on the impact of probability weighting, I assume that the stock price expectation of the employee is equal to the statistical expectation. As an alternative reference point I also consider the Black-Scholes option value with maturity equal to T.

The remaining two parameters are the volatility of the firm's stock price and the degree of probability weighting, which is captured by the parameter δ in the weighting function. Table 2.1 presents experimental results on the value of the weighting parameter δ. These estimates are relatively homogenous and suggest that values at about $\delta = 0.65$ are plausible.[32] I analyze the fit of the model for a grid of values for δ which encompasses

[31] Other common vesting schedules which stipulate the right to exercise a maximum of 25% of the options per annum over the first four years of the options life are thus assumed here to be evaluated as if all of options would become exercisable at $T = 4$. A more complex model with different time periods would have to specify an aggregation rule across time. It is thus doubtful that this would be an improvement over the current tractable model.

[32] In a large study on individual decision making, Gonzales and Wu (1999) document that there is considerable heterogeneity in probability weighting across individuals. They conclude, however, that

Table 2.1: Estimates of the parameter δ in the Tversky and Kahneman (1992) probability weighting function.

Study	**Parameter estimate**
Tversky and Kahneman (1992)	$\delta = 0.61$ (gains), $\delta = 0.69$ (losses)
Camerer and Ho (1994)	$\delta = 0.56$ (gains)
Wu and Gonzales (1996)	$\delta = 0.71$ (gains)
Abdellaoui (2000)	$\delta = 0.60$ (gains), $\delta = 0.70$ (losses)
Bleichrodt and Pinto (2000)	$\delta = 0.67$ (gains)

the most plausible values, as well as the case of no probability weighting, $\delta = 1$. I also use a grid for the volatility of the firm.

4.2 Calibration results

It is the key idea of this paper that overweighting small probabilities of large gains makes options attractive. Thus, for a given degree of probability weighting, we would expect higher evaluations of options if the underlying stock price distribution is more skewed, which is captured by firm volatility given our lognormal distributional assumption. Likewise, for a given level of firm volatility, we would expect option valuations to increase in the degree of probability weighting.

Table 2.2 shows the ratio of certainty equivalent to Black-Scholes value for a reference point equal to the Black-Scholes value (Panel A) and for the simplified intrinsic value calculation (Panel B). The certainty equivalent is calculated according to the definition in Proposition 1. In both panels, the results confirm the intuition: the more individuals overweight small probabilities (captured by δ) and the more small chances of large gains there are (captured by firm volatility), the more attractive options become. For all but the highest values of δ, the ratio of certainty equivalent to Black-Scholes value increases

the Tversky and Kahneman (1992) weighting function "...provide[s] an excellent, parsimonious fit to the median data."

Table 2.2: Calibration results. The table shows the ratio of certainty equivalent and Black-Scholes value for one option for different combinations of probability weighting and firm volatility. The model predicts employee stock option plans if this ratio exceeds one (shaded fields in the table). The calculations assume a lognormal stock price distribution with $T = 4$ years and $r = 5\%$. The strike price of the option K is set equal to the grant date stock price P_0. Preference parameters are $\alpha = 0.88$ and $\lambda = 2.25$.

Panel A: Ratio of certainty equivalent to Black-Scholes value for one option when the reference point equals the Black-Scholes value.

Probability weighting (δ = 1 implies no weighting)	Firm volatility									
	20%	*25%*	*30%*	*35%*	***40%***	*45%*	*50%*	*60%*	*70%*	*80%*
0.40	**1.37**	**1.62**	**1.91**	**2.27**	**2.70**	**3.24**	**3.90**	**5.73**	**8.57**	**12.98**
0.50	**1.10**	**1.24**	**1.40**	**1.58**	**1.80**	**2.05**	**2.35**	**3.12**	**4.20**	**5.73**
0.60	0.89	0.96	**1.03**	**1.12**	**1.22**	**1.33**	**1.46**	**1.77**	**2.19**	**2.73**
0.63	0.85	0.90	0.96	**1.03**	**1.11**	**1.20**	**1.30**	**1.55**	**1.86**	**2.28**
0.65	0.82	0.85	0.90	0.95	**1.01**	**1.08**	**1.16**	**1.35**	**1.59**	**1.90**
0.68	0.81	0.82	0.84	0.88	0.93	0.98	**1.04**	**1.18**	**1.36**	**1.59**
0.70	0.80	0.80	0.81	0.83	0.85	0.89	0.93	**1.03**	**1.17**	**1.33**
0.75	0.77	0.77	0.78	0.78	0.79	0.80	0.80	0.82	0.87	0.94
0.80	0.75	0.74	0.74	0.74	0.74	0.74	0.74	0.75	0.76	0.77
0.90	0.71	0.69	0.68	0.67	0.66	0.65	0.64	0.62	0.61	0.60
1.00	0.67	0.65	0.63	0.61	0.59	0.57	0.56	0.53	0.50	0.47

Panel B: Ratio of certainty equivalent to Black-Scholes value for one option when the reference point equals the expected intrinsic value $P_0 e^{rT} - K$.

Probability weighting (δ = 1 implies no weighting)	Firm volatility									
	20%	*25%*	*30%*	*35%*	***40%***	*45%*	*50%*	*60%*	*70%*	*80%*
0.40	**1.35**	**1.60**	**1.90**	**2.27**	**2.72**	**3.27**	**3.95**	**5.83**	**8.73**	**13.22**
0.50	**1.10**	**1.26**	**1.45**	**1.66**	**1.91**	**2.19**	**2.53**	**3.36**	**4.52**	**6.13**
0.60	0.90	**1.00**	**1.11**	**1.24**	**1.37**	**1.53**	**1.69**	**2.09**	**2.58**	**3.20**
0.63	0.86	0.95	**1.04**	**1.15**	**1.27**	**1.40**	**1.54**	**1.87**	**2.26**	**2.76**
0.65	0.83	0.90	0.98	**1.07**	**1.17**	**1.28**	**1.40**	**1.67**	**2.00**	**2.39**
0.68	0.79	0.85	0.92	1.00	**1.09**	**1.18**	**1.28**	**1.50**	**1.76**	**2.07**
0.70	0.76	0.81	0.87	0.93	**1.01**	**1.08**	**1.17**	**1.35**	**1.56**	**1.80**
0.75	0.72	0.73	0.77	0.82	0.87	0.92	0.98	**1.10**	**1.23**	**1.38**
0.80	0.70	0.67	0.69	0.72	0.75	0.78	0.82	0.90	0.97	**1.06**
0.90	0.67	0.62	0.57	0.57	0.57	0.58	0.59	0.60	0.62	0.63
1.00	0.64	0.58	0.54	0.50	0.47	0.44	0.43	0.41	0.40	0.38

at an increasing rate in the firm volatility, which, by Proposition 1 implies more options at high-volatility firms, as long as the ratio is greater than 1.

In Panel A, for $\delta = 0.65$, if the reference point equals the Black-Scholes value, no options are predicted for firms with stock price volatility less than 40%. In Panel B, when the reference point is based on the expected intrinsic value, the predicted volatility cut-off is only slightly lower. There is nothing in the model that would ex ante guarantee that it can produce any quantitatively reasonable prediction as to which firms should grant options. Its ability to produce such predictions which can be directly tested on the data is a clear strength. A cut-off level of about 40% is what I find in analyzing the universe of ExecuComp firms below.

For all specifications considered, the importance of probability weighting is striking. Without it ($\delta = 1$), the certainty equivalent is never high enough for the model to predict options, irrespective of firm volatility. To understand this, note that if the employee were risk neutral, her certainty equivalent would equal the Black-Scholes value. Hence, the values for $\delta = 1$ in Table 2.2 show that without probability weighting the employee is effectively risk averse, despite the convex part in the value function over losses. As a consequence, the relation between volatility and stock options reverses: if there is no probability weighting, the scaled certainty equivalent decreases in firm volatility, just as standard concave utility models would predict. I will show in the next section that this is actually counterfactual, which further strengthens the case for the probability weighting model.

The argument presented so far implies that the bias of overweighting small probabilities can be exploited by firms. To get an idea of the magnitude of this benefit in dollar terms, assume a typical company with 20,000 non-executive employees, which grants options with a Black-Scholes value of \$5,000 per employee annually.[33] For $\delta = 0.65$ and firm

[33] These values are the mean values for firms in the ExecuComp universe over the years 1992 to 2005

volatility of 40% the value by which reduced base salaries exceed the Black-Scholes cost of options can be calculated from Panel A of Table 2.2 to be $1 million.[34] For the reference point in Panel B, benefits are slightly higher ($17 million). Hence, for typical firms, benefits from options are small, although maybe sizeable enough to cover for expenses related with setting up and administering a broad-based plan. This changes quickly for firms with higher stock price volatility. A firm with volatility of 60% can reduce base salaries by about $50 million more than what it grants to employees in Black-Scholes value. For the largest granters of employee stock options, this value-cost differential can become enormous. Over the last decade, a company like Cisco has roughly granted per annum on average $50,000 worth of options per employee for 20,000 employees at a firm volatility of 50%. Depending on the reference point this implies a value-cost gap of between $160 and $400 million per year. To be sure, these are back-of-the-envelope calculations and have to be treated as such. They suggest clearly, however, that individual biases can have important economic consequences.

4.3 Hypotheses

The results in Table 2.2 deliver testable predictions regarding employee stock option plans. I summarize these predictions in the following three hypotheses, under the maintained assumption that employees are subject to probability weighting.

Hypothesis 1. *Firms are more likely to have a broad-based stock option plan in place if the volatility of their stock price is high.*

---

as shown in Table 2.3.

[34]This is calculated as

$$20,000 \times \$5,000 \times \left[\frac{CE}{BS} - 1\right].$$

This calculation disregards externalities from granting options because I have no good way of specifying the cost function. The benefits reported are thus upper bounds for the net "profit" from granting options to employees.

Hypothesis 2. *If a firm has a broad-based employee stock option plan, then per employee stock option grants are higher for higher volatility firms.*

Hypothesis 3. *If a firm has a broad-based employee stock option plan, then per employee stock option grants increase at an increasing rate with firm volatility.*

I test these hypotheses on the universe of ExecuComp firms in the next section.

5 Empirical tests of the model

5.1 Dataset

Firms do not have to disclose details about their stock option programs to non-executive employees, which poses a challenge for empirical research in the field. Following Desai (2003) and Bergman and Jenter (2007) I estimate the number of options granted to non-executive employees based on the ExecuComp variable "pcttotopt", which provides for each executive option grant the percentage this grant represents of the total number of options granted to all employees of the firm in the fiscal year. I average the estimates for all executives in one firm-year and eliminate outliers by dropping all firm-years for which the standard deviation of the estimates is greater than 10% of the mean. The total number of options granted by the firm in a given fiscal year thus derived is denoted by n_o^{total}.

Some papers have used what I label a "broad" definition of employees (Core and Guay, 2001, Bergman and Jenter, 2007). Under this broad definition, all individuals employed by the company, except for the top executives reported in ExecuComp, are counted as employees.[35] For typical companies, employees defined in such a way almost certainly

[35] The median company reports equity compensation for the top 5 executives (min: 1, max: 9).

include a number of employees for which incentive motives for equity compensation cannot be dismissed easily (I call them "high executives"). As I want to focus exclusively on non-executive employees for which incentive considerations are negligible, I essentially follow Oyer and Schaefer (2005) and use what I label a "narrow" definition of employees, which requires an additional assumption about how far options are spread into the organization. I assume that the number of executives increases for larger firms at a decreasing rate and I take the square root of the total number of employees as an estimate of the number of high and top executives in the firm. Hence, a company with 100 employees has an estimated 10 executives for which options could have an incentive effect, whereas for a company with 10.000 employees this is the case for 100 executives.[36] To be able to quantify the number of options to high executives I further assume, following Oyer and Schaefer (2005), that 10% of the average number of options to the top executives in the ExecuComp database excluding the CEO is awarded to the average high executive not listed in the ExecuComp database.[37] The number of options to top executives can be obtained from ExecuComp directly by summing over individual grants in the firm-year. The number of options to non-executive employees is then calculated by subtracting the number of options to top executives and the number of options to high executives from the total number of options.

I define a variable *ESOplan* which indicates whether there exists a broad-based employee stock option plan for a given firm-year. *ESOplan* is 1 if the number of non-executive employee stock options is positive and greater than 0.5% of the number of shares outstanding, and zero otherwise.

My initial sample consists of all companies in the ExecuComp database for the years

[36] Oyer and Schaefer (2005) use an estimate of the number of executives within a firm which is linear in the total number of employees. Since the total number of employees in my sample is much more dispersed, this linear estimate is likely to overstate the number of executives in large firms. For a large firm with 100,000 employees, the original Oyer and Schaefer (2005) estimate of high executives would be 10,000, whereas under the approach taken here, the estimate of high executives is 316. All results are qualitatively unchanged when using the linear estimate.

[37] All results continue to hold if this percentage is set to 5% or 20%.

1992 to 2005. All balance sheet data is taken from Compustat. I drop all companies with less than 40 employees or less than two reported executives, and winsorize firm volatility (calculated by ExecuComp based on 60 month prior stock returns), the dividend yield, and Tobin's Q (calculated as book assets minus book equity plus market value of equity all over assets) at the 1% and the 99% level. I further drop all companies in the financial sector (SIC codes 6000 to 6999) and all company-years where one of the relevant parameters for the baseline specification (Table 2.5) was missing. The resulting dataset has in total 15,005 firm-years for 2,238 unique firms. Options to non-executives were granted in 8,670 firm-years. For Table 2.3 and Table 2.4, Black-Scholes values are calculated based on the average of the grant date stock price reported in ExecuComp for all grants in a given firm-year. Option maturity and risk-free rate of interest are uniformly set to 7 years and 5%, respectively. Since my analysis is based on the number of options and not their Black-Scholes value, these assumptions are not substantial for what follows.

Table 2.3 shows descriptive statistics for the pooled sample. The median firm has 5,175 employees, a market capitalization of \$1.02 billion and sales of \$1.09 billion (Panel A). Median (mean) firm volatility is 39.0 (44.4)% and Tobin's Q is 1.60 (2.10). Panel B shows stock option plan characteristics. The majority of companies (57.8%) have a broad-based employee stock option plan in place and, for the median firm, 44.7 (74.0)% of all options granted go to employees if employee is narrowly (broadly) defined.[38] In each fiscal year, companies in the sample grant options on 1.9% to 3.2% of their shares outstanding. For the typical company, the Black-Scholes value of option grants to non-executive employees is modest, with a median per employee value of \$155. Again, the distributions are highly skewed and the per employee mean value at \$5,793 is substantially higher.[39]

[38] Hall and Murphy (2000) report that 45% of U.S. firms grant options to their exempt salaried employees in 1998. This lends some additional credibility to the estimate of 44.7% employed here using the narrow employee definition.

[39] Note, however, that there is a downward bias in these numbers if not all employees but only a subset of them receive options.

Table 2.3: Descriptive statistics. Original dataset includes all firms with more than 40 employees listed in ExecuComp over the period from 1992 to 2005. All inputs are based on ExecuComp and Compustat data. Firms in the financial sector are excluded (SIC codes 6000 to 6999). Also excluded are firm-years for which any relevant items were missing. Employees are defined "broadly" as all employees of the firm except those listed in ExecuComp. Employees are defined "narrowly" by correcting the total number of employees by the executives listed in ExecuComp and other high ranking executives. The correction is based on estimating the total number of executives in a firm by taking the square root of the total number of employees. Black-Scholes values are calculated based on the average of the grant date stock price reported in ExecuComp for all grants in a given firm-year. Maturity of the options and risk-free rate of interest is uniformly set to 7 years and 5%, respectively.

Time period		*1992 - 2005*	
Number of firms		*2,238*	
Number of firm-year observations		*15,004*	
	Mean	Median	Std. Dev.
Panel A: Firm characteristics			
Number of employees	19,087	5,175	54,885
Market value of equity (millions)	\$4,050	\$1,020	\$12,000
Sales (millions)	\$5,510	\$1,090	\$18,900
Firm volatility	44.4%	39.0%	21.4%
Tobin's Q	2.10	1.60	1.48
Total return 1 year	20.5%	11.2%	58.8%
Total return 3 years	13.2%	10.8%	28.0%
Total return 5 years	11.9%	10.9%	19.6%
Dividend yield	1.2%	0.2%	1.6%
Cash flow$_{t-1}$ / Assets$_{t-2}$	10.6%	10.9%	12.7%
Cash dividends$_{t-1}$ / Assets$_{t-2}$	1.3%	0.4%	1.9%
Cash balances$_{t-1}$ / Assets$_{t-2}$	18.5%	6.6%	29.5%
Leverage$_{t-1}$	32.8%	32.5%	25.8%
Panel B: Stock option plan characteristics			
Total granted options to shares outstanding	3.2%	1.9%	16.7%
Percentage of firms with ESO plan	57.8%	100.0%	49.4%
Percent of options to CEO	14.1%	10.9%	11.8%
Percent of options to other reported executives	15.6%	13.4%	10.6%
Percent of options to employees (broad definition)	70.3%	74.0%	19.1%
Percent of options to employees (narrow definition)	42.2%	44.7%	28.6%
BS-value of options to CEO	\$1,740,182	\$548,521	\$7,046,374
Per capita BS-value to other reported executives	\$503,492	\$185,059	\$1,395,249
Per employee BS-value to employees (broad definition)	\$6,351	\$534	\$50,077
Per employee BS-value to employees (narrow definition)	\$5,793	\$155	\$50,206

For some companies, stock option grants to non-executive employees are anything but modest. For example, one of the largest granters of employee stock options in the sample, Cisco, is estimated to grant options worth on average about \$50,000 per employee, with a total annual value of option grants to non-executive employees in excess of \$1 billion.

5.2 Empirical results

According to Hypotheses 1 to 3, the model predicts that employee stock option plans are more common among high volatility firms, that higher volatility firms grant more options per employee and that per employee stock option grants increase at an increasing rate.

Sorting firms into volatility quintiles strongly confirms all these predictions (Table 2.4). The median firm in the two lowest volatility quintiles does not have a broad-based plan, while the median firm in quintiles 3 to 5 does. Firm volatility of the median firm in quintile 3 is 39.0% and thus surprisingly close to the cut-off levels predicted by the calibration results for plausible degrees of probability weighting in Table 2.2. Moving to quintiles with higher volatility, the average per employee Black-Scholes options value increases monotonically and at an increasing rate from \$255 in the bottom quintile to \$19,472 in the top quintile. The number of options per employee increases likewise. If I do not correct for other high executives and use the broad definition of employees, I find the same monotonic relation between firm volatility, per employee Black-Scholes value and number of options, which shows that the results are not an artefact of introducing a narrow employee definition. The proportion of firms with broad-based plan increases at an increasing rate from 36.1% to 82.0%. The differences between means (and medians) of the distribution of the *ESOplan* variable between adjacent quintiles are all statistically significant at the 1% level.

Note that the increase in per employee Black-Scholes value is not mechanically caused

Table 2.4: Employee stock options grants sorted by firm volatility. Original dataset includes all firms with more than 40 employees listed in ExecuComp over the period from 1992 to 2005. All inputs are based on ExecuComp and Compustat data. Firms in the financial sector are excluded (SIC codes 6000 to 6999). Also excluded are firm-years for which any relevant items were missing. Employees are defined "broadly" as all employees of the firm except those listed in ExecuComp. Employees are defined "narrowly" by correcting the total number of employees by the executives listed in ExecuComp and other high ranking executives. The correction is based on estimating the total number of executives in a firm by taking the square root of the total number of employees. Black-Scholes values are calculated based on the average of the grant date stock price reported in ExecuComp for all grants in a given firm-year. Maturity of the options and risk-free rate of interest is uniformly set to 7 years and 5%, respectively.

	Mean								
Firm volatility quintile	Firm Volatility	Percentage of firms with ESO plan	T-test for equality with previous quintile [P-value]	*Narrow definition:* Per employee BS-value	n_O	*Broad definition:* Per employee BS-value	n_O	Sales (million)	Tobin's Q
1	21.8%	36.1%	[-]	$255	57	$425	103	$7,600	1.87
2	30.9%	47.1%	0.00	$646	97	$959	147	$5,650	1.88
3	39.3%	55.2%	0.00	$2,136	260	$2,599	325	$3,680	2.01
4	51.2%	68.8%	0.00	$6,579	751	$7,310	847	$2,130	2.23
5	79.2%	82.0%	0.00	$19,472	3,188	$20,591	3,353	$1,140	2.53

	Median								
Firm volatility quintile	Firm Volatility	ESO plan at median firm in quintile	Wilcoxon rank-sum test for equality with previous quintile [P-value]	*Narrow definition:* Per employee: BS-value	n_O	*Broad definition:* Per employee: BS-value	n_O	Sales (million)	Tobin's Q
1	22.1%	no	[-]	$0	0	$138	48	$2,790	1.55
2	30.8%	no	0.00	$0	0	$298	58	$1,580	1.53
3	39.0%	yes	0.00	$97	23	$487	81	$1,020	1.53
4	50.8%	yes	0.00	$652	107	$1,230	194	$619	1.67
5	74.5%	yes	0.00	$4,633	750	$5,461	926	$325	1.82

by using higher volatilities in the Black-Scholes formula, because together with the Black-Scholes value, the number of options per employee increases with the volatility quintiles. This is not easily reconciled with any standard concave utility model. In such a model, higher volatility would decrease the value of options to the employee since she has to be compensated for bearing additional risk. As a consequence, it would likely be optimal for the firm to substitute some of the options with cash. This would – inconsistent with the data presented here – lead to fewer options, not more.[40]

Table 2.4 shows also that high volatility firms are smaller, have a higher Tobin's Q and compensate their employees more often with stock options. This is in line with anecdotal evidence that many high-tech and new-economy firms were using stock options heavily in the dot.com era.

These insights carry over also to the multivariate case. The first set of regressions I run are OLS regressions on the subsample of firms with broad-based employee stock option plans in place according to the *ESOplan* variable. (Potential concerns about selection bias are addressed in section 5.3.) I estimate various specifications of the regression equation

$$\ln\left(1+n_{o,ikt}\right)=\alpha+\beta\cdot\sigma_{ikt}+\Gamma\cdot X_{ikt}+\lambda_t+\lambda_k+\lambda_i+\varepsilon_{ikt},$$

which predicts the log of the number of options per non-executive employee in firm i, in industry k at time t.[41] For each firm-year, firm volatility is denoted by σ_{ikt} and X_{ikt} is a vector of controls, which includes at least the log of sales to proxy for firm size and Tobin's Q to control for investment opportunities in all regressions. All else equal, firms with lower stock prices have to grant more options to grant an option package with the same Black-Scholes value. I thus use the log of the average grant-date stock prices reported in ExecuComp for all grants in a respective firm-year as an additional control. λ_t is a year dummy, λ_k is an industry dummy based on the firm's three digit SIC code and λ_i is a fixed or, depending on the specification, random firm effect. All regressions use robust standard errors that allow for clustering at the firm level.

[40] See for example Holmström and Milgrom (1987).

[41] I have also run the regressions using the log of per employee Black-Scholes value as dependent variable. The results remain basically unchanged and, if anything, get even stronger. Hence, riskier companies grant more options both on a per-employee-number-of-options basis and on a per-employee-Black-Scholes-value basis.

Table 2.5: Regressions of the log of the number of employee stock options on firm volatility and control variables. Dataset is based on all firms with more than 40 employees listed in ExecuComp over the period from 1992 to 2005. Regressions consider only the subsample of firms with employee stock option plan. Firm volatility is the 60 month stock price volatility reported by ExecuComp. Log of grant-date stock price is the log of the average grant-date stock prices reported in ExecuComp for all grants to reported executives in a given firm-year. Industry dummies are based on the three digit SIC code. Robust standard errors with clustering at the firm level are reported in parentheses. Number of observations: 8,669.

	Dependent variable: Log of employee stock options per employee for firms with ESO plan					
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Firm volatility	**2.88** ***	**1.84** ***	**1.01** ***	**0.56** ***		
	(0.15)	(0.14)	(0.11)	(0.13)		
Volatility quintile 1					-	-
Volatility quintile 2					**0.12** **	**0.06** *
					(0.05)	(0.04)
Volatility quintile 3					**0.25** ***	**0.10** **
					(0.06)	(0.05)
Volatility quintile 4					**0.51** ***	**0.19** ***
					(0.07)	(0.06)
Volatility quintile 5					**0.97** ***	**0.32** ***
					(0.08)	(0.07)
Log of grant-date stock price	**-0.13** ***	**-0.30** ***	**-0.28** ***	**-0.29** ***	**-0.35** ***	**-0.30** ***
	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)
Log of sales	**-0.27** ***	**-0.22** ***	**-0.24** ***	**-0.19** ***	**-0.23** ***	**-0.19** ***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Tobin's Q	**0.28** ***	**0.24** ***	**0.10** ***	**0.07** ***	**0.25** ***	**0.07** ***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Firm fixed effects				Yes		Yes
Firm random effects			Yes			
Industry dummies		Yes			Yes	
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
(Adjusted) R^2	0.484	0.689	0.672	0.219	0.684	0.219

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

Results are presented in Table 2.5. All coefficients have the expected signs and are highly statistically significant. The impact of firm volatility on option grants is also economically significant: increasing firm volatility by one standard deviation increases the number of options granted per non-executive employee by about 40% when controlling for year and industry effects (three digit SIC code). The positive relation between volatility and options is weaker, but still highly significant in specifications with random and fixed effects (specifications (3) and (4)). Specifications (5) and (6) replace firm volatility by firm volatility quintiles. As predicted, employee stock options increase at an increasing rate and the effect of moving from quintile four to quintile five is almost as large as moving from quintile one to quintile four.

To make sure my results are robust, I include additional control variables that have been shown to be related to employee stock option grants. I include total returns to shareholders over one, three and five firm-years as reported by ExecuComp. I also include a dummy for new economy firms (SIC codes 3570-3579, 3661, 3674, 5045, 5961 and 7370-7379), since there is evidence that options are used to a larger extent in new economy firms (Ittner, Lambert, and Larcker, 2003). Following Oyer and Schaefer (2005), I use the average volatility of firms in the same three digit SIC industry, weighted by firm assets, as a proxy for labor market conditions. I also include the three year average dividend yield as reported in ExecuComp. I control for cash constraints by including cash flow (Compustat data items 14 + 18), cash dividends (data items 19 + 21), cash balances (data item 1), all over lagged assets, and leverage ((data items 9 + 34) / (data items 9 + 34 + 216)).[42] All of these controls for cash constraints are winsorized at the 1% and 99% level.[43]

Table 2.6 shows that the results for firm volatility are qualitatively unchanged when adding these additional controls. There is strong support for higher option grants in new

[42] These are the constituents of a measure of cash constraints based on work by Kaplan and Zingales (1997), Lamont, Polk, and Saa-Requiejo (2001) and Baker, Stein, and Wurgler (2003) which is also used in Bergman and Jenter (2007). Using the Kaplan-Zingales-index instead of the constituents leaves all main results unchanged.

[43] I have also used cash flow and capital expenditure over lagged assets (as in Oyer and Schaefer, 2005), as well as interest burden and cash flow shortfall (as in Core and Guay, 2001) as measures of cash constraints. Results concerning the influence of cash constraints on option grants are weaker under these measures, while the main results relating to firm volatility remain basically unchanged.

Table 2.6: Regressions of the log of the number of employee stock options on firm volatility and enlarged set of control variables. Total returns to shareholders include the monthly reinvestment of dividends and are based on the fiscal year. Some observations are lost because "Total return 5 years" is not available. New economy firms are firms in industries with SIC codes 3570-3579, 3661, 3674, 5045, 5961 and 7370-7379. Dividend yield is a three year average. Industry volatility is calculated based on the average firm volatilities for firms within the same three digit SIC industry, weighted by firm assets. Cash flow, cash dividends and cash balances are scaled by lagged assets. All other variables as in Table 2.5. Industry dummies are based on the three digit SIC code. Robust standard errors with clustering at the firm level are reported in parentheses. All regressions include year dummies. Number of observations: 7,112.

	Dependent variable: Log of employee stock options per employee for firms with ESO plan					
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Firm volatility	**1.79** ***	**1.24** ***	**0.89** ***	**0.57** ***		
	(0.18)	(0.15)	(0.14)	(0.16)		
Volatility quintile 2					**0.10** **	**0.03**
					(0.05)	(0.04)
Volatility quintile 3					**0.23** ***	**0.06**
					(0.06)	(0.05)
Volatility quintile 4					**0.34** ***	**0.14** **
					(0.07)	(0.06)
Volatility quintile 5					**0.59** ***	**0.33** ***
					(0.08)	(0.07)
Log of grant-date stock price	**-0.18** ***	**-0.43** ***	**-0.34** ***	**-0.40** ***	**-0.46** ***	**-0.36** ***
	(0.05)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)
Log of sales	**-0.21** ***	**-0.11** ***	**-0.22** ***	**-0.15** ***	**-0.11** ***	**-0.24** ***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Tobin's Q	**0.27** ***	**0.28** ***	**0.16** ***	**0.12** ***	**0.28** ***	**0.16** ***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Total return 1 year	**-0.10** ***	**-0.15** ***	**-0.14** ***	**-0.14** ***	**-0.15** ***	**-0.15** ***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)
Total return 3 years	**-0.66** ***	**-0.48** ***	**-0.30** ***	**-0.20** ***	**-0.46** ***	**-0.28** ***
	(0.08)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)
Total return 5 years	**0.12**	**0.13**	**0.26** ***	**0.30** ***	**0.11**	**0.25** ***
	(0.12)	(0.10)	(0.07)	(0.07)	(0.10)	(0.07)
New economy (dummy)	**0.90** ***	**0.93** ***	**1.20** ***		**0.92** ***	**1.24** ***
	(0.08)	(0.16)	(0.08)		(0.16)	(0.08)
Dividend yield	**-8.17** ***	**-13.16** ***	**-4.74** ***	**-4.63** **	**-13.54** ***	**-5.15** ***
	(2.38)	(1.90)	(1.55)	(1.87)	(1.89)	(1.57)
Industry volatility	**-1.52** ***	**-0.13**	**-0.16**	**0.33**	**0.12**	**-0.02**
	(0.27)	(0.26)	(0.19)	(0.21)	(0.25)	(0.18)
Cash flow$_{t-1}$ / Assets$_{t-2}$	**-0.67** ***	**-0.52** ***	**-0.30** ***	**-0.12**	**-0.65** ***	**-0.33** ***
	(0.21)	(0.16)	(0.10)	(0.11)	(0.16)	(0.10)
Cash dividends$_{t-1}$ / Assets$_{t-2}$	**4.87** ***	**4.06** ***	**3.98** ***	**3.86** ***	**4.70** ***	**4.12** ***
	(1.80)	(1.33)	(1.02)	(1.14)	(1.37)	(1.03)
Cash balances$_{t-1}$ / Assets$_{t-2}$	**1.13** ***	**1.03** ***	**0.44** ***	**0.29** ***	**1.08** ***	**0.46** ***
	(0.10)	(0.08)	(0.05)	(0.05)	(0.08)	(0.05)
Leverage$_{t-1}$	**-0.07**	**-0.63** ***	**-0.25** ***	**-0.25** ***	**-0.61** ***	**-0.23** ***
	(0.12)	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)
Firm effects			R.E.	F.E.		R.E.
Industry dummies		Yes			Yes	
(Adjusted) R^2	0.557	0.734	0.517	0.290	0.731	0.512

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

economy firms as the new economy dummy variable is highly significant in all specifications. There is no clear evidence of the relevance of cash constraints. On the one hand, cash flow is negatively related to options grants, which is expected under the cash constraint hypothesis. On the other hand, firms with high cash dividends, high cash balances and low leverage (and hence low interest payments) grant more options, which seems to contradict the view that firms grant options when they are short of cash.

The findings on industry volatility are mixed and support the predictions of the retention model only when firm fixed effects are included and industry is not otherwise controlled for. The reduction in the coefficient of volatility after including industry effects may, however, capture effects related to labor market conditions within the industry. The coefficient on past one and three year stock returns are all highly significant and negative. This contrasts explanations for employee stock option use based on employee sentiment, because these models usually assume trend extrapolation, which would predict a positive coefficient on past returns (Bergman and Jenter, 2007).

In sum, the results in Tables 2.5 and 2.6 strongly support Hypotheses 2 and 3 and show that riskier firms grant more employee stock options.[44]

To investigate whether higher firm volatility increases the probability of broad-based employee stock option plans at firms (Hypothesis 1), I estimate a linear probability model of the form

$$\Pr(ESOplan = 1) = \alpha + \beta \cdot \sigma_{ikt} + \Gamma \cdot X_{ikt} + \lambda_t + \lambda_k + \mu_i + \varepsilon_{ikt},$$

where μ_i is a random firm effect. The dependent variable is now *ESOplan*, an indicator variable which is one if there is a broad-based plan (see section 5.1 for the construction

[44] In untabulated results I also find evidence that stock option grants were more sensitive to firm risk in the years up to and including 2000 – a period where success stories of firms and investors were everywhere and where the desire to "get rich quick" and therefore the tendency to overweight small probabilities of large gains seems particularly plausible.

Table 2.7: Regressions of an indicator variable for the existence of a broad-based employee stock option plan on firm volatility and control variables. ESOplan is equal to one if there is a broad-based stock option plan at the firm in the respective firm-year. Industry dummies are based on the three digit SIC code. For the probit model, marginal effects computed at the mean are reported. Some observations for the probit model are lost because of no within-industry variation in the ESOplan variable. For the random effects probit model, McFadden's R^2 is reported. Robust standard errors with clustering at the firm level are given in parentheses.

	Dependent variable: ESOplan (dummy variable)				
	Linear Probability Model			Probit Model	
Independent variable	(1)	(2)	(3)	(4)	(5)
Firm volatility	**0.47** ***	**0.18** ***	**0.25** ***	**0.31** ***	**1.28** ***
	(0.04)	(0.04)	(0.03)	(0.05)	(0.13)
Log of sales	**-0.04** ***	**-0.03** ***	**-0.06** ***	**-0.04** ***	**-0.26** ***
	(0.01)	(0.00)	(0.00)	(0.01)	(0.02)
Tobin's Q	**0.05** ***	**0.01** ***	**0.01** ***	**0.03** ***	**0.10** ***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.02)
Firm random effects			Yes		Yes
Industry dummies		Yes		Yes	
Year dummies	Yes	Yes	Yes	Yes	Yes
Percent correctly predicted	0.670	0.734	0.647	0.638	0.638
(Adjusted or Pseudo) R^2	0.141	0.266	0.127	0.228	0.175
N	15,004	15,004	15,004	14,881	15,004

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

of this variable). The covariates I consider are the same as in the previous regressions, except for the stock price, which I drop from the set of control variables since it only determines the number of options granted but not whether or not any options are granted at all.

Table 2.7 shows that Hypothesis 1 is borne out by the data: high volatility firms are more likely to have a broad employee stock option plan. This finding is robust to including year, industry and firm effects. The signs on the control variables are as expected and robust across specifications. To provide additional support to my linear specification, I also report results from a corresponding (random effects) probit model.[45] The marginal effect of an increase in firm volatility is lower using the linear framework: increasing firm volatility by one standard deviation increases the probability of a broad-based plan by between 3.5% and 10% in the linear probability model, while the probit model predicts changes of between 6.5% and 27%.

Table 2.8 shows the results for the larger set of control variables. The coefficients and significance levels of firm volatility, log of sales and Tobin's Q are qualitatively unchanged. Again, firms with larger past returns grant fewer employee stock options. New economy firms grant significantly more options. There is no support for cash constraints influencing grant behavior and the only significant coefficient among the variables to proxy for cash constraints is on cash balances. This coefficient is positive, which implies more options at firms with higher cash balances.

Overall, the results of this section show that the simple model of pay negotiations between a firm and an employee with CPT preferences generates predictions that are consistent with observed patterns of employees stock option plans.

5.3 Robustness checks

The strong support for Hypotheses 2 and 3 in the data is based on OLS regressions on the subsample of firms that grant employee stock options. This may introduce sample selection bias. To investigate how severe this bias is, I run two alternative regressions. The first is a Tobit model on the full sample of firms and the second is a Heckman two-stage selection model.[46] Results are presented in Table 2.9. For both models, the impact of

[45] The probit model might potentially suffer from the incidental parameter problem, which is why I compute the linear specification as a benchmark. The results here suggest, however, that the incidental parameter problem is not severe.

[46] The first stage probit regressions in the Heckman model includes all variables used in Table 2.6 with industry Q (weighted average (by assets) of Tobin's Q for firms in the same three digit SIC industry) as

Table 2.8: Regressions of an indicator variable for the existence of a broad-based employee stock option plan on firm volatility and enlarged set of control variables. All variables are defined as in Table 2.7. For the probit model, marginal effects computed at the mean are reported. Some observations for the probit model are lost because of no within-industry variation in the ESOplan variable. For the random effects probit model, McFadden's R^2 is reported. Robust standard errors with clustering at the firm level are given in parentheses.

	Dependent variable: ESOplan (dummy variable)				
	Linear Probability Model			Probit Model	
Independent variable	(1)	(2)	(3)	(4)	(5)
Firm volatility	**0.16** ***	**0.15** ***	**0.10** **	**0.22** ***	**0.56** ***
	(0.05)	(0.05)	(0.04)	(0.07)	(0.17)
Log of sales	**-0.02** ***	**-0.02** ***	**-0.04** ***	**-0.02** ***	**-0.17** ***
	(0.01)	(0.01)	(0.00)	(0.01)	(0.02)
Tobin's Q	**0.04** ***	**0.02** ***	**0.02** ***	**0.03** ***	**0.11** ***
	(0.01)	(0.01)	(0.00)	(0.01)	(0.02)
Total return 1 year	**0.00**	**0.01**	**0.00**	**0.01**	**0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)
Total return 3 years	**-0.08** ***	**-0.06** ***	**-0.06** ***	**-0.09** ***	**-0.27** ***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.10)
Total return 5 years	**-0.12** ***	**-0.11** ***	**-0.12** ***	**-0.13** ***	**-0.49** ***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.14)
New economy (dummy)	**0.24** ***	**0.17** ***	**0.27** ***	**0.31** ***	**1.50** ***
	(0.02)	(0.04)	(0.02)	(0.05)	(0.12)
Dividend yield	**-3.71** ***	**-1.86** ***	**-2.43** ***	**-1.76** **	**-8.16** ***
	(0.58)	(0.66)	(0.51)	(0.78)	(1.88)
Industry volatility	**-0.05**	**-0.16** *	**-0.08**	**-0.17**	**-0.28**
	(0.08)	(0.08)	(0.06)	(0.11)	(0.25)
Cash flow_{t-1} / Assets_{t-2}	**-0.01**	**0.10** **	**0.03**	**0.04**	**-0.08**
	(0.05)	(0.05)	(0.04)	(0.08)	(0.23)
Cash dividends_{t-1} / Assets_{t-2}	**0.48**	**-0.51**	**-0.51**	**-0.80**	**-1.91**
	(0.52)	(0.47)	(0.41)	(0.62)	(1.61)
Cash balances_{t-1} / Assets_{t-2}	**0.16** ***	**0.08** ***	**0.08** ***	**0.23** ***	**0.68** ***
	(0.03)	(0.02)	(0.02)	(0.05)	(0.14)
Leverage_{t-1}	**-0.05** *	**-0.03**	**-0.03**	**-0.04**	**-0.16** *
	(0.03)	(0.03)	(0.03)	(0.04)	(0.10)
Firm random effects			Yes		Yes
Industry dummies		Yes		Yes	
Year dummies	Yes	Yes	Yes	Yes	Yes
Percent correctly predicted	0.672	0.731	0.668	0.697	0.638
(Adjusted or Pseudo) R^2	0.177	0.270	0.171	0.233	0.161
N	12,843	12,843	12,843	12,719	12,843

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

firm volatility is larger than under the OLS specification based on the subsample and the Heckman selection model shows that the OLS results presented earlier tend to understate the positive relation between firm volatility and broad-based stock option plans.

A clear limitation of this analysis is that the characteristics of employee stock option plans are estimated and not actually observed. Moreover, companies that do not issue options to the top five executives are not in the ExecuComp database and thus not part of the sample, which could lead to biased results. Bergman and Jenter (2007) who also use the estimation based on "pcttotopt" in ExecuComp perform a robustness-check of their results based on a hand-collected dataset by Core and Guay (2001). They find no evidence for systematic biases and their results are usually even stronger using the hand-collected employee stock option data.

6 The relation between stock option grants and exercises

So far I have documented that a simple model based on employees who have CPT preferences generates predictions which are surprisingly consistent with the data. In this section I argue that such a model has the potential to provide a unified framework for thinking about both stock option grants and exercises.

The key idea is that, as a default, individuals evaluate investment decisions over short horizons (they are "myopic"). This builds on the work by Benartzi and Thaler (1995), who argue that for a typical portfolio of stocks and bonds the relevant horizon is about one year. Heath, Huddard, and Lang (1999) and Odean (1998) find that option exercises are significantly related to short-term stock price run-ups, which suggests that for stock options even shorter horizons than one year may be relevant. Typical vesting schedules

an additional regressor.

Table 2.9: Robustness checks. Specifications (1) and (2) show the results from a Tobit regression of the log of the number of stock options per employee on the set of control variables used in Tables 2.5 and 2.6. Specifications (3) and (4) show the results from the second stage regression of the Heckman selection model. The first stage probit regression in the Heckman selection model includes as additional regressor the industry Q as a weighted average (by assets) of Tobin's Q for firms in the same three digit SIC industry. Industry dummies are based on the three digit SIC code. Robust standard errors with clustering at the firm level are given in parentheses.

	Dependent variable: Log of employee stock options per employee			
	Tobit		Heckman Selection	
Independent variable	(1)	(2)	(3)	(4)
Firm volatility	**2.90** ***	**2.24** ***	**1.94** ***	**1.67** ***
	(0.36)	(0.45)	(0.17)	(0.18)
Log of grant-date stock price	**-0.25** ***	**-0.30** ***	**-0.19** ***	**-0.18** ***
	(0.08)	(0.10)	(0.05)	(0.05)
Log of sales	**-0.41** ***	**-0.20** ***	**-0.20** ***	**-0.21** ***
	(0.05)	(0.06)	(0.02)	(0.03)
Tobin's Q	**0.30** ***	**0.35** ***	**0.25** ***	**0.25** ***
	(0.03)	(0.05)	(0.02)	(0.02)
Total return 1 year		**-0.04**		**-0.09** ***
		(0.07)		(0.03)
Total return 3 years		**-0.83** ***		**-0.64** ***
		(0.18)		(0.08)
Total return 5 years		**-0.85** ***		**0.16**
		(0.28)		(0.12)
New economy (dummy)		**2.13** ***		**0.86** ***
		(0.37)		(0.09)
Dividend yield		**-25.68** ***		**-7.76** ***
		(6.40)		(2.55)
Industry volatility		**-1.55** **		**-1.32** ***
		(0.73)		(0.27)
Cash flow_{t-1} / Assets_{t-2}		**0.13**		**-0.67** ***
		(0.44)		(0.21)
Cash dividends_{t-1} / Assets_{t-2}		**-1.35**		**4.99** ***
		(4.29)		(1.77)
Cash balances_{t-1} / Assets_{t-2}		**1.61** ***		**1.12** ***
		(0.20)		(0.10)
Leverage_{t-1}		**-0.75** ***		**-0.07**
		(0.25)		(0.12)
Inverse Mills ratio			**-0.96** ***	**-0.07**
			(0.05)	(0.12)
ρ			**-0.69** ***	**-0.06**
			(0.03)	(0.10)
Industry dummies	Yes	Yes		
Year dummies	Yes	Yes	Yes	Yes
(Pseudo) R^2	0.129	0.252	0.296	0.354
N	15,004	12,843	12,843	12,843

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

preclude exercising options for a period of several years and it thus seems reasonable to assume that the evaluation horizon is extended accordingly. Once the options are vested, however, the shorter "default"-horizon becomes relevant again. I argue that the shorter this horizon, the more likely is an option exercise, consistent with empirical studies that find that employee stock options are usually exercised quickly after the vesting date (Huddard and Lang, 1996).

To fix ideas let the grant date be $T = 0$, let T_1 be the vesting date and T_2 be called the horizon date. In T_1 the employee decides on whether or not to exercise the options. If she is myopic in the sense of Benartzi and Thaler (1995), she will base this decision on the possible payoffs from exercising the stock options in T_2, which are dependent on the stock price P_{T_2} given by

$$P_{T_2} = P_{T_1} \exp\left\{ \left(r - \frac{\sigma^2}{2} \right) (T_2 - T_1) + u\sigma\sqrt{T_2 - T_1} \right\}.$$

She will exercise in T_1 if the payoff from exercising, $P_{T_1} - K$, is positive and greater than the certainty equivalent for holding the options until T_2, which is implicitly defined by

$$E^{\psi}\left[v\left(\max\left(P_{T_2} - K, 0 \right) - RP \right) \right] = v\left(CE - RP \right).$$

The intuition is now that the longer the option is held, the more skewed the payoff distribution will become. Since the employee overweights small probabilities of large gains, this tends to increase the certainty equivalent and hence decreases the probability of an option exercise in T_1.

I again test the intuition by calibrating a simple benchmark model. I assume that $T_1 = 4$ and that the stock price at P_0 has increased to $P_{T_1} = P_0 e^{rT_1}$, the expected value. The option is thus in the money as $K = P_0$ and so $P_0 e^{rT_1} - K > 0$. The reference point of the employee is denoted, without loss of generality, by $RP = P_{T_1} - K + \theta$, where θ is any number with $\theta > K - P_{T_1}$. I assume $\theta = P_{T_1}\left(e^{rT_2} - 1\right)$ in the calibrations which implies that the employee's best guess about the stock price in T_2 is the expected value as seen from time T_1. I report results only for a degree of probability weighting of $\delta = 0.65$. All other parameters are the same as in Section 4.

Table 2.10 Panel A shows that the intuition is borne out by the model. For all levels of firm volatility the ratio of certainty equivalent to intrinsic value at time T_1 is strictly increasing in T_2. A ratio smaller than one indicates option exercise. Hence, for evaluation horizons smaller than a year, the model predicts exercises for all volatilities. The model also generates another plausible result: the more the options are in the money in T_1, i.e. the higher P_{T_1} relative to the strike price, the more likely is an exercise decision (Panel B). Intuitively, a higher stock price at the vesting date ceteris paribus increases the reference point for the option payoff at the horizon date, which implies that more option payoffs fall into the loss space. Hence, a lower ratio of actual stock price to strike price tends to make options unattractive to employees with CPT preferences.

Heath, Huddard, and Lang (1999) have documented that empirical exercise behavior of employees is sensitive to reference points, most notably whether or not the stock price exceeds the 52-week high stock price. They also argue that prospect theory is largely consistent with their findings. While a truly dynamic CPT model that could integrate such reference point effects is yet unavailable, the results presented here on stock option grants and exercises and the complementary work by Heath, Huddard, and Lang (1999) suggest that prospect theory has the potential to explain individual behavior in stock option programs in a unified framework.

The CPT model predicts that employees from riskier firms should be less likely to exercise their options early. This appears to be in contrast to findings from Bettis, Bizjak, and Lemmon (2005) and Huddard and Lang (1996). The study by Bettis, Bizjak, and

Table 2.10: The influence of the evaluation horizon and the moneyness of options on exercise decisions. Panels A and B show the ratio of certainty equivalent when holding the option to the intrinsic value obtained by exercising. The option is not exercised if this value is greater than one (shaded cells). The evaluation horizon is $T_2 - T_1$. Panel A assumes $P_{T_1} = P_0 e^{rT_1}$. For Panel B an evaluation horizon of 6 months is assumed. The calculations use a lognormal stock price distribution with $T_1 = 4$ years and $r = 5\%$. The strike price of the option K is set equal to the grant date stock price P_0. Preference parameters are $\alpha = 0.88$, $\lambda = 2.25$ and $\delta = 0.65$. The reference point is taken to be equal to the statistically expected value of P_{T_2}.

Panel A: Influence of the evaluation horizon on exercises of employee stock options.

Evaluation horizon (in years)	Firm volatility: *20%*	*25%*	*30%*	*35%*	***40%***	*45%*	*50%*	*60%*	*70%*	*80%*
0.10	0.23	0.29	0.34	0.38	0.42	0.45	0.48	0.54	0.59	0.63
0.25	0.37	0.44	0.49	0.53	0.58	0.61	0.65	0.71	0.77	0.83
0.50	0.51	0.57	0.63	0.68	0.73	0.77	0.81	0.89	0.97	**1.05**
0.75	0.60	0.67	0.73	0.78	0.84	0.89	0.94	**1.03**	**1.12**	**1.25**
1.00	0.68	0.75	0.81	0.87	0.93	0.99	**1.05**	**1.15**	**1.28**	**1.56**
2.00	0.93	**1.01**	**1.09**	**1.17**	**1.25**	**1.33**	**1.40**	**1.72**	**2.23**	**2.86**
4.00	**1.32**	**1.43**	**1.54**	**1.65**	**1.77**	**1.93**	**2.25**	**3.11**	**4.25**	**5.70**

Panel B: Influence of the moneyness of the option on exercises of employee stock options.

Ratio P_{T1} / K [%]	Firm volatility: *20%*	*25%*	*30%*	*35%*	***40%***	*45%*	*50%*	*60%*	*70%*	*80%*
105	**1.30**	**1.42**	**1.55**	**1.84**	**2.17**	**2.51**	**2.88**	**3.67**	**4.53**	**5.46**
110	0.84	0.92	**1.00**	**1.08**	**1.15**	**1.21**	**1.29**	**1.61**	**2.01**	**2.44**
120	0.54	0.61	0.67	0.72	0.77	0.82	0.86	0.95	**1.03**	**1.10**
130	0.40	0.47	0.52	0.57	0.61	0.65	0.68	0.75	0.82	0.89
150	0.25	0.31	0.36	0.40	0.44	0.48	0.51	0.56	0.61	0.66
200	0.13	0.16	0.19	0.22	0.25	0.28	0.31	0.36	0.40	0.44

Lemmon (2005), however, focuses on executives, which are explicitly not the focus of this study. Since it is likely that top executives and rank-and-file employees differ along many dimensions (financial literacy, expertise in assessing risks etc.), it is not clear that the results for executives carry over to non-executives. Huddard and Lang (1996) find that across their seven firms, exercises are positively related to firm volatility. Closer inspection reveals, however, that four out of seven firms in their sample have coefficients in regressions which predict a negative relation, as suggested by the present model. These coefficients are significant for two of these companies – companies which also happen to be the most volatile in the sample. It would be valuable to see results for non-executive employees on a large sample basis to accurately assess the predictions of the CPT model with respect to exercises.

7 Are firms in a special position to exploit the bias?

Firms that grant stock options do so because – as employees overvalue small probabilities of large gains – they can reduce base salaries by more than one for one. Of course, in principle, anybody can offer skewed payoffs to individuals. Indeed, lottery tickets, long-shot race-track betting or individual investments into risky option portfolios at online brokers are examples of lottery companies, bookmakers and brokers profiting from individuals desire to "hit the jackpot".[47] There are good reasons to believe, however, that firms are in a particularly good position to exploit the bias of probability weighting.

First, there is ample evidence that individuals are greatly overinvested in their own company stock in their retirement savings plans. Some authors attribute this to a default bias (Carroll, Choi, Laibson, Madrian, and Metrick, 2005). In the present case, if the default setting in a pay contract is that it includes options, individuals are likely to be reluctant to exchange the default (options) against an alternative (no options).

A second reason lies in the well-documented psychological phenomenon that individuals like to bet on things they feel confident and knowledgeable about (Heath and Tversky, 1991, Keppe and Weber, 1995). Moreover, they may adopt the so called "insider-view"

[47] For literature that links probability weighting with these phenomena, see for example, Cook and Clotfelter (1993), Hausch and Ziemba (1995) and Jullien and Salanie (2000).

(Kahneman and Lovallo, 1993), which is a tendency to favorably judge the likely success of a project if one is directly involved in it. In the present context, an employee may overweight the possibility of her stock options paying of a large amount because she is an insider in the firm, and she may focus on her employer's stock options as opposed to other skewed gambles, since she feels especially knowledgeable about her own company.

Lastly, firms have the opportunity to exchange stock options for *future pay increases*. While from a classical economic perspective reducing salary increases and cutting base salary by the same amount are the same thing, there is evidence that individuals are much more sensitive to the latter (Kahneman, Knetsch, and Thaler, 1986, Bewley, 1999). This can also explain why it may be hard for an investment bank to step in and offer stock options on a company's stock to the company's employees: the employees would have to pay cash to the bank and thus suffer a nominal loss of cash. The firm, on the other hand, can cut real wages by reducing nominal pay increases – a small procedural change that can have profound impact on the perceived attractiveness of an offered prospect.

8 Conclusion

In this chapter I show empirically, using a sample of over 2,200 U.S. firms over the years 1992 to 2005, that firms that are small, have good growth opportunities and high stock price volatility grant more stock options to their non-executive employees. The finding that higher firm volatility is associated with more options is at odds with standard agency models of compensation. The results are robust to including industry effects, showing that labor market competition or special circumstances in new economy firms are not sufficient to explain broad-based employee stock option compensation.

A model in which risk-neutral firms bargain with employees with cumulative prospect theory preferences can explain the empirical findings remarkably well. The key intuition is that probability weighting, and in particular the tendency of individuals to overweight small probabilities of large gains, makes options attractive since they come with a highly skewed payoff distribution of bounded losses and unbounded low-probability upside. This intuition is shown to be consistent with the data when the model is calibrated using parameter values from the experimental literature. The bargaining model with CPT employees predicts that (i) riskier firms are more likely to have a broad-based employee stock option plan in place (ii) per employee number of granted stock options increases with firm volatility and (iii) the rate of this increase is increasing. All hypotheses are strongly supported in the data. An attractive feature of the model is that it avoids specifying an ad hoc bias to explain a perceived empirical anomaly. Instead, it is a straightforward application of cumulative prospect theory, the most firmly established alternative to expected utility models to date.

There are two major implications: first, the model implies that firms can use employees as a source of funds. This constitutes another example of firms contractually exploiting an individual bias in decision making. Given the importance of pay contracts to basically anyone, and given the estimates of transfers of funds from employees to some firms in the range of tens of millions of dollars annually, this example shows that exploiting behavioral biases may be more widely spread and more economically profitable than commonly thought. Second, I show that the model can provide a unified framework for thinking about both employee stock option grants and stock option exercises. The impact of probability weighting is positively correlated with the evaluation horizon of individuals. The vesting period artificially lengthens the evaluation period. After options vest, individuals are likely to resort to default evaluation horizons which are usually short. Over short horizons the impact of reference points increases relative to the impact of probability weighting. As a result, for typical parameter values the model predicts early exercise

– consistent with existing empirical evidence.

The model and results in this paper suggest a number of promising avenues for future research. It may be interesting, for example, to explore whether employees sort into companies based on their degree of probability weighting, and investigate the strategic implications of such a sorting mechanism. On a more general level, probability weighting could have a profound impact on firm policies if CEOs are also subject to the bias. Investigating the relevance and implications of probability weighting for the interaction of firms and markets is an important task for future research.

Appendix

A Proof of Proposition 1

In the main text, the interest rate was set to zero to simplify the exposition. In this appendix, I incorporate interest rates. In order to prove Proposition 1, the following two Lemmas will be useful.

Lemma 1. *The prospect value of the contract (n_o, ϕ_o) does not depend on the base salary received and is homogenous of degree α in the number of options n_o if the reference point is given by $RP = n_o\theta + \phi e^{rT}$ (Assumption 1).*

Proof.

$$\begin{aligned} E^{\psi}(n_o, \phi) &\equiv E^{\psi}\left[v\left(n_o \max(P_T - K, 0) + \phi e^{rT} - RP\right)\right] \\ &= E^{\psi}\left[v\left(n_o\left(\max(P_T - K, 0) - \theta\right)\right)\right] \\ &= E^{\psi}(n_o), \end{aligned}$$

where the second equality follows from using the definition of the reference point in Assumption 1. This proves the first part of the Lemma. To prove the second part note that

$$
\begin{aligned}
E^{\psi}\left(n_o\right) = &-\lambda \int_0^{\theta+K} \left(-\left(n_o\left(\max\left(P_T-K,0\right)-\theta\right)\right)\right)^{\alpha} d\psi\left(F(P_T)\right) \\
&+ \int_{\theta+K}^{\infty} \left(n_o\left(\max\left(P_T-K,0\right)-\theta\right)\right)^{\alpha} d\psi\left(F(P_T)\right) \\
&= n_o^{\alpha} \cdot \left(-\lambda \int_0^{\theta+K} \left(-\left(\max\left(P_T-K,0\right)-\theta\right)\right)^{\alpha} d\psi\left(F(P_T)\right)\right. \\
&+ \left.\int_{\theta+K}^{\infty} \left(\max\left(P_T-K,0\right)-\theta\right)^{\alpha} d\psi\left(F(P_T)\right)\right) \\
&= n_o^{\alpha} \cdot E^{\psi}\left(1\right).
\end{aligned}
$$

□

Lemma 2. *There does not exist an optimal contract $\left(n_o^{'}, \phi^{'}\right)$ such that the participation constraint does not hold as an equality.*

Proof. The proof will proceed by contradiction. Suppose there exists an optimal contract $\left(n_o^{'}, \phi^{'}\right)$ such that

$$
E^{\psi}\left[v\left(n_o^{'}\max\left(P_T-K,0\right)+\phi^{'}e^{rT}-RP\right)\right] > E^{\psi}\left[v\left(\overline{V}e^{rT}-RP\right)\right]. \tag{6}
$$

Using the definition of the reference point in Assumption 1 and noting that the outside option $\overline{V}$ is received with certainty, we get

$$
E^{\psi}\left[v\left(n_o^{'}\left(\max\left(P_T-K,0\right)-\theta\right)\right)\right] > v\left(\overline{V}e^{rT}-n_o^{'}\theta-\phi^{'}e^{rT}\right). \tag{7}
$$

The left-hand side does not depend on the fixed wage $\phi^{'}$, while

$$
\frac{\partial}{\partial\phi}v\left(\overline{V}e^{rT}-n_o^{'}\theta-\phi^{'}e^{rT}\right) < 0.
$$

Since the RHS of (7) is continuous in ϕ, the value function has unbounded support and since there are no restrictions on ϕ, there exists $\phi^{''} < \phi^{'}$, such that

$$E^{\psi}\left[v\left(n_o^{'}\left(\max\left(P_T - K, 0\right) - \theta\right)\right)\right] = v\left(\overline{V}e^{rT} - n_o^{'}\theta - \phi^{''}\right).$$

Since the number of options is unchanged and since $\phi^{''} < \phi^{'}$, the firm pays strictly less for the contract $\left(n_o^{'}, \phi''\right)$, while still satisfying the participation constraint. Hence, $\left(n_o^{'}, \phi^{'}\right)$ cannot be optimal. □

It follows immediately from Lemma 2 that for any optimal contract (n_o^*, ϕ^*) it must be true that

$$E^{\psi}\left[v\left(n_o^{*}\left(\max\left(P_T - K, 0\right) - \theta\right)\right)\right] \cdot v\left(\overline{V}e^{rT} - n_o^{*}\theta - \phi^{*}e^{rT}\right) \geq 0. \tag{8}$$

Hence we have to consider two cases:

Case 1: $E^{\psi}\left[v\left(n_o^{*}\left(\max\left(P_T - K, 0\right) - \theta\right)\right)\right] \geq 0.$

The the certainty equivalent CE, which depends on both n_o^* and ϕ^*, is implicitly defined by

$$E^{\psi}\left[v\left(n_o^{*}\left(\max\left(P_T - K, 0\right) - \theta\right)\right)\right] \equiv \overline{E^{\psi}\left(n_o^*\right)} = \left(CE\left(n_o^*, \phi^*\right)e^{rT} - n_o^*\theta - \phi^* e^{rT}\right)^{\alpha}. \tag{9}$$

Rewriting the participation constraint using (9) gives

$$\left(CE\left(n_o^*, \phi^*\right)e^{rT} - n_o^*\theta - \phi^* e^{rT}\right)^{\alpha} = \left(\overline{V}e^{rT} - n_o^*\theta - \phi^* e^{rT}\right)^{\alpha}$$

which implies

$$CE\left(n_o^*, \phi^*\right) = \overline{V}. \tag{10}$$

From (9) we get for the certainty equivalent

$$CE\left(n_o^*, \phi^*\right) = \overline{E^{\psi}\left(n_o^*\right)}^{1/\alpha} \cdot e^{-rT} + n_o^*\theta \cdot e^{-rT} + \phi^*,$$

and since prospect value is homogenous of degree α, we have

$$\begin{aligned} CE\left(n_o^*, \phi^*\right) &= n_o^* \cdot \overline{E^{\psi}\left(1\right)}^{1/\alpha} \cdot e^{-rT} + n_o^*\theta \cdot e^{-rT} + \phi^* \\ &= n_o^* \cdot CE\left(1, 0\right) + \phi^*. \end{aligned}$$

Thus, any contract that satisfies the original participation constraint must also satisfy

$$n_o^* \cdot CE\left(1, 0\right) + \phi^* = \overline{V}. \tag{11}$$

Case 2: $E^{\psi}\left[v\left(n_o^*\left(\max\left(P_T - K, 0\right) - \theta\right)\right)\right] < 0.$

The certainty equivalent CE, which depends on both n_o^* and ϕ^*, is implicitly defined by

$$E^{\psi}\left[v\left(n_o^*\left(\max\left(P_T - K, 0\right) - \theta\right)\right)\right] \equiv \overline{E^{\psi}\left(n_o^*\right)} = -\lambda\left(-\left(CE\left(n_o^*, \phi^*\right) \cdot e^{rT} - n_o^*\theta - \phi^* \cdot e^{rT}\right)\right)^{\alpha}. \tag{12}$$

Rewriting the participation constraint using (12) and (8) gives

$$-\lambda\left(-\left(CE\left(n_o^*, \phi^*\right) \cdot e^{rT} - n_o^*\theta - \phi^* \cdot e^{rT}\right)\right)^{\alpha} = -\lambda\left(-\left(\overline{V} \cdot e^{rT} - n_o^*\theta - \phi^* \cdot e^{rT}\right)\right)^{\alpha}$$

and thus analogous to (10),

$$CE\left(n_o^*, \phi^*\right) = \overline{V}. \tag{13}$$

From (12) we get for the certainty equivalent

$$CE\left(n_o^*, \phi^*\right) = -\left(-\lambda^{-1} \cdot \overline{E^{\psi}\left(n_o^*\right)}\right)^{1/\alpha} \cdot e^{-rT} + n_o^* \theta \cdot e^{-rT} + \phi^*,$$

and since the subjective value is homogenous of degree α, we have

$$\begin{aligned} CE\left(n_o^*, \phi^*\right) &= n_o^* \cdot \left[-\left(-\lambda^{-1} \cdot \overline{E^{\psi}\left(1\right)}\right)^{1/\alpha} + \theta\right] \cdot e^{-rT} + \phi^* \\ &= n_o^* \cdot CE\left(1, 0\right) + \phi^*, \end{aligned}$$

which together with equation (13) leads to the formulation for the participation constraint as given in equation (11).

Replacing the participation constraint in (4) with (11) and solving this maximization problem in the usual way gives the optimal contract parameters (n_o^*, ϕ^*) stated in Proposition 1. ■

Chapter III

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

1 Introduction

In this chapter we explain salient features of observed compensation contracts with a simple contracting model where the manager is loss averse.[1] We parameterize this model using standard assumptions and then compare the contracts generated by the model with those actually observed for a large sample of U.S. CEOs. Our main conclusion is that a principal-agent model with loss-averse agents can approximate observed contracts far better than the standard model based on risk aversion used in the literature. In particular, the loss-aversion model can explain the prevalence of stock options, a feature that is inconsistent with the standard risk-aversion model.

The theoretical literature on executive compensation contracts is largely based on contracting models where shareholders (principal) are risk neutral and where the manager (agent) is risk averse, which is modeled with a concave utility function. Some highly

[1] This chapter is based on joint work with Ingolf Dittmann and Ernst Maug. I therefore retain the personal pronoun "we", used in the original paper, throughout this chapter. We are grateful to seminar participants at the University of Cologne, Frankfurt, Georgia State, Humboldt, Mannheim, Maryland, Tilburg, the DGF-conference in Dresden, the GEABA-conference in Tübingen, the JFI-conference on "Financial Contracting", the 6th Oxford Finance Symposium, the European Finance Association meeting in Ljubljana, and to Bo Becker, Axel Börsch-Supan, Xavier Gabaix, Gerard Hoberg, Andreas Knabe, Matjaz Koman, Roy Kouwenberg, David Larcker, Christian Laux, David De Meza, and Werner Neus for their feedback. We also thank the collaborative research centers SFB 649 on "Economic Risk" in Berlin and the SFB 504 "Rationality Concepts, Decision Making and Economic Modeling" for financial support. Ingolf Dittmann acknowledges financial support from NWO through a VIDI grant.

stylized models can explain option-type features, but quantitative approaches rely more or less entirely on a standard model with constant relative risk aversion, lognormally distributed stock prices, and effort aversion.[2] 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.[3] Dittmann and Maug find that the optimal predicted contract almost never contains any options and typically features negative base salaries. These results raise a concern for the widespread application of the model to the valuation of executive stock options and to the analysis of their design (strike price, indexing, reloading, and repricing).[4]

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

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

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

[4]Examples on design features include Hall and Murphy (2000, 2002) on the strike price, Meulbroek (2001) on the indexing of strike prices relative to benchmark variables, and Hemmer, Matsunaga, and Shevlin (1998) and Dybvig and Loewenstein (2003) on reloading.

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.[5]

The main drawback of risk-aversion approaches in explaining the prevalent use of stock options in compensation contracts is the fact that risk-averse managers gain little utility from payoffs when the value of the firm is high.[6] Whenever firm value is high, managers become wealthier and their marginal utility becomes small. This blunts any instrument for providing incentives that pays off only when firm value is high. Contracts that rely less on rewards for good outcomes ("carrots") and more on penalties for bad outcomes ("sticks") are more beneficial as they provide the same level of incentives at a lower cost. The risk-aversion model therefore predicts contracts with much higher stock holdings combined with zero or even negative salaries and option holdings. However, these predictions are at odds with observed compensation practice, where managers are paid with options, have guaranteed base salaries and entitlements to severance payments, which protect them even in case of dismissal. By comparison, loss aversion implies that managers are more averse to losses than they are attracted by gains, so they demand a premium for being exposed to losses and value the downside protection provided by options. Shareholders will therefore offer a contract that pays at least the reference wage

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

[6] This assessment relies on the standard implementation of the risk-aversion model, see Footnote 2 for the relevant literature. Other authors have pursued larger deviations from the Risk Aversion–model, which can accommodate options in a stylized setup, see for example Hemmer, Kim, and Verrecchia (1999) and the approaches by Feltham and Wu (2001) and Oyer (2004) cited above.

most of the time in order to avoid paying this premium. The loss-aversion model therefore suggests contracts that reward good outcomes rather than penalize bad outcomes and combine positive option holdings with positive fixed salaries.

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

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

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

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

Many authors apply loss aversion successfully to other questions in finance. Benartzi and Thaler (1995, 1999) develop the notion of myopic loss aversion and use it to explain the equity-premium puzzle. Gomes (2005) and Berkelaar, Kouwenberg, and Post (2004) apply the model to portfolio choice. Barberis and Huang (2001) and Barberis, Huang, and Santos (2001) apply loss aversion to the explanation of the value premium. Haigh and List (2005) find that CBOT-traders are loss averse, and more so than inexperienced students, contradicting the effect List (2004) found earlier for consumers. Coval and Shumway (2005) support the same conclusion in their study of intraday risk taking of CBOT-traders. Kouwenberg and Ziemba (2008) study the incentives and investment decisions of hedge-fund managers, and Ljungqvist and Wilhelm (2005) base their measure of issuer satisfaction in initial public offerings on loss aversion. The only application that fails to support loss aversion to the best of our knowledge is Massa and Simonov (2005) in their study of individual investor behavior. Despite the usefulness of loss aversion to analyze risk-taking incentives in many areas of finance, the only paper so far that rigorously applies loss aversion to principal-agent theory is de Meza and Webb (2007). However, they do not apply their argument to executive compensation contracts and explore a different specification from ours. To the best of our knowledge, ours is the first study that explores empirically the potential of loss aversion to explain observed compensation contracts.

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

2 The model

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

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

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

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

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

where $C(e)$ is an increasing and convex cost function. The assumption of additive separability in effort and income is conventional in the literature, and our strategy is to follow conventions in the literature for all aspects other than the modeling of preferences.[7] For this we assume preferences over wage income, $w(P_T)$, of the form[8]

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

Here, w^R denotes the reference wage. If the payoff of the contract at time T exceeds the reference wage, then the manager codes this as a gain, whereas a payoff lower than w^R is coded as a loss. We will refer to the range of the wage above w^R as the *gain space* and to the range below w^R as the *loss space*. There are three aspects that set this

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

[8] This preference specification was originally proposed by Tversky and Kahneman (1992). It has been introduced into the finance literature by Benartzi and Thaler (1995) and was used by Langer and Weber (2001), Berkelaar, Kouwenberg, and Post (2004), and Barberis and Huang (2008).

specification apart from standard concave utility specifications. First, the parameter $\lambda > 1$ gives a higher weight to payoffs below the reference wage. This reflects the observation from psychology that losses loom larger than gains of comparable size.[9] Formally, this introduces a kink in the value function at w^R and thus locally infinite risk-aversion.[10] Second, the manager treats her income from the firm separately from income from other sources, a phenomenon that is often referred to as "framing" or "mental accounting" (Thaler, 1999). Third, while $V(w(P_T))$ is concave over gains, it is convex over losses. Throughout the remainder of this paper, we will refer to a CEO with preferences of the form (2) as *loss averse* and to the corresponding principal-agent model as the loss-aversion model or, for brevity, as the LA-model. We will often compare the LA-model to the risk-aversion model (RA-model).

The standard implementation in the literature on executive compensation features preferences with constant relative risk aversion, but some papers also use constant absolute risk aversion:

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

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

where W_0 denotes wealth, γ represents the coefficient of relative risk aversion and ρ the coefficient of absolute risk aversion. Our theoretical analysis focuses on the LA-model only as the RA-model has been analyzed in many places in the literature (see Footnote 2 in the Introduction). In the empirical part we calibrate both models to the data.

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

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

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

3 Analysis

3.1 Discrete effort

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

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

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

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

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

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

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

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

The details of the proof of Proposition 1 and an implicit definition of $\widehat{P}$ are deferred to Appendix A. The proof involves three steps. The first step shows that the optimal contract can never pay off in the interior of the loss space, so $w^*(P_T)$ cannot lie strictly between $\underline{w}$ and w^R. The reason is that the agent is risk loving in the loss space, so any payment in the loss space can be improved upon by replacing it with a lottery between the lowest possible wage $\underline{w}$ and a payoff for some wage $w \geq w^R$ in the gain space. The second step shows that such lotteries are not optimal. Instead, incentives are improved if the contract always pays $\underline{w}$ if the stock price falls below some critical value $\widehat{P}$, and pays off in the gain space otherwise. The third step derives the Lagrangian and maximizes it pointwise with respect to $w(P_T)$. Equation (8) shows that for the gain space, where $P_T > \widehat{P}$, we obtain a result very similar to the familiar Holmström condition (Holmström, 1979, equation (7)) for optimal contracts in the standard concave utility model. This is intuitive, since the problem in the gain space, where preferences are concave, is not fundamentally different from a standard utility-maximizing framework.

Proposition 1 provides us with a general characterization of the optimal contract with a loss-averse manager. Figure 3.1 illustrates this contract for a typical parametrization and contrasts it with the corresponding RA-contract. For some region $P_T > \widehat{P}$ the optimal contract is continuous, monotonically increasing, and pays off only in the gain space. For $P_T \leq \widehat{P}$ the optimal contract pays off the lowest possible wage $\underline{w}$. The contract features a discontinuity at $\widehat{P}$ where the manager's wage jumps discretely from $\underline{w}$ to some value $w^*(P_T) \geq w^R > \underline{w}$.[11]

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

[11] De Meza and Webb (2007) find a similar discontinuity in a principal-agent model with loss aversion. In their specification, however, the payoff jumps from $\underline{w}$ to w^R and is flat at w^R before it possibly increases continuously. A flat payout at the reference wage w^R occurs if the slope of the line that connects $(0, \underline{w})$ and $(\widehat{P}, w^R)$ is steeper than the slope of the utility function entering the gain space. With the Kahneman and Tversky (1992) value function, 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.

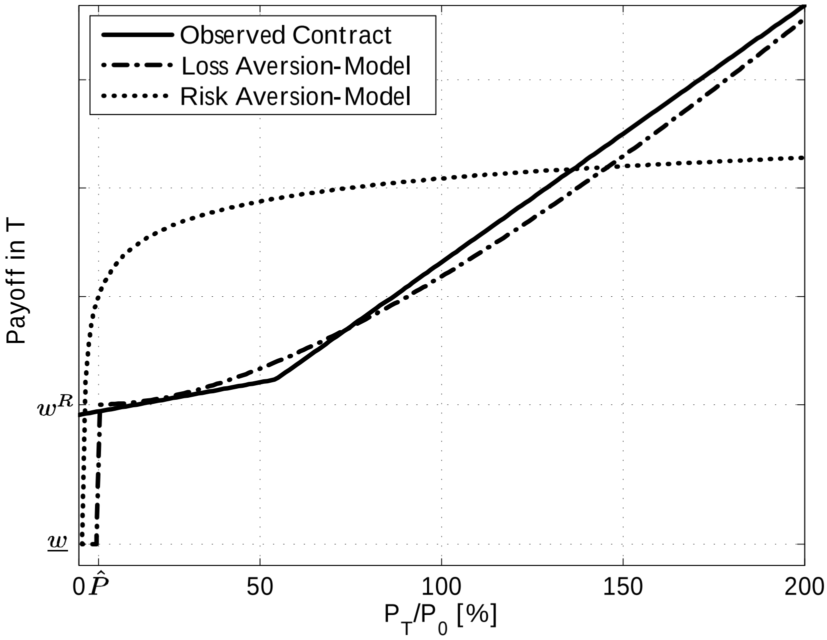

Figure 3.1: The Figure plots the loss-aversion contract (equation (8)), the risk-aversion contract, and the observed contract for a representative CEO in our sample, assuming a lognormal distribution of the terminal stock price P_T.

LA-contract (8) provides the manager with significant downside protection, punishments for extreme declines in the stock price, and increasing marginal rewards as the stock price increases. By contrast, the RA-contract provides high-powered incentives for low and intermediate stock prices and decreasing marginal rewards as the stock price increases. These qualitative features drive our empirical results for the general non-linear contracts as well as for the piecewise-linear contracts that can be implemented with stock and options.

3.2 Continuous effort

We now extend our analysis to the case where effort is continuous, so $e \in [0, \infty)$. In order to be able to solve this problem analogously to the discrete case, we have to apply the first-order approach, i.e., we replace the agent's incentive compatibility constraint (7) (more precisely, its analogue for continuous effort) with the first order condition for (7). It is always legitimate to do this if we can ensure that the manager's maximization problem when choosing her effort level is globally concave, so that the first order condition uniquely identifies the maximum of her objective function.[12] In our case, this requires that

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

This condition will not hold generally. In our setting, one issue is the convexity of the function $V\left(P_T\right)$ over the loss space. Moreover, the optimal contract $w\left(P_T\right)$ may be convex over some regions of the gain space. However, we can ensure that condition (9) holds for some cost functions C and some density functions in two ways. Firstly, equation (9) shows that this condition will be satisfied for sufficiently convex cost functions, so that $\partial^2 C\left(e\right)/\partial e^2$ is bounded from below such that (9) holds. Secondly, if the production function $P_T\left(e\right)$ is sufficiently concave (such that $\partial^2 P_T\left(e\right)/\partial e^2$ is sufficiently small for all effort levels), then (9) will also be satisfied. In the remainder of this paper we will assume that equation (9) holds. The following proposition shows that under this assumption the whole argument of the previous subsection goes through with the same implications for the optimal contract.

Proposition 2. ***(Continuous effort):*** *Assume that the agent's effort is continuous,* $e \in [0, \infty)$ *and condition (9) holds for each effort level. Then, the results from Proposition 1 continue to hold when the likelihood ratio for the discrete case,* $\Delta f\left(P_T\,|e\right)/f\left(P_T\,|\bar{e}\right)$, *is replaced by its continuous equivalent,* $f_e(P_T|e)/f(P_T|e)$.

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

4 Implementation and data

4.1 Implementation

The general loss-aversion contract. In our empirical implementation, we assume that the stock price is lognormally distributed:[13]

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

where r_f is the risk-free rate of interest, σ^2 is the variance of the returns on the stock, T the time horizon, u is a standard normal random variate and $P_0(e)$ is a strictly increasing and concave function. The expected present value of $P_T(u,e)$ under the risk-neutral density is equal to $P_0 = E[P_T \exp\{-r_f T\}]$.[14] Note that in any rational expectations equilibrium, P_0 is equal to the market value of equity at the effort level e^* chosen by the manager under the observed contract, so $P_0(e^*)$ is equal to the observed market capitalization.

We show in Appendix B that the optimal contract $w^*(P_T)$ for the problem in (5) to (7) can then be written as:

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

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

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

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

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

Hence, we can represent the non-linear LA-contract by the coefficients γ_0 and γ_1 and write it as $\mathcal{C}^{LA} = \{\gamma_0, \gamma_1\}$. This specification implies that the contract predicted by the model is strictly increasing in P_T and that it is convex as long as

$$P_T \leq \exp\{\alpha/(1-\alpha) - \gamma_0/\gamma_1\}.$$

Above this value $w^*(P_T)$ is concave. It is therefore an empirical question whether the contract described in equation (11) can explain option contracts, because the concave region may or may not be empirically relevant.

All parameters of the model given by equations (11) and (12) except γ_0 and γ_1 can be determined from standard data sources and from experimental results in the literature (see Section 4.2). We determine the remaining two parameters γ_0 and γ_1 numerically as described in the following paragraph.

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

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

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

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

This program uses a slightly more general notation as we write the wage function as $w(P_T\left|\mathcal{C}\right.)$, where $\mathcal{C}$ can refer to different types of contracts. For the time being, we only consider $\mathcal{C} = \mathcal{C}^{LA} = \{\gamma_0, \gamma_1\}$.[15] By writing P_T as in equation (10) and setting $P_0\left(e\right)$ equal to the observed value of the firm, we treat the (unknown) effort level of the CEO as given. We can then write the density without reference to the level of effort as $f(P_T)$.

Effectively, we follow Grossman and Hart (1983) and divide the solution to the optimal contracting problem into two stages, where the first stage solves for the optimal contract for a given level of effort and determines the cost of implementing this effort level. The second stage solves for the optimal contract by trading off the costs and benefits of contracts that are optimal at the first stage. We focus only on the first stage by solving program (13) to (15) as it does not depend on knowledge of the cost function $C\left(e\right)$ or of the production function $P_0\left(e\right)$. We therefore do not consider the second stage. This implies also that we cannot analyze the optimal *level* of incentives (pay for performance sensitivity) for a compensation contract, which would inevitably depend on this information. However, with our approach we can analyze the optimal *structure* of compensation contracts for any given level of incentives.

Proposition 1 provides only necessary but not sufficient conditions. We therefore

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

solve the optimization problem for different starting values in order to find the global optimum.[16] For none of the CEOs in our sample and none of the parameter constellations considered did we find any indication that there is more than one local optimum.

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

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

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

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

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

Since the observed contract is piecewise linear and expressed as a tuple of the fixed salary ϕ, the number of shares n_S, and the number of options n_O, we also calculate optimal contracts that are restricted to be piecewise linear. We will compute the piecewise linear LA-contract as the solution to program (13) to (15) and denote this contract by $\mathcal{C}_{Lin}^{LA} = \{\phi^{LA}, n_S^{LA}, n_O^{LA}\}$. Here the strike price K and the maturity T of the option grant are set equal to the strike price and maturity of the representative option that is estimated from the data. We also compare the LA-model with the RA-model and calculate optimal piecewise-linear RA-contracts, which we denote by $\mathcal{C}_{Lin}^{RA} = \{\phi^{RA}, n_S^{RA}, n_O^{RA}\}$. This is the solution to program (13) to (15) with $\mathcal{C} = \mathcal{C}_{Lin}^{RA}$, where the wage function is again piecewise linear as in (16) and preferences are given by (3) or (4).

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

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

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

Here summation is over all N CEOs in the sample. Arithmetic means over all CEOs are denoted by a bar. This metric measures the distance between the observed contract and the model contract and gives more weight to those parameters that have lower cross-sectional dispersion. D_{Lin} does not take into account fixed salaries, because these may be determined by considerations outside the model, in particular the CEO's bargaining power. In our formalization of the game shareholders have all the bargaining power, but this assumption does not affect the shape of the optimal contract.[17] If the CEO had some or all of the bargaining power then the shape of the optimal contract would still be dictated by optimal risk sharing considerations and the CEO would extract a bargaining rent through a higher base salary. For these reasons the accurate prediction of base salaries is a less important feature of the model than the prediction of the mix of stock and options. Still, we want to investigate to what extent both models predict base salaries correctly and therefore define a second metric D_{LinS} analogous to D_{Lin}, where D_{LinS} also includes the squared deviations of the base salary, $error(\phi) = \frac{\phi_i^* - \phi_i^d}{\sigma_\phi}$, where σ_ϕ is the cross-sectional standard deviation of base salaries in the sample.

A similar approach to ours was used in Carpenter (1998) and Bettis, Bizjak, and Lemmon (2005).[18] 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 (15) and the participation constraint (14) ensure that deviations from the observed value for one parameter result in deviations for the other two parameters as well. For example, an increase in the number of options increases incentives and therefore generates a lower number of shares. Hence, large deviations for one parameter result in similarly large deviations for the other parameter (or one of the other two parameters in case of D_{LinS}), so that the scaling and weighting of any single parameter relative to the other is largely inconsequential.

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

[18] 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.

4.2 Data

We identify all CEOs in the ExecuComp database who are CEO for the entire fiscal years 2004 and 2005. We also delete all CEOs who where executives in more than one company in either 2004 or 2005 and separately estimate CEOs' contracts in 2004 and in 2005. The 2004 contracts are only needed to construct the reference wage for 2005. We set P_0 equal to the market capitalization at the end of 2004 and take the dividend rate d, the stock price volatility σ, and the proportion of shares owned by the CEO n_S^d from the 2004 data, while the fixed salary ϕ^d is calculated from 2005 data.[19]

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

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

[19] 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.

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. Nevertheless, we also repeat our analysis with the minimum wage set equal to zero, an assumption that is more commonly made in the literature.

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

Table 3.1 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. The median firm value is \$2.3bn and the median moneyness K/P_0 is 0.7, so most options are clearly in the money. The median maturity is 4.4 years. The distributions of the contract parameters are highly skewed, so their means are substantially larger than their medians. We also provide the same data for 576 CEOs in 1997 in order to show that the observed parameters are broadly similar to those observed in 2005. Apart from lower firm values and lower fixed salaries we see that volatility was lower, moneyness higher and option maturities somewhat longer compared to the 2005 data set. Option holdings have almost doubled over the interval from 1997 to

Table 3.1: Description of the dataset. This table displays mean, standard deviation, and the 10%, 50% and 90% quantiles of eleven variables for our main sample of 595 CEOs from 2005 (Panel A) and a second sample of 576 CEOs from 1997 (Panel B). "Value of Contract" is the market value of the compensation package $\pi = \phi + n_S P_0 + n_O BS$, where BS is the Black-Scholes option value. All dollar amounts are given in thousands.

Variable		Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Panel A: Sample for 2005						
Stock	n_S	1.87%	5.18%	0.04%	0.31%	3.78%
Options	n_O	1.44%	1.42%	0.15%	1.03%	3.24%
Fixed Salary	Φ	$2,496	$3,107	$594	$1,675	$4,694
Value of Contract	π	$178,966	$1,887,655	$5,523	$29,837	$157,961
Non-firm Wealth	W_0	$33,285	$113,239	$2,268	$10,298	$60,858
Firm Value	P_0	$10,650,934	$30,260,334	$342,422	$2,274,781	$19,810,415
Strike Price	K	$8,243,201	$26,213,423	$242,240	$1,479,528	$13,915,001
Moneyness	K/P_0	70.06%	20.54%	40.26%	70.81%	98.94%
Maturity	T	4.58	1.30	3.39	4.44	6.01
Stock Volatility	σ	42.83%	21.42%	22.90%	36.10%	75.10%
Dividend Rate	d	1.24%	2.70%	0.00%	0.61%	3.28%
Panel B: Sample for 1997						
Stock	n_S	2.50%	6.01%	0.02%	0.28%	8.32%
Options	n_O	1.01%	1.35%	0.00%	0.56%	2.54%
Fixed Salary	Φ	$1,786	$4,454	$459	$1,141	$2,966
Value of Contract	π	$118,319	$1,046,636	$2,409	$15,528	$93,686
Non-firm Wealth	W_0	$15,270	$67,782	$1,186	$4,253	$25,807
Firm Value	P_0	$5,236,535	$11,209,383	$258,109	$1,540,377	$11,284,427
Strike Price	K	$3,777,856	$8,251,907	$192,662	$1,085,677	$8,186,544
Moneyness	K/P_0	76.27%	22.43%	47.93%	77.15%	100.00%
Maturity	T	5.58	1.86	4.10	5.22	7.34
Stock Volatility	σ	29.28%	13.11%	16.20%	26.00%	47.40%
Dividend Rate	d	1.83%	1.90%	0.00%	1.46%	4.42%

2005. We conduct our analysis for the 2005 dataset and provide the key results also for the 1997 dataset as a robustness check.

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 and variable) below the fixed salary of the previous year as a loss and we use this as a lower bound. In addition, she may also build in some part of her deferred compensation into her reference wage. Most likely, she will evaluate her securities at a substantial discount relative to their value for a well-diversified investor. This discount depends on her attitude to risk and on her framing of the wage-setting process. We therefore regard the market value of her existing contract based on the current stock price and the number of shares and options she inherited from the previous period as a (rather implausible) upper bound for the reference wage.[20] We denote the market value of her deferred compensation in 2005 based on the number of shares and options she held in 2004 by MV and write:

$$w_{2005}^{R}(\theta) = \phi_{2004} + \theta \cdot MV(n_{2004}^{S}, n_{2004}^{O}, P_{2005}) \ , \tag{18}$$

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

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

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

5 Contracts with restricted stock and options

We now describe the piecewise-linear contracts predicted by the LA-model and compare them to the contracts predicted by the standard RA-model. Minimization of program (13) to (15) is subject to two additional constraints: First, option awards can become negative (i.e. managers can be required to write options), but the manager's short position in options cannot exceed her stock holdings $n_S e^{dT}$, so $n_O > -n_S e^{dT}$.[22] This restricts the wage function to be non-decreasing. Similarly, we assume limited liability, so the base salary is limited by the manager's non-firm wealth ($\phi > -W_0$). For each CEO, we compare the observed contract with the optimal piecewise-linear contract for the LA-model and for the RA-model.

Table 3.2 Panel A summarizes the results for the LA-Model for seven different levels of the reference wage as parameterized by θ (see equation (18)). Panel B shows the results for the RA-model for seven values of the coefficient of relative risk aversion γ.[23] Panel C

[21] See Tversky and Kahneman (1992). These values have become somewhat of a standard in the literature, see for example Benartzi and Thaler (1995), Langer and Weber (2001), Berkelaar, Kouwenberg and Post (2004), Barberis and Huang (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).

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

[23] We do not consider values of γ below 0.1 in Table 3.2 as they lead to numerical problems. When the manager is risk neutral, then the optimal contract is indeterminate and the numerical problems for low values of γ reflect this indeterminacy. The literature on executive compensation has often discussed values for γ in the range between 2 and 3. Hall and Murphy (2000) use these values that seem to go back to Lambert, Larcker, and Verrecchia (1991). Lambert and Larcker (2004) more recently proposed a value as low as 0.5. A useful point of reference here is the portfolio behavior of the CEO, since very low levels of risk aversion (below 1) imply that CEOs have implausibly highly leveraged investments in the stock market. Ingersoll (2006) develops a parametrization of the RA-model that is sufficiently similar to ours

Table 3.2: This table describes the optimal piecewise-linear contract for the base model where options and salary can become negative ($n_O \geq -n_S e^{dT}$, $\phi \geq -W_0$). It shows mean and median of the three contract parameters base salary ϕ^*, stock holdings n_S^*, and option holdings n_O^* together with the mean of the errors $error(\phi) = (\phi^* - \phi^d)/\sigma_\phi$, $error(n_S) = (n_S^* - n_S^d)/\sigma_S$, and $error(n_O) = (n_O^* - n_O^d)/\sigma_O$. The table also shows mean and median of the two distance metrics D_{Lin} and D_{LinS}, and the average probability of a loss, i.e., $Prob(w^*(P_T) < w^R)$. Panel A displays the results for the loss-aversion model for seven different reference wages parameterized by θ. Panels B and C show the results for the CRRA model and, respectively, the CARA model for six levels of the CRRA risk-aversion parameter γ. For the results in Panel C, we calculate each CEO's coefficient of absolute risk aversion ρ as $\rho = \gamma/(W_0 + \pi_0)$, where π is the market value of her observed compensation package. The last row in Panel A shows the corresponding values of the observed contract.

Panel A: Loss-aversion model

θ	Obs.	Avg. Prob. of Loss	Salary			Stock			Options			D_{Lin}		D_{LinS}	
			Mean	Median	Mean Dev.	Mean	Median	Mean Dev.	Mean	Median	Mean Dev.	Mean	Median	Mean	Median
0.0	594	4.1%	-2,458	292	-1.59	2.4%	0.5%	0.10	0.8%	0.7%	-0.43	0.54	0.16	1.95	0.74
0.1	578	13.6%	3,597	1,468	0.35	1.9%	0.5%	0.02	1.4%	0.9%	-0.02	0.71	0.15	2.23	0.67
0.2	571	20.1%	2,356	1,293	-0.05	2.2%	0.6%	0.05	1.2%	0.7%	-0.14	1.44	0.40	5.30	1.56
0.4	585	31.3%	11,980	-2,891	3.03	3.4%	1.1%	0.28	-0.2%	0.1%	-1.14	2.40	0.87	14.19	3.29
0.6	586	41.1%	9,816	-6,744	2.34	4.6%	1.7%	0.53	-1.8%	-0.2%	-2.27	3.07	1.13	19.77	4.29
0.8	585	51.0%	-7,800	-8,264	-3.29	5.3%	1.8%	0.65	-2.7%	-0.3%	-2.92	3.32	1.23	18.24	4.72
1.0	582	58.3%	-31,120	-8,894	-10.73	5.6%	1.9%	0.71	-3.2%	-0.5%	-3.29	3.47	1.28	12.48	4.91
Data	595	N/A	2,496	1,675	N/A	1.9%	0.3%	N/A	1.4%	1.0%	N/A	N/A	N/A	N/A	N/A

Table 3.2 continued

Panel B: Expected-utility model with constant relative risk aversion

γ	Obs.	Salary			Stock			Options			D_{Lin}		D_{LinS}	
		Mean	Median	Mean Dev.	Mean	Median	Mean Dev.	Mean	Median	Mean Dev.	Mean	Median	Mean	Median
0.1	595	-27,754	-8,682	-9.74	5.5%	1.8%	0.70	-2.9%	-0.4%	-3.08	3.16	1.09	11.35	4.79
0.2	593	-28,448	-8,833	-9.95	5.6%	1.8%	0.72	-3.2%	-0.5%	-3.27	3.35	1.17	11.68	4.82
0.5	595	-29,035	-8,845	-10.15	5.8%	2.0%	0.75	-3.7%	-0.7%	-3.63	3.71	1.40	12.10	5.08
1.0	593	-27,864	-8,271	-9.77	5.5%	2.1%	0.70	-4.0%	-1.1%	-3.83	3.89	1.68	11.84	5.09
3.0	594	-18,277	-5,166	-6.68	4.4%	1.6%	0.49	-4.2%	-1.4%	-3.95	3.99	2.01	9.11	4.15
6.0	585	-8,668	-1,229	-3.58	3.3%	1.0%	0.26	-3.3%	-1.1%	-3.38	3.40	1.83	5.98	2.81
20.0	487	1,017	971	-0.42	2.5%	0.5%	0.05	-2.6%	-0.6%	-2.78	2.81	1.54	3.00	1.69

Panel C: Expected-utility model with constant absolute risk aversion

γ	Obs.	Salary			Stock			Options			D_{Lin}		D_{LinS}	
		Mean	Median	Mean Dev.	Mean	Median	Mean Dev.	Mean	Median	Mean Dev.	Mean	Median	Mean	Median
0.1	595	-31,632	-9,207	-10.99	5.9%	1.9%	0.78	-3.7%	-0.5%	-3.59	3.67	1.26	12.87	5.25
0.2	595	-31,390	-9,104	-10.91	5.8%	1.9%	0.75	-3.7%	-0.6%	-3.60	3.68	1.34	12.78	5.27
0.5	595	-30,392	-9,016	-10.59	5.5%	2.0%	0.70	-3.7%	-0.8%	-3.63	3.70	1.49	12.47	5.14
1.0	595	-28,458	-8,271	-9.96	5.2%	2.0%	0.63	-3.7%	-1.0%	-3.63	3.69	1.60	11.86	4.93
3.0	594	-21,073	-6,087	-7.58	4.4%	1.6%	0.49	-3.8%	-1.3%	-3.71	3.75	1.83	9.70	4.28
6.0	595	-13,772	-3,160	-5.24	3.7%	1.2%	0.35	-3.5%	-1.1%	-3.48	3.51	1.86	7.44	3.53
20.0	590	-1,530	539	-1.29	2.4%	0.7%	0.10	-2.5%	-0.7%	-2.79	2.80	1.53	3.59	1.94

shows the same results for the RA-model for constant-absolute-risk-aversion preferences (equation (4)), where the coefficient of absolute risk aversion is chosen so that relative risk aversion corresponds to the values in Panel B.[24] For each model we show the means and medians of the contract parameters predicted by the models and the scaled mean deviations of these predicted parameters from their observed counterparts (referred to as errors in equation (17)).

Both parameterizations of the RA-model predict negative base salaries and negative option holdings, so optimal RA-contracts are concave, confirming what we expected based on the theoretical analysis above (see also Figure 3.1). Both versions of the RA-model predict larger stock holdings, although the scaled deviations are smaller here because the cross-sectional standard deviations of stockholdings is 5.2% and therefore almost four times as large as the standard deviation of option holdings, which is 1.4% (see Table 3.1). Given the similarity of the two parameterizations of the RA-model, we will focus on one model from now on. The CRRA-model performs better than the CARA-model in terms of the metric D_{LinS} for all levels of risk aversion, and also better in terms of D_{Lin} for lower levels of risk aversion. We want to make sure not to bias our analysis in favor of the LA-model and therefore focus our analysis and all comparisons on the CRRA-version from now on.

The performance of the LA-model is very sensitive to the assumed reference wage. For lower values of the reference wage ($\theta = 0$ to $\theta = 0.2$) the LA-model predicts values for all contract parameters that are broadly consistent with the data. The scaled deviations

but includes investments in the stock market. Using his equation (8) and assuming a risk premium on the stock market as low as 4% and a standard deviation of the market return of 20% gives an investment in the stock market (including exposure to the stock market through holding securities in his own firm) equal to $1/\gamma$. E.g., $\gamma = 0.1$, the lowest value considered in Table 3.2, would imply that the CEO invests ten times her wealth in the stock market. We do not wish to take a restrictive stance in order not to bias our analysis in favor of the LA-model and therefore allow for levels of risk aversion as low as 0.1, even though we regard such values as highly implausible.

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

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

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

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

We can illustrate this point with the help of Figure 3.1 above. Effectively, the piecewise-linear contract attempts to approximate the general non-linear contract as well as possible. As we increase the reference wage, the discontinuity of the general non-linear contract moves to the right, i.e., $\hat{P}$ becomes larger and moves more towards the center of the distribution. This is reflected in the average probability of loss, $\Pr(w(P_T) \leq w^R)$, in Table 3.2 Panel A. The optimal non-linear contract is locally concave at $\hat{P}$: It jumps discretely and then has a very small, positive slope. If the jump at $\widehat{P}$ is in the center of the distribution, this local concavity is important and the best approximation with a piecewise-linear contract is achieved through a concave contract with negative option holdings. We further analyze the relationship between the optimal general contract and the optimal piecewise-linear contract in Section 6 below.

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

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

An important limitation of the analysis in Table 3.2 is the fact that it confounds two aspects of our problem. First, we analyze and compare different approaches to modeling attitudes to risk. Second, we also vary the overall attitude to risk as we change the reference wage, respectively, the degree of relative risk aversion. It therefore does not seem warranted to compare *all* parameterizations of the LA-model with *all* parameterizations of the RA-model. Instead, it is more sensible to compare the two models based on *comparable* parameterizations that hold the overall attitude to risk constant in a meaningful way. Then we can be sure that differences between the models do not reflect implicit differences in the overall attitude to risk. We therefore compare parameterizations that generate the same valuation of the observed contract by the same CEO. We define the certainty equivalent of model M, CE^M, from $E\left(V^M\left(w^d\left(P_T\right)\right)\right) = V(CE^M)$. We fix θ to determine the reference wage of each CEO and then define an equivalent degree of relative risk aversion γ_e from

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

We refer to the value of γ_e that satisfies (19) as the equivalent degree of relative risk aversion, because it holds the certainty equivalent constant. A straightforward implication of this step is that we also hold the risk premium paid by shareholders, $E(w^d) - C(w^d)$, constant for both models. For each CEO and for each θ we calculate the equivalent γ_e and the optimal RA-contract with $\gamma = \gamma_e$. Table 3.3 compares the two models.

Table 3.3 Panel A reports the mean and the median difference $D^{RA}_{Lin} - D^{LA}_{Lin}$ of the distance metric D_{Lin} between the two models (as of now the RA-model refers to the CRRA-preferences). The verdict based on the mean and median of D_{Lin} as well as that based on the median of D_{LinS} is clear and independent of the overall attitude to risk: The LA-model dominates the RA-model for the entire range of reference wages.

Table 3.3: Comparison of loss-aversion model with matched risk-aversion model. This table compares the optimal loss-aversion contract with the equivalent optimal (constant relative) risk-aversion contract where each CEO has constant relative risk aversion with parameter γ, which is chosen such that both models predict the same certainty equivalent for the observed contract (equation (19)). Contracts are piecewise linear, and options and salary can become negative ($n_O \geq -n_S e^{dT}, \phi \geq -W_0$). Panel A shows the average equivalent γ , mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequency of these differences being positive. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5% and 10% level. Panel B shows the frequency of optimal option holdings, the frequency of positive optimal salaries, and the frequency of both (options and salary) being positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

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

Panel B: Positive option holdings and positive base salaries

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

The distribution of D_{LinS} is skewed, so we sometimes obtain different indications for means and medians. Note that the mean of $D^{RA}_{LinS} - D^{LA}_{LinS}$ is never significantly different from zero when it is negative, so the RA-model never dominates the LA-model for any parametrization and any test. However, the RA-model fits the data better than the LA-model according to D_{LinS} for a small number of observations (3% - 23% of the sample), some of which generate extreme deviations for the LA-model. We investigate this in more detail in Table 3.10 below and show that the large deviations occur primarily for owner-CEOs who own a large fraction of their companies.

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

Table 3.3 Panel B reports how successful the two models are in explaining the two stylized facts that fixed salaries and option holdings are almost always positive for observed CEO pay contracts. The LA-model predicts positive option holdings for 91% of the sample for $\theta = 0.1$, the value that also yields the best approximation overall. Moreover, the LA-model predicts positive salaries for the majority of all CEOs when $\theta \leq 0.2$ and then it also predicts simultaneously positive option holdings and positive base salaries. By contrast, the number of cases where the RA-model predicts simultaneously positive option holdings and positive salaries is virtually zero. The model reduces options and exchanges them for more stock and lower salaries until either the restriction on salaries ($\phi \geq -W_0$) or the restriction on option holdings ($n_O \geq -n_S e^{dT}$) 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. Altogether, the LA-model can generate the qualitative characteristics of observed contracts for the majority of the CEOs in our

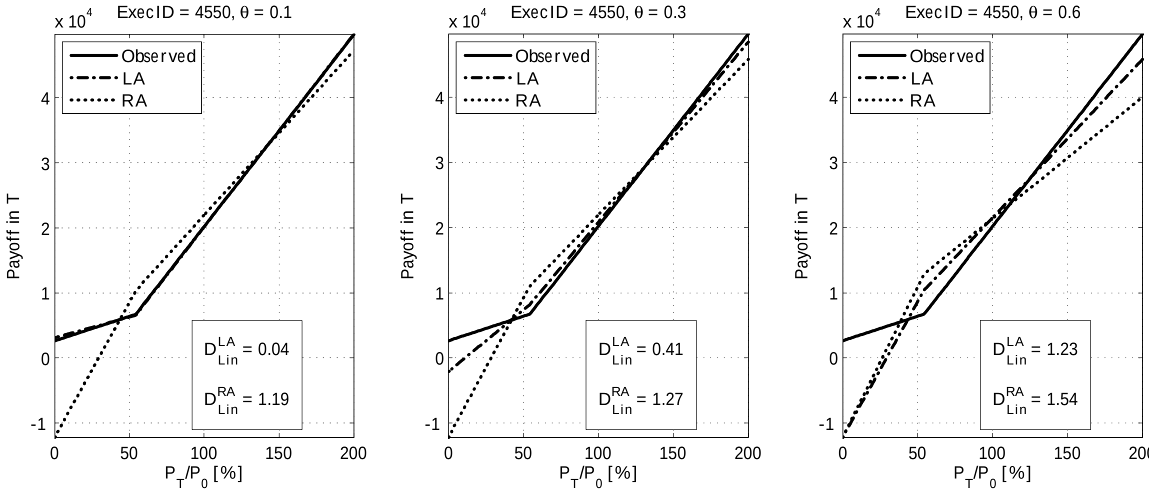

Figure 3.2: The figure shows the observed contract, the LA-contract and the RA-contract for the CEO with ExecID # 4550 for $\theta = 0.1$, $\theta = 0.3$, and $\theta = 0.6$. The horizontal axis shows the terminal stock price P_T as a percentage of the current stock price P_0. The vertical axis displays the total payoff $w(P_T)$ for each type of contract.

sample, provided we parameterize the model appropriately. The RA-model is clearly inferior on this dimension.

Figure 3.2 illustrates the results from Tables 3.2 and 3.3 for the case of a typical CEO and provides also a visual impression of the distance metric and the corresponding observed and predicted wage functions. The figure shows the optimal LA-contract, the optimal RA-contract and the observed contract for the same CEO for $\theta = 0.1$, $\theta = 0.3$, and for $\theta = 0.6$. For $\theta = 0.1$, the LA-contract and the observed contract are visually indistinguishable with a value of $D^{LA}_{Lin} = 0.04$ for the distance metric. By contrast, the corresponding RA-contract is concave and differs substantially from the observed contract, which is reflected in a higher value of the distance metric of $D^{RA}_{Lin} = 1.19$. For $\theta = 0.3$, the LA-model predicts a convex contract with positive option holdings, but with a negative base salary. Here the LA-model still performs much better than the RA-model. For $\theta = 0.6$ both models perform poorly, but the deterioration is somewhat stronger for the LA-model than it is for the RA-model, even though the LA-model still dominates.

Finally, we observe that our results on optimal contracts rely entirely on risk-sharing considerations. In particular, shareholders' objective to reduce the CEO's rents never plays a role in our analysis. Both, the theoretical contract (11) and the observed contract may provide the agent with a positive rent [25]. Any rent the agent receives in the observed contract is preserved in our calibrations, because the participation constraint in our numerical work (14) ensures that the agent's utility from the optimal contract is never lower than her utility from the observed contract. Empirically, constraint (14) is always binding in our sample, so the optimal contract provides the agent with exactly the same rent as the observed contract.

6 General non-linear loss-aversion contracts

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

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

[25] Our preference specification (2) implies that the agent's lowest possible utility is bounded away from minus infinity, so rents cannot be precluded (see Proposition 2 in Grossman and Hart, 1983). In the observed contract, rents could additionally be caused by rigid salaries (i.e. liquidity constraints) or managerial power.

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

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

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

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

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

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

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

We have no reliable method to evaluate individual dismissal probabilities for CEOs. We estimate the average probability of dismissal by calculating the frequency with which CEOs in the ExecuComp database leave the company within a given four-year period,

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

Table 3.4: This table describes the optimal non-linear loss-aversion contract. The table shows the average slope of the wage function below the observed strike price Δ_{Low}, the average slope of the wage function above the observed strike price Δ_{High}, and the frequency with which $\Delta_{High} > \Delta_{Low}$. In addition, the table shows (i) the average dismissal probability, defined as the probability with which the contract pays the minimum wage $\underline{w}$, (ii) the incentives from dismissals that are generated by the drop to the minimum wage $\underline{w}$, and (iii) the mean inflection quantile, which is the quantile at which the curvature of the optimal wage function changes from convex to concave. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

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

where the recorded reason is "resigned." we repeat this for all four-year periods between 1995 and 2004 and obtain an average dismissal probability of 7.4%. This number is inferred from a cross-section and the *ex ante* probabilities may well vary across CEOs. However, we have no reliable way of modeling this heterogeneity here, so we can only compare the mean generated by the model with the mean in the data.

Table 3.4 reports the average slopes Δ_{Low} and Δ_{High}, the dismissal probability, and the quantile of the inflection point for different parameterizations. We also report the percentage of those CEOs where $\Delta_{High} > \Delta_{Low}$. The contracts predicted by the LA-model are mostly convex by both measures of convexity. The slope in the upper range, Δ_{High} is almost always higher than the slope in the lower range, Δ_{Low}. Similarly, almost all of the probability mass for this contract lies to the left of the inflection point, rendering the concave part of the contract irrelevant.

The dismissal probabilities are unrealistically high for the LA-model once the reference point becomes sufficiently high ($\theta-$values above 0.5). This aspect underlines our earlier assessment that high parameterizations with high reference wages lead to poor performance of the LA-model. As the reference wage increases, the threat of dismissals becomes more important. Intuitively, CEOs with a higher reference wage demand a higher compensation, and they receive it in the sense that their compensation while they are employed is larger. However, then incentives are provided to a lesser extent through the slope of the wage function (note how Δ_{Low} and Δ_{High} both tend to decline as w^R increases) and to a larger extent through the threat of dismissals (column seven in Table 3.4).

In principle, the optimal non-linear contract (11) could be approximated with a sufficiently large number of options with different strike prices, where option holdings are negative for some strike prices to approximate the discrete jump and the concave part of the wage function for very high wages. In practice however, we do not observe contracts with negative option holdings. This raises the question how costly it is to restrict the contract shape to being piecewise linear, i.e. implementable by fixed salary, stock and one option grant. In Table 3.5 we therefore compare the optimal non-linear contract (11) with the optimal piecewise-linear contract. For both contracts, the table shows the average slopes Δ_{Low} and Δ_{High} and the distance metric D_{NonLin}, which parallels our definition of D_{Lin}:[27]

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

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

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

Here, $\Delta_{Low}^{d,i}$ and $\Delta_{High}^{d,i}$ represent the slopes of the observed contract corresponding to (20) and (21) and $\bar{\Delta}_{Low}^{d}$ and $\bar{\Delta}_{Low}^{d}$ denote their sample averages. In addition, Table 3.5 shows how much shareholders could save (as a proportion of total observed compensation) if they could recontract and replace the observed contract with the contract predicted by the models. These savings from recontracting are defined as

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

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

Table 3.5 shows that the accuracy (i.e. the negative of the average D_{NonLin}) of the general contract is higher than the accuracy of the piecewise-linear contract except for $\theta = 0$. For low reference wages the difference is small, but it increases as the reference wage increases. By construction, the savings relative to the status-quo of the optimal general contract are higher than the savings of the piecewise-linear contract.

The savings are not substantial for either version of the contract. This is important, because it shows that even where the distance between the observed contracts and the predicted contracts appears large in terms of the metrics developed above, the savings are insubstantial, particularly for the piecewise-linear contract. The difference in savings between the piecewise-linear contract and the general non-linear contract is small: It is 0.4% for $\theta = 0$ of total compensation costs and 4.6% for $\theta = 0.4$, or \$1.37 million for the median CEO with a pay package worth \$29.8 million. This is about 0.06% of the value of the median company. These savings have to be related to the costs of writing and enforcing such a general contract. We conclude that the benefits of incentive provision through CEO dismissals with big drops in compensation rather than through

Table 3.5: This table compares the optimal piecewise-linear loss-aversion contract with the optimal non-linear loss-aversion contract. For piecewise-linear contracts, options and salary can become negative ($n_O \geq -n_S e^{dT}$, $\phi \geq -W_0$), while the minimum wage equals minus the CEO's wealth ($\underline{w} = -W_0 e^{r_f T}$) for non-linear contracts. For both models, the table shows the average slope of the wage function below the observed strike price, $n_S e^{dT}$ and Δ_{Low} , respectively, the average slope of the wage function above the observed strike price, $n_S e^{dT} + n_O$ and Δ_{High}, respectively, and the average distance metric D_{NonLin}. In addition, the table shows the savings $[E(w^d(P_T)) - E(w^*(P_T))]/E(w^d(P_T))$ the models predict from switching from the observed contract to the optimal contract. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

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

high-powered wage functions is negligible for most companies.

7 Robustness checks

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

Table 3.6 reports the main results of Table 3.3 if we reduce the estimate of wealth by

Table 3.6: This table contains the main results from repeating our analysis shown in Table 3.3 when we decrease or increase our wealth estimates by a factor of two. For Panel A, our wealth estimate W_0 is multiplied by 0.5. For Panel B, it is multiplied by 2. Both panels show the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequencies that option holdings and salary are both positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level.

Panel A: Results if wealth is reduced by 50%

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

Panel B: Results if wealth is increased by 100%

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

50% (Panel A) and if we increase it by 100% (Panel B). The results are qualitatively very similar to those reported in Table 3.3 for the base case. The mean and median difference of $D_{Lin}^{RA} - D_{Lin}^{LA}$ is significantly positive for all levels of the reference wage. The results are more pronounced compared to those in Table 3.3 if wealth is higher and somewhat less strong but statistically still highly significant if wealth is lower. This is so because the values for the average equivalent γ are increased if wealth is higher and reduced if wealth is lower. In the CRRA-model absolute risk aversion is lower if wealth is higher, so the equivalent γ must be higher with higher wealth, and we observe already in the discussion of Table 3.2 that very low levels of risk aversion improve the performance of the RA-model. The mean and median differences of D_{LinS} exhibit the same patterns as in the base case and the median difference favors the LA-model in all cases. Also, the percentage of CEOs for whom the fixed salary as well as option holdings are positive is hardly affected by the changes in wealth considered. Overall, none of our results seems to be affected by measurement errors of CEO wealth.

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

Table 3.7 reports the results for the model with tighter restrictions and has the same structure as Table 3.3. Comparison of the restricted model with the unrestricted base case from Panel B of Tables 3.3 shows that the restrictions have a much stronger impact on the RA-model than they have on the LA-model. This is not surprising given that the LA-model already generates non-negative base salaries and option holdings in most cases, so that tightening the constraints has no impact. However, the RA-model is still not able to generate positive salaries and positive option holdings simultaneously. One

Table 3.7: This table contains the results from repeating the analysis shown in Table 3.3 with the stricter constraints that option holdings and salaries must be non-negative ($n_O \geq 0$, $\phi \geq 0$). The table compares the optimal loss-aversion contract with the equivalent optimal (constant relative) risk-aversion contract where each CEO's risk aversion parameter γ is chosen such that both models predict the same certainty equivalent for the observed contract (equation (19)). Panel A shows the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequency of these differences being positive. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level. Panel B shows the frequency of positive optimal option holdings, the frequency of positive optimal salaries, and the frequency of both (options and salary) being positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

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

Panel B: Positive option holdings and positive base salaries

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

of the two new constraints always binds: Either option holdings are equal to zero, then salary is positive, or the predicted salary is zero and option holdings are positive. The median values of D_{Lin} and D_{LinS} in Table 3.7, Panel A show that for most CEOs the LA-model still dominates the RA-model for almost all values of the reference wage w^R. In terms of the means of $D^{RA}_{Lin} - D^{LA}_{Lin}$ the accuracy of the RA-model increases and is higher than the accuracy of the LA-model on average in most cases. This shows that ruling out concave contracts and negative base salaries improves the performance of the RA-model significantly and for a minority of cases the RA-model now dominates. We conclude that the RA-model is only able to generate positive salaries or positive option holdings if we impose this as a restriction on the maximization problem, but even with these assumptions the LA-model still dominates the RA-model for the typical CEO.

Data from 1997. The data for the 2005 cross-section of CEOs on ExecuComp may be special. As a robustness check we repeat our analysis for 1997 (see Table 3.1 Panel B for descriptive statistics on these CEOs).

Table 3.8 shows the results for 1997, which are very similar to those for the 2005 sample in Table 3.3. The percentage of CEOs where $D^{RA}_{Lin} > D^{LA}_{Lin}$ and $D^{RA}_{LinS} > D^{LA}_{LinS}$ depend less on the reference wage than they do for the 2005 sample. Both models are now better at predicting positive option holdings and positive fixed salaries, but the RA-model still cannot predict both contract features simultaneously, while the results for the LA-model are better for the 1997 dataset than they are for the 2005 dataset in this respect. The 1997 data therefore lead to very similar conclusions and, if anything, strengthen the case for the LA-model.

Preference parameters. We check to what extent our results are sensitive to our assumptions on the preference parameters. We have based our discussion on the estimates

Table 3.8: This table contains the results from repeating the analysis shown in Table 3.3 for data for 1997. The table compares the optimal loss-aversion contract with the equivalent optimal (constant relative) risk-aversion contract where each CEO's risk aversion parameter γ is chosen such that both models predict the same certainty equivalent for the observed contract (equation (19)). Contracts are piecewise linear, and options and salary can become negative ($n_O \geq -n_S e^{dT}$, $\phi \geq -W_0$). Panel A shows the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequency of these differences being positive. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level. Panel B shows the frequency of positive optimal option holdings, the frequency of positive optimal salaries, and the frequency of both (options and salary) being positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems.

Panel A: Accuracy

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

Panel B: Positive option holdings and positive base salaries

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

of α, β, and λ from the experimental literature. These estimates might be inappropriate for the study of CEOs, so we check the robustness of our results with respect to different values for the preference parameters.

Table 3.9 reports the results of a comparative static analysis in terms of the preference parameters where the reference wage w^R is set to last year's fixed salary plus 10% of the risk-neutral value of last year's stock and option holdings (i.e. $\theta = 0.1$). We report only the results for the piecewise-linear model. From the metric D_{Lin} we can see that the LA-model performs better if we increase the loss-aversion parameter λ, whereas the performance of the model deteriorates for increases in the curvature of the value function, i.e., for reductions in α and β. Increases in α and β make the value function locally risk neutral, so this result is similar to the improvement with convergence to risk neutrality noted earlier. For high α-values and β-values the attitude to risk depends then only on the degree of loss aversion λ, but unlike risk aversion loss aversion is a local property of the value function in the neighborhood of the reference point. The results of Table 3.9 therefore show that it is this local property that is responsible for the better performance of the LA-model, which improves further if this aspect is emphasized (higher λ, α and β). Conversely, for a lower degree of loss aversion and stronger curvature of the value function (lower λ, α and β) the value function becomes more similar to that of the standard CRRA-model with $\gamma = 1 - \alpha$ in the gain space, where more than 95% of the probability mass lies for the base scenario in Table 3.9 (Table 3.2 Panel A). The performance of the LA-model deteriorates accordingly and becomes more similar to that of the RA-model.

Owners versus managers. As a last robustness check we try to identify those observations where the LA-model performs consistently poorly. We split the sample into a subsample with the 54 owner-executives who own 5% or more of the shares of their firm and a subsample with the remaining 541 CEOs who own less than 5% of their firm. Table

Table 3.9: This table describes the optimal piecewise-linear loss-aversion contract for different values of the parameters λ, α, and β of the value function. The reference wage w^R is set equal to last year's fixed salary plus 10% of the risk-neutral value of last year's stock and option holdings, i.e. $\theta = 0.1$ in equation (18). Panel A shows the results for the parameter λ, Panel B for α, and Panel C for β. Options and salary can become negative ($n_O \geq -n_S e^{dT}$, $\phi \geq -W_0$). The table shows mean and median of the three contract parameters base salary ϕ^*, stock holdings n_S^*, and option holdings n_O^*. In addition, it displays mean and median of the distance metric D_{Lin}. Some observations are lost because of numerical problems.

Panel A: Loss-aversion parameter λ

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

Panel B: Gain-space curvature α

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

Panel C: Loss-space curvature β

β	Obs.	Salary		Stock		Options		D_{Lin}	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.60	267	-4,585	-1329	0.0224	0.0070	-0.0016	0.0018	1.1690	0.2942
0.70	349	-6,244	-1,396	0.0341	0.0092	-0.0142	0.0015	2.3636	0.3923
0.80	388	163	940	0.0356	0.0057	-0.0053	0.0068	1.6406	0.1434
0.88	578	3,597	1,468	0.0191	0.0047	0.0141	0.0095	0.7132	0.1478
0.95	508	2,584	1,636	0.0204	0.0053	0.0153	0.0097	0.6294	0.1297

Table 3.10: This table contains the main results from repeating our analysis shown in Table 3.3 when we split our sample according to the stock ownership of the CEOs. Panel A displays the results for CEOs who own more than 5% of their firm's equity, while Panel B displays the corresponding results for the remaining CEOs in our sample. Both panels show the average equivalent γ, mean and median of the difference between the two distance metrics D_{Lin} and D_{LinS} between the RA-model and the LA-model, and the frequencies that option holdings and salaries are both positive. Results are shown for eleven different reference wages parameterized by θ. Some observations are lost because of numerical problems. ***, **, * denote significance of the T-test for zero mean and, respectively, the Wilcoxon signed rank test for zero median at the 1%, 5%, and 10% level.

Panel A: Results for owner-managers (n_S >= 5%)

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

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

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

3.10 displays the results for the two subsamples; it provides a breakdown of the results shown in Table 3.3.

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

8 Conclusion

We develop a principal-agent model with a loss-averse agent in order to explain observed executive compensation contracts. We derive the optimal contract and show that it can be characterized by an upward sloping function that is convex over the relevant region for plausible parameterizations and by a firing rule for the manager. We parameterize this model in a way that is standard in the literature and calibrate it to observed contracts.

We find that the loss-aversion model performs better on several dimensions compared to the risk-aversion model.

- Contracts predicted by the loss-aversion model are much closer to observed contracts than contracts predicted by the risk-aversion model.

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

- The loss-aversion model predicts positive option holdings in line with observed contracts for most CEOs, whereas the risk-aversion model always predicts concave contracts with negative option holdings.
- The loss-aversion model predicts positive base salaries, whereas the risk-aversion model implies that the majority of CEOs should invest some of their private wealth in their firms without receiving a base salary.

Our results are of particular importance to the substantial literature on the design and the valuation of executive stock options that relies on variants of the risk-aversion model. Our analysis suggests that for these applications the loss-aversion model is more relevant than the risk-aversion model. Our analysis also gives some guidance regarding relevant ranges of the reference wage: predicted contracts most closely resemble observed contracts for relatively low reference wages that are set close to the previous fixed salary.

Our analysis relies on stylized contracts that abstract from a number of features of observed contracts. The simplest and probably most innocuous assumption restricts the number of option grants to one. Multiple strike prices would allow for a better approximation of the piecewise-linear contract to the optimal non-linear 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 compensation items are not related to stock price performance, so they only bias our estimate of fixed compensation downward. The risk-aversion model predicts lower levels of fixed compensation compared to the loss-aversion model, so such a downward bias in estimating fixed compensation biases our results against the loss-aversion model. Finally, we ignore severance provisions, again for lack of data, but our discussion in Section 6 suggests that our analysis can potentially help to explain the widespread use of severance arrangements. If we assume that the loss-aversion model is correct, then the benefits from threatening the CEO with dismissal and an associated drop in compensation are small and probably outweighed by the costs of a governance structure that could enforce such a contract.

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

Appendix

A Proofs

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

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

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

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

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

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

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

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

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

which implies that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

From our parametric form of P_T in equation (10), we have that $\ln\left(P_T\right)$ is normally distributed with mean $\mu\left(e\right) = \ln\left(P_0\left(e\right)\right) + \left(r_f - \frac{\sigma^2}{2}\right)T$ and standard deviation $\sigma\sqrt{T}$. The density $f\left(P_T\left|e\right.\right)$ of the lognormal distribution is then:

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

and the likelihood ratio is

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

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

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

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

allows us to write:

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

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

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

Chapter IV

The Optimal Design of Stock Options: A Loss-Aversion Approach

1 Introduction

In this chapter we investigate the gap between the recommendations of academics as well as practitioners on one side and common practice in most countries on the other side with respect to the design of executive stock options.[1] Common sense concurs with economic theory that two types of options would be particularly desirable, namely premium options and indexed options. Yet, premium options and indexed options are rarely observed in practice, and most options are granted at the money. We build a simple efficient-contracting model where we analyze optimal strike prices and then calibrate the model to a sample of CEOs. We show that there is in fact little mystery here: Companies could reduce their compensation costs by less than 0.5% if they replaced at the money options with premium options, so setting strike prices optimally is a secondary consideration from the point of view of designing efficient pay packages.

The case for premium options rests on the notion that CEOs and senior executives should only be rewarded for the value they help to generate. Consider a typical option

[1] This chapter is based on joint work with Ernst Maug. I therefore retain the personal pronoun "we", which is used in the original paper, throughout this chapter. We thank Stefan Hirth and seminar participants at the University of Aarhus for helpful comments. We are grateful to the collaborative research center SFB 504 "Rationality Concepts, Decision Making and Economic Modeling" for financial support.

grant with a maturity of 10 years, where options are typically exercised after about 7 years. If the stock price appreciates by 7% p.a., then the expected price in 7 years will be 61% above the current price, so at the money options – by far the most common type of option issued in the U.S. – provide the CEO with a significant windfall profit. By contrast, for a typical firm in our sample, a premium option, struck at the expected stock price in 7 years, would cost the firm almost 30% less than an otherwise identical at the money option.[2] If we adjust for the fact that premium options provide less incentives, by increasing the number of premium options, then the firm could still save 12% by replacing at the money options with premium options.[3] Given that the typical CEO in our sample has option grants worth about \$11 million, this is a substantial component of compensation costs, which the boards of directors and their compensation committees should be concerned about. The argument for indexed options follows a similar line of reasoning and argues that CEOs should be paid only for performance and not for luck.[4] We do not investigate indexed options in this chapter and argue in the conclusion why we expect that the logic of our analysis in this chapter carries over to the case of indexed options as well.

Two views have emerged that explain the gap between observed practice and recommendations based on the argument outlined above. The *rent-extraction view* takes the absence of indexed options and premium options as evidence for the hypothesis that managers capture the pay-setting process and extract unearned rents.[5] The *efficient-*

[2] We use the following parameters: normalize the price and strike price to 100, set the dividend yield to zero, a risk-free rate of 4%, maturity of 7 years and a volatility of 40%. Then the Black-Scholes value of the option is 48.73. With a strike-price of 161 (the expected stock price after 7 years) the Black-Scholes value drops to 34.63.

[3] In this example, we define incentives in terms of pay-performance sensitivity. The value of the CEO's options change for a small change in the stock price by $n_o N(d1)$, where n_o is the number of options granted and $N(d1)$ is the option delta. The option delta is 0.79 for at the money options and 0.64 for premium options. Hence, the number of options has to be increased by a factor of 1.248 to keep pay-performance sensitivity constant.

[4] See for example, Rappaport (1999) and Bebchuk and Fried (2004).

[5] Bertrand and Mullainathan (2001), Bebchuk and Fried (2004), particularly pp. 142-146.

contracting view holds that this reading of the evidence is one-sided. In particular, it is not clear why the U.S., which is held to be the country with the best developed corporate governance practices, should deviate furthest from the best practice for setting executive pay.[6]

The debate between these two views suffers from the fact that there is no accepted model of efficient contracting that can serve as a normative benchmark and that can also accommodate the use of stock options as part of the optimal contract. The standard principal-agent model with effort aversion, lognormal stock prices, and CRRA-preferences, which was conventionally used in the literature, cannot accommodate stock options.[7] Our analysis here relies on the loss-aversion model introduced in Chapter III, which yields combinations of stock and options that correspond broadly to the proportions observed in practice. Our contribution to the literature is that we are the first to analyze optimal strike prices in an efficient-contracting model that can endogenously generate positive option holdings and positive base salaries.

We apply the model to a sample of 724 U.S. CEOs. We calibrate the model individually and compute the optimal strike price for each CEO. We present comparative static analyses for a representative CEO as well as for the whole sample, where we also vary our assumptions about the reference wage in the loss-aversion model. We find that premium options are optimal for higher values of the reference wage but not for the lower values suggested by the results in Chapter III. We find that premium options are optimal for volatile firms and for firms that rely heavily on options as a form of incentive compensation. We then compare the costs to shareholders of contracts with optimal strike prices

[6] See Aggarwal, Erel, Stulz, and Williamson (2007) for a recent cross-country study on corporate governance quality.

[7] To the best of our knowledge the earliest use of this model for the analysis of compensation contracts is Lambert, Larcker, and Verrecchia (1991). The model was used for the analysis of optimal strike prices by Hall and Murphy (2000, 2002). Dittmann and Maug (2007) show that this model cannot accommodate stock options and that it generally predicts concave contracts.

according to the model with the costs of observed contracts. The savings from setting strike prices optimally are always small, even for those parameterizations where premium options become optimal for most companies.

The savings from switching from the observed contracts to the contracts prescribed by the model also admit another interpretation, where we adopt the model as a normative benchmark and regard the savings from recontracting as a measure of the inefficiency of pay-setting. According to the rent-extraction view, these savings should be related to indicators of the quality of corporate governance. We find that the CEO's pay slice, suggested by Bebchuk, Cremers, and Peyer (2007) as a measure of CEO power, is consistently related to the savings from recontracting. There is also weak evidence that savings are higher in firms with higher agency costs, such as R&D intensive firms and firms with high Tobin's Q. The governance index of Gompers, Ishii, and Metrick (2003) has no predictive power. While there is some evidence that internal governance is statistically significant, the economic significance is small. Hence, there is no evidence from our approach to suggest that there are large inefficiencies in the structure of CEO compensation contracts.

The reason why our results are different from those in the simple numerical example above is that our model looks at all components of the compensation package. Recontracting then does not just lead to a redesign of the option component of the contract in which a given number of options with low strike prices is replaced with more options that have higher strike prices. Instead, contemporaneous with a change in the strike price, the optimal portion of incentives from options and shares is adjusted (to meet the incentive compatibility constraint), and the base salary is changed accordingly (to meet the CEO's participation constraint).

We develop our analytic approach in the next Section 2. Section 3 describes the data set used in our empirical analysis and Section 4 contains our main results. Section 5 provides some robustness checks and Section 6 concludes.

2 The analytic approach

We develop an efficient-contracting model to analyze optimal strike prices on the basis of the loss-aversion model in Chapter III.[8] The strategy is to calibrate the model and numerically derive optimal contracts for each individual CEO in a large cross-sectional data set. In solving for the optimal contract, we endogenize the number of shares, the number of options, the base salary, and the strike price of the option. We can thus explicitly analyze how a change in one contract parameter influences the optimal choice of the other parameters.

2.1 The model

The model is a version of the static hidden action principal-agent model (Holmström, 1979). At time $t = 0$ a risk-neutral firm makes a take-it-or-leave-it offer of a contract to a loss-averse and effort-averse CEO. The CEO accepts the contract if it provides her with at least the same value (net of effort costs) as her exogenous outside opportunity.[9] If the CEO accepts the contract, she can exert non-contractible and costly effort, which enhances the expected value of the firm at time $t = T$. Any uncertainty about the firm value is resolved at time T and the CEO is paid according to the contract.

The non-contractibility of CEO effort introduces a trade-off for the firm between efficient risk-sharing and providing the CEO with an incentive to exert effort by making her pay contingent on firm value.[10] Under standard technical assumptions, the optimal contract can be shown to be a monotonically increasing function of the time T firm value

[8] The exposition of the model in this paper will focus on the essential parts. For further details and some methodological choices (such as using risk-neutral valuation or the validity of the first-order approach) see the detailed discussion in Chapter III. For details on risk-neutral pricing see also Dittmann and Maug (2007) and Cai and Vijh (2005).

[9] We do not use the term "utility" here because we are working in a loss-aversion framework.

[10] This is strictly true for risk-averse CEOs. The argument also holds empirically for the loss-averse CEOs in our set-up, since all are effectively risk averse in the sense that their certainty equivalent for the observed contract is lower than the expected value of the contract.

P_T.

To make this model operational, we assume specific functional forms for the technology, admissible contracts, and CEO preferences. For the **technology**, we assume that the value of the firm at time $t = T$, denoted by P_T, is lognormally distributed and that CEO effort, denoted by e, shifts the mean of the distribution of stock prices:

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

where r_f is the risk-free rate of interest, σ is the annualized standard deviation of stock returns, u is a standard normal random variate, and $P_0(e) = e^{-r_f T} E[P_T]$ is a strictly increasing and concave function.[11] To guarantee internal consistency of our approach, we use risk-neutral pricing throughout. Hence, the stock price is expected to appreciate annually at the risk-free rate. Note that in any rational-expectations equilibrium, P_0 is equal to the market value of equity at the effort level e^* chosen by the manager under the observed contract. We assume rational expectations, so $P_0(e^*)$ is equal to the observed market capitalization of the firm.

Admissible contracts are denoted by $w(P_T)$ and specify the pay-off to the CEO at time T as a function of firm value. As is standard in the literature, we restrict ourselves to stylized linear contracts that consist of stock, options, and base salary

$$w(P_T) = \phi e^{r_f T} + n_S P_T + n_O \max(P_T - K, 0), \tag{2}$$

where ϕ denotes fixed salary (which in our formulation we assume to be paid at $t = 0$), n_S is the number of shares, expressed as a fraction of all shares outstanding, n_O is the number of stock options (where the number of shares outstanding is normalized to one),

[11] For ease of the exposition, we will submerge reference to u and e. We also do not include dividend yields here. Dividend yields will, however, be integrated into our numerical implementation.

and K is the strike price of the option. We use the superscript "d" to denote observed contract parameters ("data") and superscript "$*$" to denote optimal contract parameters chosen by our model.

Regarding **preferences**, CEOs are assumed to be loss averse, so they evaluate outcomes of risky gambles relative to a reference point (Kahneman and Tversky, 1979 and Tversky and Kahneman, 1992). Following the literature, we assume the following parametric form:

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

Here w^R is the reference point and outcomes above the reference point are coded as gains and outcomes below the reference point are coded as losses. The reference point is assumed to be exogenous in what follows.[12] The parameters α and β determine the curvature of the value function over the gain space and the loss space, respectively. CEOs are risk averse over gains and risk seeking over losses. Finally, $\lambda \geq 1$ is the coefficient of loss aversion, which governs the steepness of the value function over losses. For values of $\lambda > 1$ the aversion to losses of all sizes is higher than the attraction to equal-sized gains.

In the absence of clear guidance from the literature, we assume that the reference point w^R is based on last year's pay package. More specifically we assume

$$w_t^R\left(\theta\right)=\phi_{t-1}+\theta\cdot MV(n_{t-1}^S,n_{t-1}^O,P_t). \tag{4}$$

Hence, the reference point equals last year's base salary plus θ times the market value of the share and option portion of last year's contract evaluated at today's stock price. For $\theta = 0$, the reference point equals last year's base salary, while for $\theta = 1$, the reference

[12]For a treatment of endogenous reference points see for example de Meza and Webb (2007).

point equals the risk-neutral value of last year's contract evaluated today.

We do not incorporate probability weighting. This is primarily a technical assumption to keep the underlying theoretical model tractable. However, there is also some research from decision scientists which suggests that individuals can "learn their way out" of distorting probabilities (van de Kuilen and Wakker, 2006, van de Kuilen, 2008). For loss aversion, on the other hand, there is strong evidence that professional traders are not less, and, if anything, more loss averse than inexperienced subjects (Haigh and List, 2005, Coval and Shumway, 2005). Hence there seems to be at least some support for the assumptions used here that CEO loss aversion is a stable effect for most professionals, whereas probability weighting may not be.

2.2 Analyzing optimal contracts

It is our aim to analyze optimal strike prices in a model where they have to be jointly determined with base salaries, the number of shares, and the number of stock options. Following Dittmann and Maug (2007) and the treatment in Chapter III, we use the set-up developed in the previous section to show that the optimal structure of compensation contracts can be derived numerically for individual CEOs based on observable contracts. We proceed under the null hypothesis that observed contracts $w^d(P_T)$ are indeed optimal. Then it should not be possible to replace $w^d(P_T)$ by a contract $w^*(P_T)$ that gives the same value and incentives to the CEO and costs less to the firm. Formally, both, the firm and the CEO would agree to replace $w^d(P_T)$ with a new contract $w^*(P_T)$ that solves

$$\min_{\{\phi, n_S, n_O, K\}} \pi\left(w\left(P_T\right)\right) \equiv \phi + n_S P_0 + n_O BS_0\left(K\right) \tag{5}$$

such that

$$\int V\left[w^*\left(P_T\right)\right] f\left(P_T\right) dP_T \geq \int V\left[w^d\left(P_T\right)\right] f\left(P_T\right) dP_T \tag{6}$$

$$\int V\left[w^*\left(P_T\right)\right] \frac{\partial f\left(P_T\right)}{\partial P_0} dP_T \geq \int V\left[w^d\left(P_T\right)\right] \frac{\partial f\left(P_T\right)}{\partial P_0} dP_T \tag{7}$$

$$\phi \geq -W_0,\ n_S \geq 0,\ n_O \geq 0. \tag{8}$$

The cost of the contract to the company, $\pi\left(w\left(P_T\right)\right)$, is approximated by the value a risk-neutral investor would pay for the contract. This value is given in (5). In our model the value of stock options is given by their Black-Scholes value, which we denote by BS.

Since the CEO is not allowed to hedge the risk imposed on her by the stock and options in her contract, the value of the contract to the CEO depends on the CEO's preferences. By equation (6), the new contract has to provide at least the same expected value to the CEO as the old one.

Incentives are the sensitivity of the CEO's expected utility with respect to the observed market value P_0. Equation (7) states that the algorithm should only consider contracts where the effort incentives of the CEO are at least as high as those under the observed contract $w^d\left(P_T\right)$. For a risk-neutral CEO ($\alpha = \beta = \lambda = 1$), this definition of incentives becomes the widely-studied pay-performance sensitivity.

We further assume that stock options and shares in the contract are bounded by zero, which means that the CEO cannot write options on her company and that she cannot short her company's stock. We allow for negative base salaries, which can be interpreted as the CEO investing in her own company from her own non-firm-related wealth. A conservative lower bound on ϕ is thus her total outside wealth W_0. There are no negative base salaries in observed contracts. We argue, however, that a good model should endogenously generate positive base salaries. Imposing $\phi \geq 0$ does not change anything material as we show in Section 5.

Given our assumptions about technology, admissible contracts, and CEO preferences, and given the fact that we can observe actual CEO pay contracts, we can numerically solve program (5) to (8) for individual CEOs. The solution to the program is a tuple $(\phi^*, n_S^*, n_O^*, K^*)$, consisting of the optimal base salary, the optimal numbers of shares and options, and the optimal strike price of the option.

The optimal contracts generated by the model can then be compared to observed contracts $\left(\phi^d, n_S^d, n_O^d, K^d\right)$. We define total savings as the reduction in expected compensation costs to the firm from switching from observed to optimal contracts as:

$$\text{Total savings } \equiv \frac{\pi\left(w^d\left(P_T\right)\right) - \pi\left(w^*\left(P_T\right)\right)}{\pi\left(w^d\left(P_T\right)\right)}. \tag{9}$$

If total savings are positive, $w^d\left(P_T\right)$ has an inefficient structure and cannot be optimal. Clearly, we do not expect a contract suggested by a highly stylized model to conform to contracts observed in reality. However, we use the savings from (9) as a metric for the difference between observed contracts and optimal contracts suggested by the model. It is these savings, which our numerical procedure maximizes.

Solving program (5) to (8) is numerically demanding because it involves searching over four dimensions (ϕ, n_S, n_O, K). Our numerical routine reliably solves problems with up to three parameters.[13] We therefore solve for (ϕ^*, n_S^*, n_O^*) using a minimization routine *given* a strike price K, and then let K vary over a grid of strike prices. For sufficiently fine grids this approach is equivalent to the one-step search over four dimensions.[14] We

[13] We use a sequential quadratic programming method implemented in the Matlab routine "fmincon."

[14] The fineness of the grid is bounded by the available computing power. Solving the model for our sample of 724 CEOs takes about 8 hours for one single value of K.

define our grid relative to the actual market value of the firm according to

$$K \equiv \psi P_0. \tag{10}$$

In our benchmark specification we use steps of 0.125 for $\psi \in [0, 4]$. Hence, we solve 33 optimization problems for each CEO.

If premium options are indeed optimal, then we should expect $\psi^* > 1$ for most CEOs in our sample. If in addition, actual pay contracts are grossly inefficient, savings from (9) should be substantial. We will test these implications on our dataset.

3 Data

3.1 Observed contracts

We identify all CEOs in the ExecuComp database who are CEO at least from January 2004 to December 2005. We restrict ourselves to CEOs in order to avoid multiple observations from one firm that are likely to be correlated. We also delete all CEOs who were executives in more than one company in either 2004 or 2005. We estimate the CEOs' contracts in 2005. We also evaluate their contracts for 2004 separately in order to construct the reference wage for 2005. We set P_0 equal to the market capitalization at the end of 2004 and take the dividend yield d, the stock price volatility σ^2, and the proportion of shares owned by the CEO n_S from the 2004 data, while the fixed salary ϕ is calculated from 2005 data.[15] The numbers of shares and options, n_S and n_O, include the CEO's total holdings of stock-based compensation, and not just the most current grant of stock and options. This is important because Hall and Liebman (1998) have shown that almost all

[15] 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.

incentives for CEOs come from their holdings of stock and options and not merely from current grants.

We estimate the option portfolio held by the CEO from 2004 data using the procedure proposed by Core and Guay (2002). We then map this option portfolio into one representative option by first setting the number of options n_O equal to the sum of the options in the option portfolio. Then we determine the strike price K and the maturity T of the representative option such that n_O representative options have the same market value and the same Black-Scholes option delta as the estimated option portfolio. We take into account the fact that most CEOs exercise their stock options before maturity by multiplying the maturity of the individual options in the estimated portfolio by 0.7 before calculating the representative option (see Huddart and Lang, 1996, and Carpenter, 1998). The maturity T determines the contracting period and the risk-free rate r_f is the U.S. government bond rate from January 2005 with maturity closest to T. After deleting 4 CEOs with stock volatility exceeding 250% and 2 companies with a dividend yield greater than 20% the raw data set contains 913 CEOs.

We estimate the portion of each CEO's wealth that is not tied up in securities of his or her company from historical data for a subsample of 496 CEOs who have a history of at least five years (as executive of any firm) in the ExecuComp database. We cumulate the CEO's income from salary, bonus, and other compensation payments, add the proceeds from sales of securities, and subtract the costs from exercising options. For this subsample, the median ratio of non-firm wealth to the risk-neutral value of the CEO's pay package (including fixed salary, stock and options) is 0.34. We therefore estimate each CEO's non-firm wealth W_0 by calculating the risk-neutral value of the CEO's pay package and then set W_0 equal to 34% of this value. This procedure introduces some noise into the estimation of wealth.[16] However, we will show below that for the majority of CEOs and

[16] It is therefore not different from (or indeed likely to be more accurate than) other procedures such as,

optimal contracts, the lower bound on base salaries are not binding, which is why we are not concerned about small measurement errors in wealth.[17]

3.2 Preference parameters

To specify the preference parameters in equation (3), we use the experimental evidence in Tversky and Kahneman (1992) and set $\lambda = 2.25$, $\alpha = 0.88$, and $\beta = 0.88$. To specify the reference wage, we use results from Chapter III, in which the model used here is calibrated to the cross-section of a subset of our 913 CEOs, where it is shown that the model fits the data well for $\theta = 0.1$. We then perform robustness checks to demonstrate that our results are not sensitive to this choice of the reference wage.

3.3 Measures of agency costs and managerial discretion

In order to test the hypothesis that inefficient structures of executive compensation contracts are systematically related to agency costs or managerial discretion, we use a range of different indicators. As a first set of measures, we hypothesize that managerial power and agency costs are likely to be higher in firms where a substantial part of the value is tied up in growth options. We use Tobin's Q, which we define as market value of equity (Compustat data item 25 $\times$ data item 199) plus the book value of assets (data item 6) minus book equity (data item 60 + data item 74), as well as expenses for research and development (R&D, data item 43) as our proxies for growth options. All variables are scaled by the book value of assets.

A second hypothesis is that managers can more easily divert cash if it is abundant (Jensen, 1986). Moreover, excess cash flow could be a sign of organizational slack, which

for example, Hall and Knox (2004) who estimate wealth as the greater of six times annual compensation or \$3 million.

[17] Unlike for models with constant relative risk aversion utility functions, loss aversion does not imply a relationship between wealth and the attitude to risk from compensation, so measurement errors of wealth are less important for our model.

is also likely to be associated with contractual inefficiencies. We use cash flow shortfall as a measure of cash flow available to management and define it as common plus preferred dividends (data item 19 + data item 21) plus cash flow from investing activities (data item 311) less cash flow from operating activities (data item 308). Again, all variables are scaled by the book value of assets. We argue that this internally generated cash is controlled by insiders and not accessible to outsiders which makes it valuable as a measure of managerial discretion.[18]

A third set of measures are measures of governance problems. Bebchuk, Cremers, and Peyer (2007) provide evidence that governance problems regarding contracting about compensation are related to the *CEO pay slice*. This measure is defined as the percentage of total compensation of the top 5 managers paid to the CEO. They define this as "the relative importance of the CEO within the top executive team in terms of ability, contribution, or power."[19] Hence, we would expect contractual inefficiencies detected in our model to be positively correlated to the CEO pay slice, which we compute using the total compensation ("TDC1") reported in ExecuComp for the top 5 executives.[20] As another measure of potential governance problems, we use the corporate governance index proposed by Gompers, Ishii, and Metrick (2003), the "GIM-Index", from Andrew Metrick's website.

Lastly, we control for size. It may be easier for managers to entrench themselves in larger firms. At the same time, larger firms may be under particular scrutiny from institutional investors, analysts, and the press, and they are more likely to rely on the services of specialized pay consultants. We therefore have no strong prior about the sign of the relation between assets and contractual inefficiency. We use the log of lagged total

[18] See also Core and Guay (2001) and Bergman and Jenter (2007) for similar uses of this measure.

[19] Bebchuk, Cremers and Peyer (2007), p. 1.

[20] We disregard all firms with less than 5 reported executives and use only the top 5 highest paid officers for companies that report compensation for more than 5 officers.

book assets in our regressions to control for size.

By using the method of Core and Guay (2002), we look at total holdings of stock and options at time $t = 0$ (2005 in our implementation), which are made up of the current tranche of stock and options and several tranches received in the past. To reflect this in our agency and managerial discretion controls, we take three-year averages for all these variables. We also winsorize all variables at the 1% and the 99% level. For the GIM-Index, which is only available biannually, we take the average of the 2002 and 2004 values.

3.4 Descriptive statistics

From the raw data set with 913 companies, we drop 142 financial companies (SIC code 6000 to 6999) and 43 observations for which at least one variable (assets, Q, cash flow shortfall, or one of their components) was missing. Of the remaining 724 companies we have a value for the CEO pay slice for 607 companies, the GIM-Index for 597 companies, and R&D for 424 companies.

Table 4.1 presents descriptive statistics. In Panel A, the median CEO in the sample holds 0.31% of the shares and stock options on 1.05% of the shares. The fixed salary for 2005 (including most bonus components) is about 1.5 million dollars and the total value of the contract (including all current holdings of stock and stock options) is 27.4 million dollars. The market value of equity is about 2 billion dollars for the median firm. These variables are skewed and means are considerably larger than medians. The moneyness of the observed contract, K^d/P_0, is 0.7, which reflects the fact that stock prices tend to appreciate over our sample period. Panel B shows three-year averages of the agency and managerial discretion proxies. Total book assets of the median firm in the sample is about 1.5 billion dollars and Tobin's Q is 1.53. Cash flow shortfall is negative, which means that the median firm spends less cash on dividends and investments than its net

Table 4.1: Descriptive statistics. The table shows descriptive statistics for our dataset of 724 managers who were CEO in 2005. Stock and options are the total number of shares and options the CEO holds at the beginning of the year, normalized by firm value. Moneyness and maturity report the strike price and maturity of a hypothetical option grant with the same value and option delta as the actual option portfolio held by the CEO. Stock volatility and dividend yield are taken directly from ExecuComp. Tobin's Q is defined as market value of equity plus book assets minus book equity all over assets. Cash flow shortfall is common plus preferred dividends plus cash flow from investing activities less cash flow from operating activities all over assets. All dollar amounts are given in thousands.

	Mean	Std. Dev.	10% Quantile	Median	90% Quantile	N
Panel A: Observed contracts						
Stock	1.93%	5.12%	0.03%	0.31%	4.34%	724
Options	1.45%	1.58%	0.15%	1.05%	3.21%	724
Fixed Salary	$2,209	$2,698	$558	$1,503	$4,108	724
BS-value of options	$23,694	$43,806	$1,103	$11,036	$52,925	724
Value of Contract	$166,033	$1,751,514	$4,939	$27,352	$153,591	724
Firm Value	$9,243,453	$29,800,000	$357,345	$1,983,262	$17,300,000	724
K^d / P_0	69.22%	21.45%	38.73%	70.17%	99.11%	724
Maturity	4.59	1.23	3.39	4.48	6.04	724
Stock Volatility	45.65%	22.17%	24.70%	39.10%	78.30%	724
Dividend Yield	0.96%	1.36%	0.00%	0.40%	2.80%	724
Panel B: Agency and Managerial Discretion Proxies						
3-year avg. of Assets_{t-1}	$6,830,569	$17,300,000	$263,397	$1,490,352	$15,500,000	724
3-year avg. of Tobin's Q_{t-1}	1.86	1.06	1.04	1.53	3.18	724
3-year avg. of Cash flow shortfall_{t-1}	-1.83%	6.87%	-9.02%	-2.11%	5.46%	724
3-year avg. of CEO pay slice	39.14%	8.96%	27.84%	39.13%	50.92%	607
3-year avg. of $\text{R\&D}_{t-1}/\text{Assets}_{t-1}$	4.76%	5.65%	0.00%	2.50%	12.60%	424
Avg. over GIM-Index 2002 and 2004	9.46	2.52	6.00	9.00	13.00	597

cash flow from operations. The median CEO pay slice is 39.1% and expenses for research and development amount to 2.5% of book assets for the median firm. The median value for the GIM-Index is 9, which is the same as in the sample used by Gompers, Ishii, and Metrick (2003). Our sample is thus not biased towards either better or worse governed firms.

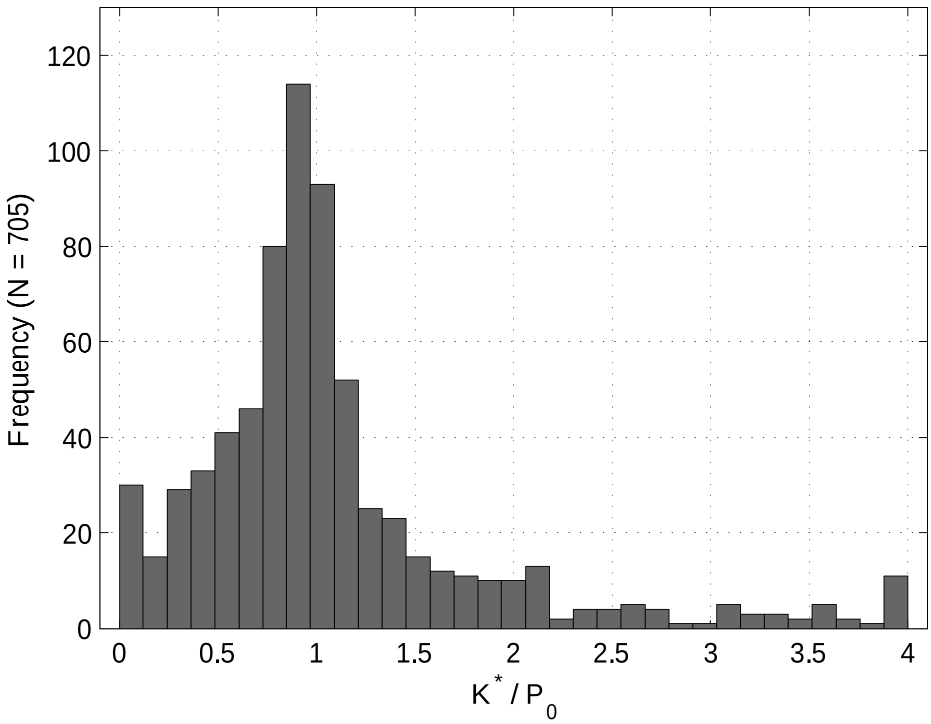

Figure 4.1: This figure shows a histogram of individually optimal strike prices K^* relative to the actual market value of the firm P_0 across all CEOs in the sample.

4 Results

4.1 Are premium options optimal?

We analyze the optimality of premium options by solving program (5) to (8) over a grid of candidate strike prices for each CEO. Optimal contracts from this procedure give the same incentives and value to the CEO as the observed contract. Effectively, we are analyzing the optimal structure of the contract the firm would offer if it could renegotiate the entire contract of the CEO including all holdings of stock-based pay granted in the past. If premium options are optimal, then the optimal strike price in the renegotiated contract should be above the current stock price.

Figure 4.1 shows a histogram of the resulting optimal strike prices, K^*, scaled by

Table 4.2: Optimal strike prices and savings. The table shows the distribution of optimal strike prices when the model is individually optimized for all CEOs in the sample. P_0 is the observed stock price. K^* is the derived optimal strike price. The table also reports total savings the firm can generate by switching from observed contract to derived optimal contract. "Savings with at the money options" is the component of total savings that could be realized by optimizing over stocks, options and fixed salaries for observed strike prices. "Savings from endogenous strike price" are incremental savings that can be realized by endogenizing the strike price. Number of observations: 705.

	Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Panel A: Optimal strike prices					
K^* / P_0	105.3%	77.5%	25.0%	87.5%	200.0%
Percent with K^* not larger than P_0	68.2%	-	-	-	-
Panel B: Potential savings					
Total savings	0.79%	1.35%	0.01%	0.17%	2.51%
Savings with at the money options	0.46%	1.12%	0.00%	0.06%	1.22%
Savings from endogenous strike price	0.34%	0.61%	0.00%	0.03%	1.25%

firm value, P_0, across our sample of CEOs.[21] Optimal strike prices cluster heavily at or slightly below the current stock price, and for a large majority of CEOs, K^* is smaller than P_0. The distribution is skewed to the right with only a minority of strike prices above the current stock price. Hence, there is no support from our model for the view that premium options are generally optimal. To the contrary, if our model is correct, Figure 4.1 suggests that at the money options are optimal for the average company.

Table 4.2 further analyzes the evidence from Figure 4.1. Panel A shows that the median strike price is at 87.5% of the current strike price and that the average is slightly higher at 105.3%. Almost 70% of firms should grant options with strike prices not higher than P_0.[22]

[21] We lose 19 CEOs (2.6 percent of our sample) because of numerical problems.

[22] The model predicts all-share contracts for a small number (about 7 percent) of CEOs and companies. Interestingly, this includes Warren Buffett at Berkshire Heathaway, who neither holds options in the predicted, nor in the observed contract.

While there is strong evidence that premium options are not optimal for most companies, they seem to be optimal for some. In general, Figure 4.1 suggests some dispersion in optimal strike prices. We therefore ask why we do not see more dispersion in strike prices across observed contracts. Panel B of Table 4.2 provides a potential answer. The savings that a firm could generate by replacing the observed contract with the optimal contract are small. On average, firms could save only 0.79%, or about $218,000 of the $27.35 million granted to the median CEO. The median firm could save as little as 0.17%. These savings can be broken down into two components. The first component are the savings the firm could realize by adjusting only the structure of the contract (base salary, stock, and options) without also adjusting the strike prices of the options (we call this component "savings from at the money options"). These savings would be on average 0.46%, which is about 60% of the total savings with endogenous strike price. The second component are the incremental savings from endogenizing the strike price. We find that on average savings of only 0.34%, or $93,000 for the median CEO, can properly be attributed to premium options. As a consequence, the costs of implementing a contract with a tailor-made strike price are very likely to outweigh the benefits in terms of more efficient contracts. Examples for such costs include direct costs for compensation consultants and indirect costs related to negotiating the magnitude of the premium.

4.2 Comparative static analysis

In our model optimal strike prices cannot be considered independently of the other contract parameters. Changing the strike price also induces a change in all other contract parameters, because the CEO has to be kept at her reservation value while maintaining incentives. In particular, tougher performance goals through higher strike prices will have to come with some form of additional compensation, because increasing the strike price

Table 4.3: Representative CEO. This table presents the characteristics of the observed contract of the representative CEO in our sample. The CEO is representative on the dimensions: firm volatility, moneyness of the options, base salary, and incentives from options as a fraction of total incentives.

Name: Hans Helmerich	*ExecuComp execid: 462*
Company: HELMERICH & PAYNE INC	*ExecuComp permid: 5581*
Stock	0.53%
Options	1.95%
Fixed Salary	\$1,376
Value of Contract	\$21,627
Firm Value	\$1,447,267
K^d / P_0	71.57%
Maturity	4.00
Stock Volatility	40.60%
Dividend Yield	1.12%

of options reduces their value to the CEO. We explore these trade-offs by analyzing the comparative statics of our model.

To demonstrate the workings of the model we will make use of a representative CEO in our sample. The representative CEO is chosen to match as closely as possible the medians of firm volatility, base salary, moneyness of the options, and incentives granted by options as a fraction of total incentives. Incentives from options ("IncOpt") are calculated on a risk-neutral basis as

$$IncOpt = \frac{n_O N(d_1)}{n_O N(d_1) + n_S e^{dT}}, \tag{11}$$

where $N(d_1)$ is the Black-Scholes delta of the option. There is exactly one CEO in our sample who is in the third quintile of each of these four variables (volatility, base salary, moneyness, and IncOpt). We provide statistics of this "representative" CEO in Table 4.3.

The representative CEO is below the median CEO in terms of the value of the contract and his base salary compared to the respective sample medians, and has somewhat higher-powered incentives.

We first conduct a comparative static analysis for the representative CEO. Using only

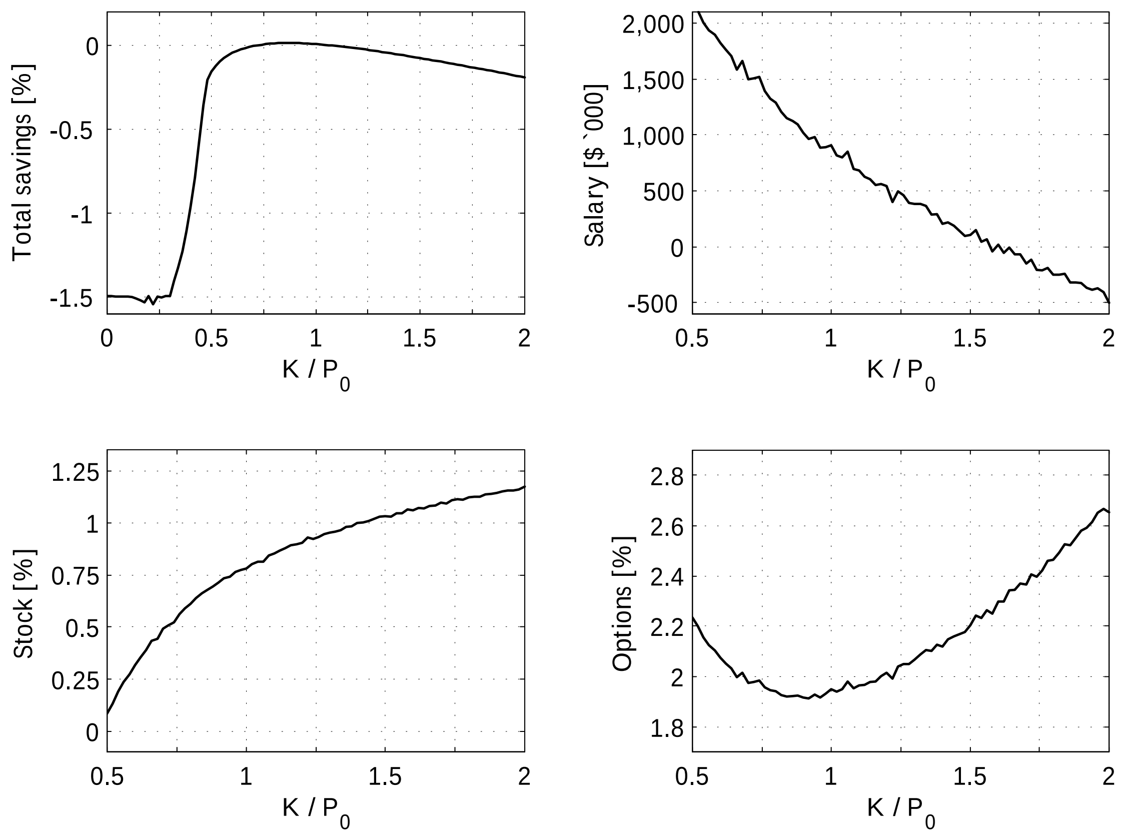

Figure 4.2: This figure shows how total savings, base salary, shares, and options predicted by the model change with a change in the strike price for the representative CEO in our sample (see Table 4.3 for parameters). The observed strike price is at $K^d/P_0 = 0.72$. The optimal strike price is at $K^*/P_0 = 0.88$.

one CEO allows us to use a finer grid for our strike prices and we choose 100 equally spaced values for K/P_0 between zero and two. Figure 4.2 shows total savings and predicted base salaries, stock, and options at each candidate strike price K/P_0. The top left plot presents the total savings at each candidate strike price. There is a unique value of $K/P_0 = 0.88$ that maximizes savings. This value is slightly above the observed strike price, which is 0.72. For $0.72 \leq K/P_0 \leq 1.08$, there exist contracts with positive savings, outside this interval no contract can make both the firm and the CEO better off, thus reinforcing our claim that our model is consistent with observed strike prices being optimal.

Base salaries (top right plot) are always declining and predicted share holdings (bottom left) are always increasing as K/P_0 increases. For values of $K/P_0 \geq 0.9$ we see that predicted stock option holdings (bottom right) are increasing. This is intuitive: increasing the strike price decreases incentives per option because it reduces the probability to see the options in the money at maturity. Hence, to keep the CEO at the incentive level of the observed contract, the number of options has to increase. Simultaneously, since stock options with higher strike prices are riskier, shares get relatively more attractive in terms of providing incentives per unit of risk. Hence, there is a substitution effect between shares and options, which is why increasing the strike price leads to both higher stock and higher option holdings. Finally, the CEO has to be kept at her reservation value, and more shares and options are granted with higher strike prices of the option, so the base salary has to be lower to satisfy the participation constraint.

For values of $K/P_0 < 0.9$ predicted stock option holdings are decreasing with the strike price. To understand this we show in Figure 4.3 the optimal non-linear contract for our representative CEO (solid line), which was derived in Chapter III. The horizontal axis depicts the stock price at maturity relative to the current stock price and the vertical axis is the total pay-off from the pay package at maturity. Above a unique cut-off value, the optimal non-linear contract is monotonically increasing and convex.[23] The optimal linear contract we derive by solving program (5) to (8) tries to approximate the non-linear contract as closely as possible over the relevant range of possible realizations of the stock price at maturity.[24] Figure 4.3 shows this for three different linear contracts, which are optimal conditional on the candidate strike prices $K/P_0 = 0.5$, $K/P_0 = 1$ and $K/P_0 = 1.5$, respectively. Increasing the strike price to stock price ratio from 1 (dashed line) to 1.5

[23]As was shown in Chapter III, the optimal non-linear contract becomes eventually concave. This concave region is not empirically relevant for the representative CEO over the range considered here. Still, for sufficiently high share prices, optimal stock option holdings might decrease again.

[24]The expected value of the stock at maturity is 109.77, the median is 78.94, and the mode is 66.94.

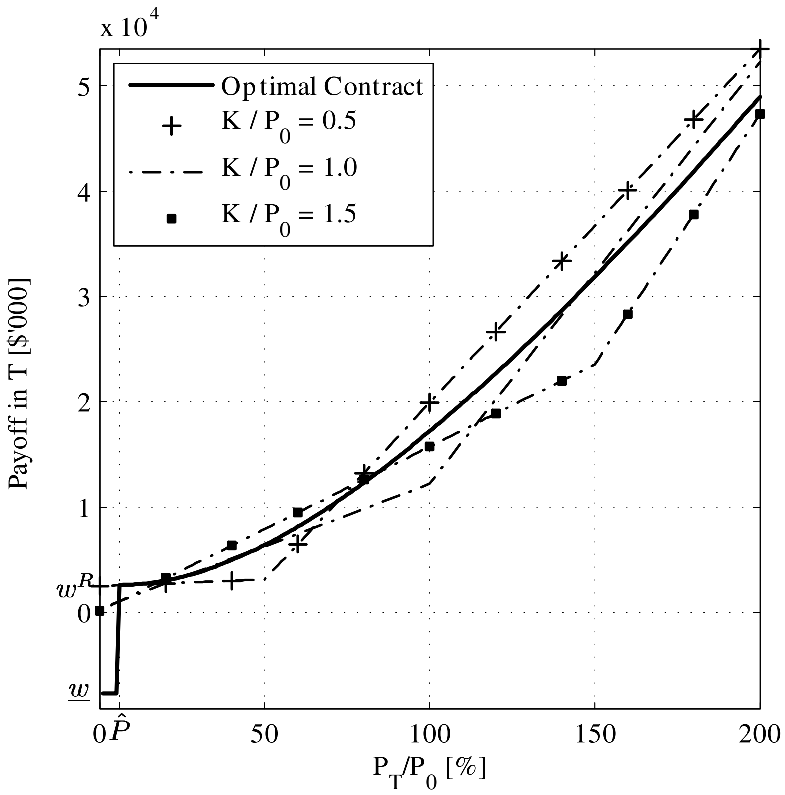

Figure 4.3: This figure shows the optimal contract for the non-linear loss-aversion model and the optimal linear contract for three different values of K/P_0 for the representative CEO in our sample (see Table 4.3 for parameters).

(dashed line with diamonds) leads to more options, more stock, and lower base salaries. Decreasing the ratio from 1 to 0.5 (dashed line with plus-sign) decreases predicted stock holdings, increases bases salary, and increases predicted options, consistent with Figure 4.2.

We have restricted the range for which we present base salaries, stock, and options in Figure 4.2 and show only values of $K/P_0 \geq 0.5$. The reason is that for very low strike prices the option delta $N(d_1)$ approaches unity and stock options become effectively like restricted stock. For the representative CEO, $N(d_1)$ is 0.92 at $K/P_0 = 0.5$ and our numerical routines cannot reliably distinguish between stock and options for lower strike prices. The algorithm can still reliably compute total savings, which approach those for an all-stock contract for lower strike prices.

We check the validity of the conclusions reached for a single representative CEO for the whole sample. Table 4.4 shows total savings and mean and median contract parameters if we uniformly set the ratio of strike price to stock price K/P_0 to the same value for the entire cross-section of CEOs. We report the percentage of firms for which our lower bounds on base salary, stock and options are binding, as well as the percentage of firms with negative predicted base salaries.

Savings are highest when strike prices are at or slightly above the current stock price, consistent with what was observed in Figure 4.1. Savings get smaller for both, higher and lower strike prices, and they are even negative, for the median firm, for options that are far in the money. Savings are negative whenever it is not possible to find an optimal contract at the given strike price that satisfies both the participation constraint and the incentive constraint and costs less to the firm than the observed contract. Effectively, by stipulating a certain strike price rather than solving for it, we impose an additional constraint on program (5) to (8) that can sometimes not be satisfied by the observed contract.

Over most of the range considered here, higher strike prices are associated with both a higher number of shares and a higher number of options. Base salaries increase with the strike price for strike prices below the current stock price and decrease for higher strike prices. For both very low and very high strike prices, the percentage of CEOs who should receive negative base salaries (invest into their own company) increases. Note that the lower bounds on the contract parameters rarely bind and that our model generates interior solutions.

Table 4.4: Cross-sectional comparative static results. This table shows means and medians for base salaries, stock, and stock options of the optimal contract for the same strike price (scaled by the stock price) across the sample. The table also reports the percentage of observations for which the lower bound on base salary ($\phi \geq -W_0$), stock ($n_S \geq 0$), and stock options ($n_O \geq 0$) is binding, as well as the percentage of firms with negative predicted base salary. This table includes 503 firms for which all 33 optimizations were successful. Of the original 705 firms, 118 firms are lost because one optimization was unsuccessful, 46 are lost because two optimizations were unsuccessful, and 38 because more than two optimizations were unsuccessful.

K / P_0	phi = $-W_0$	phi < 0	$n_S = 0$	$n_O = 0$	Total savings		Salary		Stock		Options	
					Mean	Median	Mean	Median	Mean	Median	Mean	Median
0.25	0%	44%	2%	17%	0.06%	-0.01%	-$5,301	-$1,571	2.16%	0.82%	1.41%	0.32%
0.50	0%	21%	1%	14%	0.36%	0.04%	$2,579	$405	2.23%	0.69%	1.48%	0.77%
0.75	0%	16%	0%	10%	0.61%	0.12%	$4,671	$1,462	2.36%	0.75%	1.50%	0.87%
1.00	0%	14%	0%	10%	0.69%	0.17%	$4,940	$1,443	2.48%	0.82%	1.54%	0.90%
1.25	0%	16%	0%	10%	0.72%	0.16%	$4,584	$1,318	2.57%	0.85%	1.61%	0.97%
1.50	0%	18%	0%	9%	0.72%	0.14%	$4,187	$1,178	2.68%	0.93%	1.66%	1.03%
1.75	0%	20%	0%	10%	0.71%	0.11%	$3,528	$986	2.72%	0.98%	1.82%	1.13%
2.00	0%	22%	0%	10%	0.69%	0.08%	$3,207	$871	2.77%	1.03%	1.97%	1.20%
2.25	0%	25%	0%	9%	0.67%	0.07%	$2,730	$749	2.81%	1.07%	2.15%	1.31%
2.50	0%	29%	0%	10%	0.64%	0.04%	$2,405	$630	2.85%	1.12%	2.34%	1.40%
2.75	0%	33%	0%	10%	0.62%	0.04%	$2,018	$522	2.89%	1.17%	2.54%	1.49%
3.00	0%	37%	0%	10%	0.59%	0.02%	$1,661	$455	2.92%	1.21%	2.78%	1.64%
3.25	0%	40%	0%	9%	0.56%	0.01%	$1,383	$403	2.95%	1.24%	3.01%	1.76%
3.50	0%	42%	0%	9%	0.54%	0.00%	$912	$340	2.98%	1.26%	3.26%	1.88%
3.75	0%	44%	0%	10%	0.51%	0.00%	$339	$301	3.01%	1.26%	3.53%	1.93%
4.00	0%	45%	0%	10%	0.48%	0.00%	$409	$206	3.03%	1.27%	3.87%	2.06%

4.3 Which companies should use premium options?

We now turn to the question which companies should use premium options. We conduct two types of regressions to investigate which company characteristics and which characteristics of the CEO explain the optimal moneyness of the options. First, we run a logit regression of an indicator function for premium options predicted by our model, on all observed contract characteristics. The indicator is one if $K^* > P_0$ and zero else. Columns (1) and (2) of Table 4.5 present results. Second, we use the relative premium $\left(K^* - K^d\right)/P_0$ itself as an independent variable (columns (3) and (4)).

In regression (1), premium options are positively associated with observed option holdings and negatively associated with observed stock holdings. Column (2) adds the fraction of incentives granted through options, $IncOpt$, as an additional regressor. $IncOpt$ is highly significant and positively related to premium options. Introducing $IncOpt$ also changes the sign on both stock and option holdings, which may well be due to collinearity of $IncOpt$ with observed stock holdings (Spearman's $\rho = 0.79$) and its negative correlation with option holdings ($\rho = 0.34$). Running specification (2) without shares and options as independent variables leaves the sign and significance of $IncOpt$ unchanged. Moneyness (K^d/P_0), which was insignificant before, becomes significant when controlling for incentives. In both regressions, firm volatility is significantly positive, indicating that premium options are predicted predominantly for riskier firms. Since high volatility firms are also firms with a substantial upside potential for stock options payoffs, they lose less incentives from granting stock options with high strike prices.

The distribution of $\left(K^* - K^d\right)/P_0$ is very skewed, so we use median regressions in specifications (3) and (4).[25] The results are consistent with the results from the logit model. However, the coefficient on moneyness is now negative and highly significant. In

[25] Using OLS regressions does not materially affect our results.

Table 4.5: Regression of predicted premium on observed contract parameters. The table shows the result of a logit regression of an indicator of a predicted premium option ($K^* > P_0$) and a median regression of the difference between the observed and predicted strike price scaled by the current market value of the firm on observed contract parameters. The interest rate is the U.S. government bond rate from January 2005 with maturity closest to the maturity of the representative options of each CEO. Maturity (T) is the calculated maturity of the representative option. IncOpt is the fraction of incentives that come from stock options and is calculated as in equation (11). P-values are given in parentheses. Coefficients are multiplied by 1,000. Number of observations: 705.

	Dependent variable			
	$I_{K^* > P0}$		$(K^*\text{-}K^d) / P_0$	
Independent variable	(1)	(2)	(3)	(4)
Stock	**-80.68** ***	**10.52** ***	**-4.45** ***	**-0.11**
	(0.00)	(0.01)	(0.00)	(0.68)
Options	**18.77** **	**-22.61** ***	**3.15** ***	**-0.92**
	(0.01)	(0.01)	(0.00)	(0.13)
Base salary	**0.00**	**0.00**	**0.00**	**0.00** **
	(0.43)	(0.15)	(0.80)	(0.03)
Firm volatility$_{t-1}$	**2.30** ***	**2.72** ***	**0.25** ***	**0.59** ***
	(0.00)	(0.00)	(0.00)	(0.00)
Dividend yield$_{t-1}$	**-38.61** ***	**-37.13** ***	**-4.56** ***	**-2.48** ***
	(0.00)	(0.00)	(0.00)	(0.00)
Interest rate	**-30.96**	**-189.36**	**9.56**	**3.98**
	(0.82)	(0.28)	(0.69)	(0.80)
Maturity	**0.11**	**0.43**	**-0.01**	**0.00**
	(0.64)	(0.17)	(0.76)	(0.96)
K^d / P_0	**0.71**	**1.21** **	**-0.73** ***	**-0.80** ***
	(0.12)	(0.03)	(0.00)	(0.00)
IncOpt		**8.14** ***		**1.16** ***
		(0.00)		(0.00)
Pseudo R-squared	0.19	0.37	0.15	0.32

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

specification (4), stock and stock options become insignificant when $IncOpt$ is included in the regression, consistent with $IncOpt$ capturing the information about incentives in the number of shares and options, respectively. The results from specifications (3) and (4) suggest that firms with options far in the money should increase strike prices in order

Table 4.6: Cross-sectional variation of contracts. The table shows the means, medians and interquartile ranges of important contract characteristics. It also shows the median change between observed and optimal value when the variable of interest was in the first or the fifth quintile, respectively, of its distribution over all observed contracts. IncOpt is the fraction of incentives that come from stock options and is calculated as in equation (11). The change in IncOpt, moneyness, share holdings, and options are calculated as difference between optimal and observed value. Changes in base salary are defined as percentage changes relative to observed base salaries.

Variable	Mean		Median		Interquartile range		Median change if observed value in…	
	Observed	Optimal	Observed	Optimal	Observed	Optimal	Quintile 1	Quintile 5
K / P_0	69.15%	105.32%	70.11%	87.50%	31.29%	50.00%	43.67%	-6.43%
IncOpt	56.74%	64.42%	73.00%	62.85%	41.47%	24.36%	28.49%	-22.94%
Base salary	$2,204	$4,857	$1,491	$1,324	$1,586	$2,136	25.61%	-38.39%
Stock	1.79%	1.61%	0.31%	0.56%	0.86%	1.00%	0.13%	-0.85%
Options	1.47%	1.96%	1.07%	1.17%	1.38%	1.84%	0.03%	0.22%

to re-incentivize the CEO. They also suggest that firms that grant incentives primarily through stock options should decrease this fraction by increasing strike prices.

The regressions in Table 4.5 have to be interpreted with caution because the observations for each firm are already at an internal optimum. In order to support the conclusions from the previous regressions, we also look at the cross-sectional variation of contract parameters in Table 4.6. The table shows that our model predicts a relatively homogenous mix of incentives from options and shares across executives. The median firm should decrease incentives from stock options, while on average firms should increase option incentives. Overall, the distribution of incentives from options is less dispersed for optimal contracts than for observed contracts, and on average about 60% of incentives should be granted by stock options. The last columns show that firms with the highest option incentives in observed contracts should reduce these incentives, while the firms with the lowest option incentives in observed contracts should increase them. The same tendency for extreme firms to revert back to the median is also observed for the moneyness of options. A possible interpretation is that changes in the environment have moved the parameter K/P_0 away from the optimum. Since stock prices tended to increase over the period we consider, we would expect the predicted change to be larger for firms with low moneyness than for firms with high moneyness, which is consistent with the last column in Table 4.6.

Our model thus prescribes a more homogenous mix of incentives between stock and options across CEOs. Our results are therefore more consistent with a "one size fits all"-approach to compensation, where firms use at the money strike prices and adjust stock and option holdings in line with those of the median company.

4.4 Governance implications

If the difference between observed contracts and optimal contracts from the model reflects inefficiencies in pay-setting, then we should see a relationship between these and measures of agency costs and managerial discretion.

We run median regressions of total savings on Tobin's Q, cash flow shortfall, the fraction of total top 5 compensation that goes to the CEO ("CEO pay slice"), research and development, and the corporate governance index proposed by Gompers, Ishii, and Metrick (2003). We control for firm volatility, dividend yield, and firm size. We run median regressions to address concerns about outliers in the savings variable and the independent variables. We also include industry dummies based on 30 Fama-French industries.

Table 4.7 shows that savings generated by our model are indeed systematically related to measures of agency problems and managerial power. Savings are higher in firms with higher Q, higher free cash flows, higher CEO pay slice, and higher R&D spending.

Table 4.7: Agency costs and contractual inefficiency. Median regression of total savings on stock volatility, dividend yield, and proxies for agency costs and managerial discretion. Tobin's Q, assets, cash flow shortfall, and research and development (R&D) are three-year averages. CEO pay slice is the fraction of pay to top 5 executives that goes to the CEO. GIM-Index is an average over the years 2002 and 2004. Industry dummies are based 30 Fama-French industries for all specifications. P-values are reported in parentheses. All coefficients are multiplied by 1,000.

	Dependent variable: Log (1 + savings)					
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Firm volatility$_{t-1}$	**12.32** ***	**11.97** ***	**9.60** ***	**12.62** ***	**18.09** ***	**11.84** ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Dividend yield$_{t-1}$	**-9.62**	**-7.98**	**-11.42**	**-10.77**	**-41.21**	**3.89**
	(0.55)	(0.47)	(0.33)	(0.34)	(0.22)	(0.59)
Log of Assets$_{t-1}$	**0.02**	**-0.06**	**-0.05**	**-0.10**	**0.42** **	**0.04**
	(0.85)	(0.47)	(0.61)	(0.27)	(0.03)	(0.49)
Log of Tobin's Q$_{t-1}$	**1.68** ***			**1.20** ***		
	(0.00)			(0.00)		
Cash flow shortfall$_{t-1}$		**-7.36** ***		**-10.71** ***		
		(0.00)		(0.00)		
CEO pay slice			**3.57** **	**5.10** ***		
			(0.03)	(0.00)		
R&D$_{t-1}$					**20.04** ***	
					(0.00)	
GIM-Index						**0.04**
						(0.13)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.101	0.098	0.101	0.110	0.107	0.106
N	705	705	590	590	415	582

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

All coefficients on these variables are statistically significant, although economic significance is small, as was expected given the very low levels of total savings for the sample. For example, increasing Tobin's Q by one standard deviation (1.06) increases savings by 0.25%.

While there may be other possible explanations, agency costs and managerial power are consistent with all observed effects. Managers are harder to monitor and likely to have more leeway if more of their company value is attributed to future cash flows (high Tobin's Q and high R&D). Moreover, the negative coefficient on cash flow shortfall (the difference between cash dividends and net cash flow from investment and net cash flow from operating activities all over assets) is consistent with Jensen's (1986) agency cost of cash flow hypothesis: the less cash from operations is used to invest or to pay dividends, the more is available for managers to divert and the higher is the likelihood of organizational slack.

Among the governance variables, CEO pay slice is highly significant, while the GIM-Index has the correct sign but is insignificant. This suggests a particular role for internal governance. The GIM-Index predominantly measures external governance, which is relevant for takeovers and may influence managers' tendency to extract rents from their investment policy or acquisition policy. CEO pay slice, on the other hand, is related to internal governance and measures the balance of power between the CEO and the board. The pay-setting process is likely to depend on this balance of power and the sign and significance of CEO pay slice lends additional support to the empirical validity of our model.

5 Robustness checks

We perform a number of robustness checks on our model specification as well as on our sample. We first solve program (5) to (8) but impose the tighter restriction $\phi \geq 0$. Table 4.8, which has the same structure as Table 4.2, shows the resulting distribution of optimal strike prices. The results are essentially unchanged. The mean strike price is now at 108% and the median is now at precisely 100%. Savings from recontracting are 0.77% compared to 0.79% before. Since base salaries are restricted, adjusting the level of fixed pay is not as effective anymore, and a larger fraction of savings are due to efficient setting of the

Table 4.8: Optimal strike prices and savings for the model with restricted base salary. The table shows the distribution of optimal strike prices when the model is individually optimized for all CEOs in the sample. The base salary is restricted to be positive ($\phi \geq 0$). P_0 is the observed stock price. K^* is the derived optimal strike price. The table also reports total savings the firm can generate by switching from observed contract to derived optimal contract. "Savings with at the money options" is the component of total savings that could be realized by optimizing over stock, options, and fixed salaries for observed strike prices. "Savings from endogenous strike price" are incremental savings that can be realized by endogenizing the strike price. Number of observations: 704.

	Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Panel A: Optimal strike prices					
K^* / P_0	108.0%	76.5%	25.0%	100.0%	200.0%
Percent with K^* not larger than P_0	70.7%	-	-	-	-
Panel B: Potential savings					
Total savings	0.77%	1.29%	0.01%	0.16%	2.49%
Savings with at the money options	0.15%	0.53%	0.00%	0.00%	0.37%
Savings from endogenous strike price	0.62%	1.16%	0.00%	0.09%	2.10%

strike price. We also repeat the analysis of Table 4.7 with the additional restriction on base salaries and find similar results (not tabulated).

In a second check we want to know how much our results depend on the assumed reference wage. We therefore change the parameter θ in equation (4) and investigate higher levels of the reference wage. We perform this analysis for the representative CEO first. Figure 4.4 shows that savings tend to increase with the reference wage and that premium options become more attractive for higher reference wages. For the representative CEO $K^*/P_0 = 1.64$ for $\theta = 0.3$, but total savings and the savings from endogenizing the strike price are still small. For higher levels of the reference wage savings from recontracting are higher overall, but become less dependent on the strike price. For $\theta = 0.7$, savings are in the range of 4%, but incremental savings from endogenizing the strike price essentially disappear and almost all savings can be generated by optimizing the structure of the

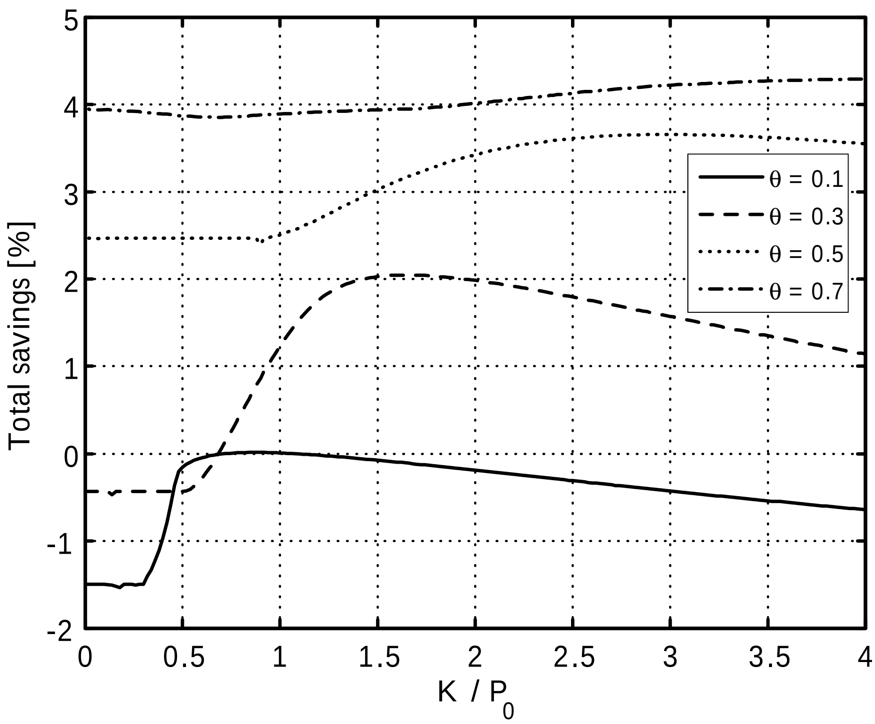

Figure 4.4: This figure shows the savings the firms could generate for a given strike price by switching from the observed to the optimal contract for the representative CEO. Four different reference point parameterizations (θ) are considered.

contracts (salary, stock, and options) alone.

In Table 4.9 Panel A, we repeat this analysis for the whole sample of CEOs. We calculate optimal contracts for a strike price equal to the current stock price (P_0), for a strike price equal to the expected stock price $E(P_T)$ for each CEO, and, as a basis for comparison, for the observed strike price. We report mean values for total savings, salary, stock, and options. The results show that total savings increase with the reference point, and that the level of savings is even smaller for the whole sample than for the representative CEO. Incremental savings from endogenizing the strike price are largest for $\theta = 0.3$ and decrease with higher reference points. The optimal levels of salary, stock, and options change with the reference point specification. For high reference points, predicted contracts do not resemble observed contracts anymore and the model predicts

Table 4.9: Comparative static results for preference parameters. This table shows means for base salaries, stock, and stock options of the optimal contract when preference parameters are varied. Results are presented for three different strike prices for the whole 2005 sample. The expected value of the stock price is calculated for each individual firm as $E[P_T] = P_0 e^{r_f T}$. Number of observations: 705.

	Total savings			Salary			Stock			Options		
	$K = K^d$	$K = P_0$	$K = E[P_T]$	$K = K^d$	$K = P_0$	$K = E[P_T]$	$K = K^d$	$K = P_0$	$K = E[P_T]$	$K = K^d$	$K = P_0$	$K = E[P_T]$
Panel A: $\alpha = \beta = 0.88$; $\lambda = 2.25$												
$\theta = 0.1$	0.43%	0.69%	0.72%	$2,202	$2,253	$1,763	1.70%	1.87%	1.94%	1.45%	1.51%	1.57%
$\theta = 0.3$	1.44%	1.88%	2.03%	-$171	$3,609	$4,183	1.89%	1.63%	1.64%	1.19%	1.82%	2.04%
$\theta = 0.5$	2.25%	2.47%	2.66%	-$4,131	-$990	$478	2.42%	2.19%	2.04%	0.50%	0.95%	1.37%
$\theta = 0.7$	2.78%	2.85%	2.97%	-$6,243	-$5,422	-$4,437	2.55%	2.53%	2.50%	0.26%	0.36%	0.54%
Panel B: $\alpha = \beta = 0.88$; $\theta = 0.1$												
$\lambda = 1.75$	0.39%	0.54%	0.56%	$1,695	$1,515	$844	1.83%	2.01%	2.11%	1.28%	1.31%	1.32%
$\lambda = 2.25$	0.43%	0.69%	0.72%	$2,202	$2,253	$1,763	1.70%	1.87%	1.94%	1.45%	1.51%	1.57%
$\lambda = 3.00$	0.52%	0.88%	0.93%	$2,353	$2,500	$2,031	1.66%	1.81%	1.88%	1.49%	1.61%	1.67%
Panel C: $\theta = 0.1$; $\lambda = 2.25$												
$\alpha = \beta = 0.80$	1.17%	1.23%	1.23%	-$1,229	-$1,882	-$2,518	2.22%	2.41%	2.46%	0.72%	0.71%	0.73%
$\alpha = \beta = 0.88$	0.43%	0.69%	0.72%	$2,202	$2,253	$1,763	1.70%	1.87%	1.94%	1.45%	1.51%	1.57%
$\alpha = \beta = 0.95$	0.22%	0.64%	0.73%	$4,124	$3,985	$3,683	1.34%	1.61%	1.68%	1.86%	1.84%	1.90%

low base salaries, more stock, and less options.[26]

Table 4.9 also shows robustness checks with respect to the other preference parameters. In Panel B, we vary the loss-aversion parameter λ. For higher degrees of loss aversion, savings increase slightly and exchanging stock for higher base salary and more stock options becomes attractive. This is the core of our model: loss-averse CEOs value downside protection. Panel C shows that increasing the degree of diminishing sensitivity, by increasing the curvature parameters α and β in the value function, diminishes the attractiveness of stock options. The payoff distribution from options is skewed to the right and a higher curvature of the value function makes these payoffs less attractive to the CEO. Overall, it seems save to conclude that our claim, that the absence of premium options is not a puzzle in terms of efficiency, is not affected by changing our assumptions about the preference parameters.

As a last robustness check we generate a dataset for the year 1997, which is otherwise identical to the dataset we used before. Table 4.10 presents descriptive statistics. Compared to 2005, base salaries and stock option holdings are lower for the 1997 sample, while stock holdings are larger. Firm value, book assets and stock price volatility are also lower in the 1997 sample compared to 2005. Among the managerial discretion proxies, cash flow shortfall is now positive: the median firm spends more on dividends and investment than its cash flow from operations (0.99% of book assets). In 2005, cash flow shortfall was negative (-2.11% of book assets).

Optimal strike prices for the 1997 sample are slightly lower than for the 2005 sample, as can be seen from Table 4.11.[27] The mean and the median firm should optimally grant stock options with a strike price at about 75% of the current 1997 stock price and premium

[26] See Chapter III for a more detailed analysis and discussion of the relationship between the reference point and the fit of the model.

[27] The results presented here are based on a slightly coarser grid with 17 equally-spaced values of ψ between 0 and 4 (steps of 0.25).

Table 4.10: Descriptive statistics for the 1997 sample. The table shows descriptive statistics for our dataset of 887 managers who were CEO in 1997. Stock and options are the total number of shares and options the CEO holds at the beginning of the year, normalized by firm value. Moneyness and maturity report the strike price and maturity of a hypothetical option grant with the same value and option delta as the actual option portfolio held by the CEO. Stock volatility and dividend yield are taken directly from ExecuComp. Tobin's Q is defined as market value of equity plus book assets minus book equity all over assets. Cash flow shortfall is common plus preferred dividends plus cash flow from investing activities less cash flow from operating activities all over assets. All dollar amounts are given in thousands.

	Mean	Std. Dev.	10% Quantile	Median	90% Quantile	N
Panel A: Observed contracts						
Stock	3.33%	6.95%	0.03%	0.43%	11.05%	887
Options	1.07%	1.34%	0.00%	0.64%	2.67%	887
Fixed Salary	$1,461	$3,619	$421	$964	$2,474	887
BS-value of options	$10,433	$26,746	$0	$3,332	$24,217	887
Value of Contract	$95,473	$847,607	$2,466	$14,253	$99,496	887
Firm Value	$4,268,301	$11,600,000	$203,112	$992,527	$8,073,254	887
K^d / P_0	76.14%	24.00%	43.35%	77.75%	100.00%	887
Maturity	5.45	1.78	4.03	5.18	7.00	887
Stock Volatility	32.03%	13.83%	17.30%	29.10%	51.60%	887
Dividend Yield	1.50%	1.81%	0.00%	0.96%	4.10%	887
Panel B: Agency and Managerial Discretion Proxies						
3-year avg. of Assets_{t-1}	$3,288,462	$6,923,044	$161,038	$850,169	$8,406,138	887
3-year avg. of Tobin's Q_{t-1}	1.90	1.19	1.02	1.54	3.20	887
3-year avg. of Cash flow shortfall_{t-1}	2.01%	8.16%	-5.80%	0.99%	10.65%	887
3-year avg. of CEO pay slice	35.96%	8.10%	26.52%	35.73%	45.28%	707
3-year avg. of $\text{R\&D}_{t-1}/\text{Assets}_{t-1}$	5.54%	8.21%	0.00%	2.60%	14.14%	470
GIM-Index 1995	9.05	2.75	5.00	9.00	13.00	591

options should be granted only for a minority of 12.6%. This reinforces our claim that premium options are not generally optimal. Savings are even smaller and the average firm could save less than 0.2% from switching to optimal contracts. Hence, the results from the 1997 sample are also consistent with optimal strike prices being a second-order issue in terms of efficiency.

In Table 4.12 we use median regressions to investigate whether our corporate governance results are robust. As before in Table 4.7, CEO pay slice and R&D are positively associated with possible savings. The GIM-Index, for which we use the value of the

Table 4.11: Optimal strike prices and savings for the 1997 sample. The table shows the distribution of optimal strike prices when the model is individually optimized for all CEOs in the 1997 sample. P_0 is the observed stock price. K^* is the derived optimal strike price. The table also reports total savings the firm can generate by switching from observed contract to derived optimal contract. "Savings with at the money options" is the component of total savings that could be realized by optimizing over stock, options, and fixed salaries for observed strike prices. "Savings from endogenous strike price" are incremental savings that can be realized by endogenizing the strike price. Number of observations: 866.

	Mean	Std. Dev.	10% Quantile	Median	90% Quantile
Panel A: Optimal strike prices					
K^* / P_0	75.8%	48.6%	0.0%	75.0%	125.0%
Percent with K^* not larger than P_0	87.4%	-	-	-	-
Panel B: Potential savings					
Total savings	0.18%	0.49%	0.00%	0.03%	0.39%
Savings with at the money options	0.09%	0.28%	0.00%	0.01%	0.21%
Savings from endogenous strike price	0.10%	0.29%	0.00%	0.00%	0.20%

GIM-Index of 1995 (because the index is not available for either 1996 or 1994), has the correct sign, but is insignificant (p-value 0.12). Tobin's Q is positively related to possible savings in specification (1), indicating that inefficiencies are more pronounced for firms with more growth options, but insignificant (p-value 0.15) in specification (4). A notable difference to Table 4.7 is that cash flow shortfall is now positive and significant, which implies that firms with less cash flow have less efficient contracts. We have seen in Table 4.10 that overall cash flow shortfall was much higher in 1997 and it may be possible that this influences our results.

Overall, Table 4.12 supports our results from the 2005 sample, which suggest that internal governance matters. Our results regarding proxies for agency problems are either weaker or inconsistent with those for the later period.

Table 4.12: Agency costs and contractual inefficiency for the 1997 sample. Median regression of total savings on stock volatility, dividend yield, and proxies for agency costs and managerial discretion. Tobin's Q, assets, cash flow shortfall, and research and development (R&D) are three-year averages. CEO pay slice is the fraction of pay to top 5 executives that goes to the CEO. GIM-Index is the value of the GIM-Index in the year 1995. Industry dummies are based on 30 Fama-French industries for all specifications. P-values are reported in parentheses. All coefficients are multiplied by 1,000.

	Dependent variable: Log (1 + savings)					
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Firm volatility$_{t-1}$	**4.79** ***	**4.50** ***	**4.07** ***	**4.22** ***	**4.21** ***	**4.44** ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Dividend yield$_{t-1}$	**3.12** *	**2.14**	**1.52**	**2.37**	**-1.10**	**5.01**
	(0.10)	(0.30)	(0.53)	(0.35)	(0.77)	(0.18)
Log of Assets$_{t-1}$	**0.13** ***	**0.11** ***	**0.10** ***	**0.13** ***	**0.10** ***	**0.12** ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log of Tobin's Q$_{t-1}$	**0.15** **			**0.12**		
	(0.01)			(0.15)		
Cash flow shortfall$_{t-1}$		**0.84** ***		**0.97** **		
		(0.01)		(0.01)		
CEO pay slice			**0.92** **	**0.74** **		
			(0.01)	(0.04)		
R&D$_{t-1}$					**3.82** ***	
					(0.00)	
GIM-Index						**0.02**
						(0.12)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.078	0.078	0.071	0.073	0.091	0.083
N	866	866	689	689	458	575

*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

6 Conclusion

This chapter calibrates a principal-agent model with a loss-averse manager to a sample of U.S. CEOs and finds that premium options are optimal for higher assumed levels of the reference wage, but not for lower values of the reference wage. The model predicts that options granted in the past that are now deep in the money should be replaced by at the money options. Overall, the case for premium options is only weak, because the savings these options would generate for shareholders through more efficient compensation contracts are small: generally, they are less than 0.5% for our benchmark model and less than 2% for most alternative specifications we consider. Hence, the size of the inefficiency from not using premium options – if there is any – is small.

We calibrate the model to each individual CEO in our sample, which gives us the opportunity to also analyze the cross-sectional variation in compensation practice. Surprisingly, this variation should be less than observed rather than more. Firms in the lowest quintile with respect to the moneyness of their options should increase it, whereas those in the top quintile should reduce it. Similar conclusions also hold for other parameters of the compensation contract, like the use of stock or the size of base salaries. The model is therefore consistent with the conclusion that "one size fits all" and that pay practices should be even more similar across companies rather than more diverse.

We also interpret the savings from recontracting as an indication of potentially inefficient corporate governance and find mixed evidence. While potential savings from switching to another contract are consistently related to measures of CEO power, and therefore suggest that internal governance matters to some extent, they are small. With the exception of R&D expenses, other measures of the quality of governance and agency problems were either not consistently related to these savings over time, or not significant at all.

This chapter only investigates premium options, but we believe that similar results hold also for other types of options, such as indexed options. The fundamental intuition behind our results is that the costs of the contract are largely determined by the outside option of the CEO and by the need to provide incentives to the CEO. Any change in the contract that reduces the value of stock options to the CEO forces an increase in another component of pay. Similarly, any change in the design of stock options that reduces the incentives they provide has to be offset by increasing the number of either options or shares.

We caution the reader that we cannot conclude from this exercise that CEOs do not extract rents in the pay-setting process. We take the observed compensation as a reflection of CEOs' outside options and can therefore not address the question whether the size of total compensation is adequate, which is the subject of another literature.[28] However we do conclude that there is little indication that the structure of observed compensation contracts reflects inefficient governance.

[28] For example, Gabaix and Landier (2008).

References

[1] Abdellaoui, Mohammed, 2000, Parameter-free elicitation of utility and probability weighting functions, *Management Science*, 46, 11, 1497-1512.

[2] Abdellaoui, Mohammed, Vossmann, Frank, and Martin Weber, 2005, Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty, *Management Science*, 51, 9, 1384-1399.

[3] Abowd, John and Michael Bognanno, 1995, International differences in executive and managerial compensation, In: Freeman, Richard and Lawrence Katz, *Differences and changes in wage structures*, University of Chicago Press, 67-103.

[4] Abowd, John and David Kaplan, 1999, Executive compensation: six questions that need answering, *Journal of Economic Perspectives*, 13, 4, 145-168.

[5] Aggarwal, Reena, Erel, Isil, Stulz, René, and Rohan Williamson, 2007, Do U.S. firms have the best corporate governance? A cross-country examination of the relation between corporate governance and shareholder wealth, Working Paper, Fisher College of Business.

[6] Allais, Maurice, 1953, Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole americaine, *Econometrica*, 21, 4, 503-546.

[7] Arellano, Manuel and Stephen Bond, 1991, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58, 2, 277-297.

[8] Ashraf, Nava, Camerer, Colin, and George Loewenstein, 2005, Adam Smith, behavioral economist, *Journal of Economic Perspectives*, 19, 3, 131-145.

[9] Baker, Malcom, Ruback, Richard, and Jeffrey Wurgler, 2006, Behavioral corporate finance, In: Eckbo, Espen (Ed.), *Handbook of corporate finance: Empirical corporate finance*, Volume A, Ch. 4., Handbooks in Finance Series, Elsevier/North-Holland.

[10] Baker, Malcom, Stein, Jeremy, and Jeffrey Wurgler, 2003, When does the market matter? Stock prices and the investment of equity-dependent firms, *Quarterly Journal of Economics*, 118, 3, 969-1006.

[11] Barberis, Nicholas and Ming Huang, 2001, Mental accounting, loss aversion, and individual stock returns, *Journal of Finance*, 56, 4, 1247-1292.

[12] Barberis, Nicholas and Ming Huang, 2004, The loss aversion / narrow framing approach to the equity premium puzzle, Working Paper, Yale School of Management.

[13] Barberis, Nicholas and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *American Economic Review*, forthcoming.

[14] Barberis, Nicholas, Huang, Ming, and Tano Santos, 2001, Prospect theory and asset prices, *Quarterly Journal of Economics*, 116, 1, 1-53.

[15] Barberis, Nicholas and Richard Thaler, 2003, A survey of behavioral finance, In: Thaler, Richard (Ed.), *Advances in behavioral finance, Volume II*, Princeton University Press.

[16] Bebchuk, Lucian, Cremers, Martijn and Urs Peyer, 2007, CEO centrality, NBER working paper 13701.

[17] Bebchuk, Lucian and Jesse Fried, 2004, *Pay without performance – The unfulfilled promise of executive compensation*, Harvard University Press.

[18] Bebchuck, Lucian and Yaniv Grinstein, 2005, The growth of executive pay, *Oxford Review of Economic Policy*, 21, 2, 283-303.

[19] Bell, David, 1982, Regret in decision making under uncertainty, *Operations Research*, 30, 5, 961-981.

[20] Benartzi, Shlomo and Richard Thaler, 1995, Myopic loss aversion and the equity premium puzzle, *Quarterly Journal of Economics*, 110, 1, 73-92.

[21] Benartzi, Shlomo and Richard Thaler, 1999, Risk aversion or myopia? Choices in repeated gambles and retirement investments, *Management Science*, 45, 3, 364-381.

[22] Bens, Daniel, Nagar, Venky, Skinner, Douglas, and Franco Wong, 2003, Employee stock options, EPS dilution and stock repurchases, *Journal of Accounting and Economics*, 36, 1-3, 51-90.

[23] Bergman, Nittai and Dirk Jenter, 2007, Employee sentiment and stock option compensation, *Journal of Financial Economics*, 84, 3, 667-712.

[24] Berkelaar, Arjan, Kouwenberg, Roy, and Thierry Post, 2004, Optimal portfolio choice under loss aversion, *Review of Economics and Statistics*, 86, 4, 973-987.

[25] Berle, Adolf and Gardiner Means, 1932, *The modern corporation and private property*, New York, Macmillan.

[26] Bertrand, Marianne and Sendhil Mullainathan, 2001, Are CEOs rewarded for luck? The ones without principal are, *Quarterly Journal of Economics*, 116, 3, 901-932.

[27] Bettis, Carr, Bizjak, John, and Michael Lemmon, 2001, Managerial ownership, incentive contracting, and the use of zero-cost collars and equity swaps by corporate insiders, *Journal of Financial and Quantitative Analysis*, 36, 3, 345-370.

[28] Bettis, Carr, Bizjak, John, and Michael Lemmon, 2005, Exercise behavior, valuation, and the incentive effects of employee stock options, *Journal of Financial Economics*, 76, 445-470.

[29] Bewley, Truman, 1999, *Why wages don't fall during a recession*, Harvard University Press.

[30] Bleichrodt, Han and Jose Luis Pinto, 2000, A parameter-free elicitation of the probability weighting function in medical decision analysis, *Management Science*, 46, 11, 1485-1496.

[31] Bowman, David, Minehart, Deborah, and Matthew Rabin, 1997, Loss aversion in a consumption-savings model, *Journal of Economic Behavior & Organization*, 38, 2, 155-178.

[32] Brickley, James, 2003, Empirical research on CEO turnover and firm-performance: A discussion, *Journal of Accounting and Economics*, 36, 1-3, 227-233.

[33] Cai, Jie and Anand Vijh, 2005, Executive stock and option valuation in a two state-variable framework: Allowing optimal investment of outside wealth in the riskfree asset and the market portfolio, *Journal of Derivatives*, 12, 9-27.

[34] Camerer, Colin, 2000, Prospect theory in the wild: Evidence from the field, In: Kahneman, Daniel and Amos Tversky (Eds.), *Choices, Values, and Frames*, Cambridge University Press.

[35] Camerer, Colin, Babcock, Linda, Loewenstein, George, and Richard Thaler, 1997, Labor supply of New York City cabdrivers: One day at a time, *Quarterly Journal of Economics*, 112, 2, 407-441.

[36] Camerer, Colin, Loewenstein, George, and Matthew Rabin, 2003, *Advances in behavioral economics*, Princeton University Press.

[37] Carpenter, Jennifer, 1998, The exercise and valuation of executive stock options, *Journal of Financial Economics*, 48, 2, 127-158.

[38] Carroll, Gabriel, Choi, James, Laibson, David, Madrian, Brigitte, and Andrew Metrick, 2005, Optimal defaults and active decisions, NBER Working Paper 11074.

[39] Conyon, Martin and Joachim Schwalbach, 2000, European differences in executive pay and corporate governance, *Zeitschrift für Betriebswirtschaft*, Ergänzungsheft 1/2000 Corporate Governance, 97-114.

[40] Core, John and Wayne Guay, 2001, Stock option plans for non-executive employees, *Journal of Financial Economics*, 61, 2, 253-287.

[41] Core, John and Wayne Guay, 2002, Estimating the value of stock option portfolios and their sensitivities to price and volatility, *Journal of Accounting Research*, 40, 3, 613-640.

[42] Core, John, Guay, Wayne, and Robert Verrecchia, 2003, Price versus non-price performance measures in optimal CEO compensation contracts, *Accounting Review*, 78, 4, 957-981.

[43] Cook, Philip and Charles Clotfelter, 1993, The peculiar scale economies of lotto, *American Economic Review*, 83, 3, 634-643.

[44] Coval, Joshua and Tyler Shumway, 2005, Do behavioral biases affect prices?, *Journal of Finance*, 60, 1, 1-34.

[45] De Meza, David and David Webb, 2007, Incentive design under loss aversion, *Journal of the European Economic Association*, 5, 1, 66-92.

[46] DellaVigna, Stefano, 2007, Psychology and economics: evidence from the field, NBER Working Paper W13420.

[47] DellaVigna, Stefano and Ulrike Malmendier, 2006, Paying not to go to the gym, *American Economic Review*, 96, 3, 694-719.

[48] Dekel, Eddie, 1986, An axiomatic characterization of preferences under uncertainty: weakening the independence axiom, *Journal of Economic Theory*, 40, 2, 304-318.

[49] Desai, Mehir, 2003, The divergence between book and tax income, In: Poterba, James (Ed.), *Tax Policy and the Economy*, 17, MIT Press.

[50] Devers, Cynthia, Wiseman, Robert, and Michael Holmes, 2007, The effects of endowment and loss aversion in managerial stock option valuation, *Academy of Management Journal*, 50, 1, 191-208.

[51] Dittmann, Ingolf and Ernst Maug, 2007, Lower salaries and no options? On the optimal structure of executive pay, *Journal of Finance*, 62, 1, 303-343.

[52] Dunn, Lucia, 1996, Loss aversion and adaptation in the labor market: Empirical indifference functions and labor supply, *Review of Economics and Statistics*, 78, 3, 441-450.

[53] Dybvig, Philip and Mark Loewenstein, 2003, Employee reload options: Pricing, hedging, and optimal exercise, *Review of Financial Studies*, 16, 1, 145-171.

[54] Edmans, Alex, Gabaix, Xavier, and Augustin Landier, 2007, A multiplicative model of optimal CEO incentives in market equilibrium, Working paper, New York University.

[55] Engel, Ellen, Hayes, Rachel, and Xue Wang, 2003, CEO turnover and properties of accounting information, *Journal of Accounting and Economics*, 36, 1-3, 197-226.

[56] Farrell, Kathleen and David Whidbee, 2003, Impact of firm performance expectations on CEO turnover and replacement decisions, *Journal of Accounting and Economics*, 36, 1-3, 165-196.

[57] Feltham, Gerald and Martin Wu, 2001, Incentive efficiency of stock versus options, *Review of Accounting Studies*, 6, 1, 7-28.

[58] Fishburn, Peter, 1984, SSB utility theory: an economic perspective, *Mathematical Social Sciences*, 8, 1, 63-94.

[59] Fox, Craig and Kelly See, 2003, Belief and preference in decision under uncertainty, In: Hardman, David and Laura Macchi, *Thinking: psychological perspectives on reasoning, judgment and decision making*, John Wiley & Sons Ltd, 273-314.

[60] Gabaix, Xavier and David Laibson, 2006, Shrouded attributes, consumer myopia, and information suppression in competitive markets, *Quarterly Journal of Economics*, 121, 2, 505-540.

[61] Gabaix, Xavier and Augustin Landier, 2008, Why has CEO pay increased so much?, *Quarterly Journal of Economics*, 123, 1, 49-100.

[62] Gomes, Francisco, 2005, Portfolio choice and trading volume with loss averse investors, *Journal of Business*, 78, 2, 675-706.

[63] Gompers, Paul, Ishii, Joy, and Andrew Metrick, 2003, Corporate governance and equity prices, *Quarterly Journal of Economics*, 118, 1, 107-155.

[64] Gonzales, Richard and George Wu, 1999, On the form of the probability weighting function, *Cognitive Psychology*, 38, 129-166.

[65] Grossman, Sanford and Oliver Hart, 1983, An analysis of the principal-agent problem, *Econometrica*, 51, 1, 7-45.

[66] Gul, Faruk, 1991, A theory of disappointment aversion, *Econometrica*, 59, 3, 667-686.

[67] Hackbarth, Dirk, 2008, Managerial traits and capital structure decisions, *Journal of Financial and Quantitative Analysis*, forthcoming.

[68] Hall, Brian, 1998, A better way to pay CEOs?, In: Yermack, David and Jennifer Carpenter (Eds.), *Executive Compensation and Shareholder Value: Theory and Evidence*, Kluwer Academic Publishers, 33-44.

[69] Hall, Brian and Thomas Knox, 2004, Underwater options and the dynamics of executive pay-to-performance sensitivities, *Journal of Accounting Research*, 42, 2, 365-412.

[70] Hall, Brian and Jeffrey Liebman, 1998, Are CEOs really paid like bureaucrats?, *Quarterly Journal of Economics*, 112, 3, 653-691.

[71] Hall, Brian and Kevin Murphy, 2000, Optimal exercise prices for executive stock options, *American Economic Review*, 90, 2, 209-214.

[72] Hall, Brian and Kevin Murphy, 2002, Stock options for undiversified executives, *Journal of Accounting and Economics*, 33, 2, 3-42.

[73] Hall, Brian and Kevin Murphy, 2003, The trouble with stock options, *Journal of Economic Perspectives*, 17, 3, 49-70.

[74] Hallock, Kevin and Craig Olson, 2006, The value of stock options to non-executive employees, NBER working paper 11950.

[75] Haigh, Michael and John List, 2005, Do professional traders exhibit myopic loss aversion? An experimental analysis, *Journal of Finance*, 49, 1, 523-534.

[76] Haubrich, Joseph, 1994, Risk aversion, performance pay, and the principal-agent-problem, *Journal of Political Economy*, 102, 2, 258-275.

[77] Haubrich, Joseph and Ivilina Popova, 1998, Executive compensation: A calibration approach, *Economic Theory*, 12, 3, 561-581.

[78] Hausch, Donald and William Ziemba, 1995, Efficiency in sports and lottery betting markets, In: Jarrow, Robert, Maksimovic, Voijslav, and William Ziemba (Eds.), *Handbook of Finance*, Amsterdam: North Holland.

[79] Heath, Chip, Huddart, Stephen, and Mark Lang, 1999, Psychological factors and stock option exercise, *Quarterly Journal of Economics*, 114, 2, 601-627.

[80] Heath, Chip and Amos Tversky, 1991, Preference and belief: Ambiguity and competence in choice under uncertainty, *Journal of Risk and Uncertainty*, 4, 5-28.

[81] Heaton, James, 2002, Managerial optimism and corporate finance, *Financial Management*, 31, 2, 33-45.

[82] Hemmer, Thomas, Kim, Oliver, and Robert Verrecchia, 1999, Introducing convexity into optimal compensation contracts, *Journal of Accounting and Economics*, 28, 3, 307-327.

[83] Hodge, Frank, Rajgopal, Shiva, and Terry Shevlin, 2006, How do managers value stock options and restricted stock?, Working paper, Washington Business School.

[84] Holmström, Bengt, 1979, Moral hazard and observability, *Bell Journal of Economics*, 10, 1, 74-91.

[85] Holmström, Bengt, 1982, Moral hazard in teams, *Bell Journal of Economics*, 13, 2, 324-340.

[86] Holmström, Bengt and Steven Kaplan, 2001, Corporate governance and merger activity in the united states: making sense of the 1980s and 1990s, *Journal of Economic Perspectives*, 15, 2, 121-144.

[87] Holmström, Bengt and Paul Milgrom, 1987, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica*, 55, 2, 303-328.

[88] Holmström, Bengt and Paul Milgrom, 1991, Multitask principal-agent analyses: incentive contracts, asset ownership, and job design, *Journal of Law, Economics, and Organization*, 7, 24-52.

[89] Huddard, Stephen and Mark Lang, 1996, Employee stock option exercises: An empirical analysis, *Journal of Accounting and Economics*, 21, 1, 5-43.

[90] Inderst, Roman and Holger Müller, 2005, A theory of broad-based option pay, CEPR Discussion Paper 4878.

[91] Inderst, Roman and Holger Müller, 2007, Safeguarding specific human capital investments, Working paper, New York University.

[92] Ingersoll, Jonathan, 2006, The subjective and objective evaluation of incentive stock options, *Journal of Business*, 79, 2, 453-488.

[93] Ingersoll, Jonathan, 2008, Non-monotonicity of the Tversky-Kahneman probability-weighting function: A cautionary note, *European Financial Management*, forthcoming.

[94] Ittner, Christopher, Lambert Richard, and Larcker David, 2003, The structure and performance consequences of equity grants to employees of new economy firms, *Journal of Accounting and Economics*, 34, 89-127.

[95] Jensen, Michael, 1986, Agency costs of free cash flow, corporate finance, and takeovers, *American Economic Review*, 76, 2, 323-329.

[96] Jensen, Michael and William Meckling, 1976, Theory of the firm: managerial behavior, agency costs and ownership structure, *Journal of Financial Economics*, 3, 4, 305-360.

[97] Jensen, Michael and Kevin Murphy, 1990, Performance pay and top management incentives, *Journal of Political Economy*, 98, 2, 225-264.

[98] Jensen, Michael and Kevin Murphy, 2004, Remuneration: where we've been, how we got to here, what are the problems, and how to fix them, ECGI-Finance Working Paper No. 44/2004.

[99] Jewitt, Ian, 1988, Justifying the first-order approach to principal-agent problems, *Econometrica*, 56, 5, 1177-1190.

[100] Jones, Derek, Kalmi, Panu, and Mikko Mäkinen, 2006, The determinants of stock option compensation: Evidence from Finland, *Industrial Relations*, 45, 3, 437-468.

[101] Jullien, Bruno and Bernard Salanie, 2000, Estimating preferences under risk: The case of racetrack bettors, *Journal of Political Economy*, 108, 3, 503-530.

[102] Kahle, Kathleen, 2002, When a buyback isn't a buyback: open market repurchases and employee options, *Journal of Financial Economics*, 63, 235-261.

[103] Kahneman, Daniel, Knetsch, Jack, and Richard Thaler, 1986, Fairness as a constraint on profit seeking: Entitlements in the market, *American Economic Review*, 76, 4, 728-741.

[104] Kahneman, Daniel and Dan Lovallo, 1993, Timid choices and bold forecasts: A cognitive perspective on risk taking, *Management Science*, 39, 1, 17-31.

[105] Kahneman, Daniel and Amos Tversky, 1979, Prospect Theory: An analysis of decision under risk, *Econometrica*, 47, 2, 263-292.

[106] Kahneman, Daniel and Amos Tversky, 1984, Choices, values, and frames, *American Psychologist*, 39, 4, 341-350.

[107] Kahnemann, Daniel and Amos Tversky, 2000, *Choices, values, and frames*, Cambridge University Press.

[108] Kaplan, Steven, 1994, Top executive rewards and firm performance: A comparison of Japan and the United States, *Journal of Political Economy*, 102, 3, 510-545.

[109] Kaplan, Steven and Luigi Zingales, 1997, Do investment-cash flow sensitivities provide useful measures of financing constraints?, *Quarterly Journal of Economics*, 112, 1, 169-215.

[110] Kedia, Simi and Abon Mozumdar, 2002, Performance impact of employee stock options, Working paper, Harvard Business School.

[111] Keppe, Hans-Jürgen and Martin Weber, 1995, Judged knowledge and ambiguity aversion, *Theory and Decision*, 39, 1, 51-77.

[112] Knetsch, Jack, 1989, The endowment effect and evidence of nonreversible indifference curves, *American Economic Review*, 79, 5, 1277-1284.

[113] Knetsch, Jack and Jack Sinden, 1984, Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value, *Quarterly Journal of Economics*, 99, 3, 507-521.

[114] Kouwenberg, Roy and William Ziemba, 2008, Incentives and risk taking in hedge funds, *Journal of Banking and Finance*, forthcoming.

[115] Kroumova, Maya and James Sesil, 2006, Intellectual capital, monitoring, and risk: What predicts the adoption of employee stock options?, *Industrial Relations*, 45, 4, 734-752.

[116] Lambert, Richard and David Larcker, 2001, How do employees value (often incorrectly) their stock options, available at: knowledge@wharton.

[117] Lambert, Richard and David Larcker, 2004, Stock options, restricted stock, and incentives, Working paper, University of Pennsylvania.

[118] Lambert, Richard, Larcker, David, and Robert Verrecchia, 1991, Portfolio considerations in valuing executive compensation, *Journal of Accounting Research*, 29, 1, 129-149.

[119] Lamont, Owen, Polk, Christopher, and Jesus Saa-Requejo, 2001, Financial constraints and stock returns, *Review of Financial Studies*, 14, 2, 529–554.

[120] Langer, Thomas and Martin Weber, 2001, Prospect Theory, mental accounting, and differences in aggregated and segregated evaluation of lottery portfolios, *Management Science*, 47, 5, 716-733.

[121] Lazear, Edward, 2005, Output-based pay: incentives or sorting, *Research in Labor Economics*, 23, 1, 1-25.

[122] List, John, 2004, Neoclassical Theory versus Prospect Theory: Evidence from the marketplace, *Econometrica*, 72, 2, 615-625.

[123] Ljungqvist, Alexander and William Wilhelm, 2005, Does Prospect Theory explain IPO market behavior?, *Journal of Finance*, 49, 4, 1759-1790.

[124] Loewenstein, George and Daniel Adler, 1995, A bias in the prediction of tastes, *Economic Journal*, 105, 431, 929-937.

[125] Loomes, Graham and Robert Sugden, 1982, Regret theory: an alternative theory of rational choice under uncertainty, *Economic Journal*, 92, 368, 805-824.

[126] Machina, Mark, 1982, "Expected utility" analysis without the independence axiom, *Econometrica*, 50, 2, 277-324.

[127] Malmendier, Ulrike and Geoffrey Tate, 2005a, CEO overconfidence and corporate investment, *Journal of Finance*, 90, 6, 2661-2700.

[128] Malmendier, Ulrike and Geoffrey Tate, 2005b, Does overconfidence affect corporate investment? CEO overconfidence measures revisited, *European Financial Management*, 11, 5, 649-659.

[129] Malmendier, Ulrike and Geoffrey Tate, 2008, Who makes acquisitions? CEO overconfidence and the market's reaction, *Journal of Financial Economics*, forthcoming.

[130] Margiotta, Mary and Robert Miller, 2000, Managerial compensation and the cost of moral hazard, *International Economic Review*, 41, 3, 669-719.

[131] Massa, Massimo and Andrei Simonov, 2005, Behavioral biases and investment, *Review of Finance*, 9, 4, 483-507.

[132] McNeil, Barbara, Pauker, Stephen, Sox, Harold, and Amos Tversky, 1982, On the elicitation of preferences for alternative therapies, *New England Journal of Medicine*, 306, 21, 1259-62.

[133] Meulbroek, Lisa, 2001, The efficiency of equity-linked compensation: Understanding the full cost of awarding executive stock options, *Financial Management*, 30, 2, 5-30.

[134] Mirrlees, James, 1974, Notes on welfare economics, information, and uncertainty, In: Balch, Michael, McFadden, Daniel, and Shih-yen Wu (Eds.), *Essays on economic behavior under uncertainty*, Amsterdam: North-Holland.

[135] Mirrlees, James, 1976, The optimal structure of incentives and authority within an organization, *Bell Journal of Economics*, 7, 1, 105-131.

[136] Murphy, Kevin, 1999, Executive compensation, *Handbook of Labour Economics*, 3, 2, 2486-2563.

[137] Myagkov, Mikhail and Charles Plott, 1997, Exchange economies and loss exposure: Experiments exploring Prospect Theory and competitive equilibria in market environments, *American Economic Review*, 87, 5, 801-828.

[138] Nagaoka, Sadao, 2005, Determinants of the introduction of stock options by Japanese firms: Analysis from the incentive and selection perspectives, *Journal of Business*, 78, 6, 2289-2315.

[139] Neilson, William, 1992, A mixed fan hypothesis and its implications for behaviour towards risk, *Journal of Economic Behavior and Organization*, 19, 2, 197-211.

[140] Neumann, John von and Oskar Morgenstern, 1944, *Theory of games and economic behavior*, Princeton University Press.

[141] Odean, Terance, 1998, Are investors reluctant to realize their losses?, *Journal of Finance*, 53, 5, 1775-1798.

[142] Oyer, Paul, 2004, Why do firms use incentives that have no incentive effects?, *Journal of Finance*, 59, 4, 1619-1650.

[143] Oyer, Paul and Scott Schaefer, 2005, Why do some firms give stock options to all employees? An empirical examination of alternative theories, *Journal of Financial Economics*, 76, 1, 99-133.

[144] Oyer, Paul and Scott Schaefer, 2006, Costs of broad-based stock option plans, *Journal of Financial Intermediation*, 15, 4, 511-534.

[145] Plott, Charles, 1996, Rational individual behavior in markets and social choice processes: the Discovered Preference Hypothesis, In: Arrow, Kenneth, Colombatto, Enrico, Perlman, Mark, and Christian Schmidt (Eds.), *The rational foundations of economic behavior*, London: Macmillan and New York: St Martin's Press, 225-250.

[146] Plott, Charles and Kathryn Zeiler, 2005, The willingness to pay-willingness to accept gap, the "endowment effect", subject misconceptions, and experimental procedures for eliciting valuations, *American Economic Review*, 95, 3, 530-545.

[147] Polkovnichenko, Valery, 2005, Household portfolio diversification: A case for rank-dependent preferences, *Review of Financial Studies*, 18, 4, 1467-1502.

[148] Post, Thierry, van den Assem, Martijn, Baltussen, Guido, and Richard Thaler, 2007, Deal or no deal? Decision making under risk in a large-payoff game show, *American Economic Review*, 98, 1, 38-71.

[149] Prelec, Drazen, 1998, The probability weighting function, *Econometrica*, 66, 3, 497-527.

[150] Prendergast, Canice, 1999, The provision of incentives in firms, *Journal of Economic Literature*, 37, 1, 7-63.

[151] Prendergast, Canice, 2002, The tenuous trade-off between risk and incentives, *Journal of Political Economy*, 110, 5, 1071-1102.

[152] Rabin, Matthew, 1998, Psychology and economics, *Journal of Economic Literature*, 36, 1, 11-46.

[153] Rabin, Matthew, 2000, Risk aversion and expected-utility theory: A calibration theorem, *Econometrica*, 68, 5, 1281-1291.

[154] Rabin, Matthew, 2002, A perspective on psychology and economics, *European Economic Review*, 46, 657-685.

[155] Rappaport, Alfred, 1999, New thinking on how to link executive pay with performance, *Harvard Business Review*, 72, 91-101.

[156] Ray, Debraj, 2002, The time structure of self-enforcing agreements, *Econometrica*, 70, 2, 547-582.

[157] Rayo, Luis and Gary Becker, 2007, Evolutionary efficiency and happiness, *Journal of Political Economy*, 115, 2, 302-337.

[158] Rogerson, William, 1985, The first-order approach to principal-agent problems, *Econometrica*, 53, 6, 1357-1367.

[159] Ross, Stephen, 1973, The economic theory of agency: the principal's problem, *American Economic Review*, 63, 2, 134-139.

[160] Samuelson, William and Richard Zeckhauser, 1988, Status quo bias in decision making, *Journal of Risk and Uncertainty*, 1, 4-5, 7-59.

[161] Sautner, Zacharias and Martin Weber, 2005, Stock options and employee behavior, Working paper, University of Mannheim.

[162] Sawers, Kimberly, Wright, Arnold, and Valentina Zamora, 2006, Loss aversion, stock-based compensation and managerial risk-seeking behavior, Working paper, Boston College.

[163] Schoemaker, Paul, 1982, The expected utility model: its variants, purposes, evidence and limitations, *Journal of Economic Literature*, 20, 2, 529-563.

[164] Segal, Uzi and Avi Spivak, 1990, First order versus second order risk aversion, *Journal of Economic Theory*, 51, 1, 111–125.

[165] Starmer, Chris, 2000, Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk, *Journal of Economic Literature*, 38, 2, 332-382.

[166] Thaler, Richard, 1980, Toward a positive theory of consumer choice, *Journal of Economic Behavior and Organization*, 1, 1, 39-60.

[167] Thaler, Richard, 1993, *Advances in behavioral finance*, Russell Sage Foundation.

[168] Thaler, Richard, 1999, Mental accounting matters, *Journal of Behavioral Decision Making*, 12, 3, 183-206.

[169] Thaler, Richard, 2005, *Advances in behavioral finance, Volume II*, Princeton University Press.

[170] Thaler, Richard and William Ziemba, 1988, Pari-mutuel betting markets: Racetracks and lotteries, *Journal of Economic Perspectives*, 2, 2, 161-174.

[171] Tversky, Amos and Daniel Kahneman, 1974, Judgment under uncertainty: heuristics and biases, *Science*, 185, 4157, 1124-1131.

[172] Tversky, Amos and Daniel Kahneman, 1986, Rational choice and the framing of decisions, *Journal of Business*, 59, 4, 251-278.

[173] Tversky, Amos and Daniel Kahneman, 1991, Loss aversion in riskless choice: A reference-dependent model, *Quarterly Journal of Economics*, 106, 4, 1039-1061.

[174] Tversky, Amos and Daniel Kahneman, 1992, Advances in Prospect Theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5, 4, 297-323.

[175] Van de Kuilen, Gijs and Peter Wakker, 2006, Learning in the Allais paradox, *Journal of Risk and Uncertainty*, 33, 155-164.

[176] Van de Kuilen, Gijs, 2008, Subjective probability weighting & the Discovered Preference Hypothesis, *Theory and Decision*, forthcoming.

[177] Wakker, Peter and Horst Zank, 2002, A simple preference foundation of cumulative prospect theory with power utility, *European Economic Review*, 46, 7, 1253-1271.

[178] Weisbach, Michael, 1988, Outside directors and CEO turnover, *Journal of Financial Economics*, 20, 1, 431-460.

[179] Wu, George, Zhang, Jiao, and Richard Gonzalez, 2004, Decision under risk, In: Koehler, Derek and Nigel Harvey (Eds.), *Blackwell Handbook of Judgement and Decision Making*, Blackwell, 399-423.

[180] Yermack, David, 1995, Do corporations award CEO stock options effectively?, *Journal of Financial Economics*, 39, 2-3, 237-269.

Kurzlebenslauf

Oliver Spalt

Geboren am 27. März 1979 in Esslingen am Neckar.

Anschrift: Wilhelm-Blos-Strasse 10, 68199 Mannheim.

Schulausbildung und akademischer Werdegang

2006 - 2008	UNIVERSITÄT MANNHEIM
	Wissenschaftlicher Mitarbeiter am Lehrstuhl für Corporate Finance (Professor Ernst Maug, Ph.D.) und Doktorand am Center for Doctoral Studies in Business
2004 - 2006	HUMBOLDT-UNIVERSITÄT ZU BERLIN
	Wissenschaftlicher Mitarbeiter am Lehrstuhl für Konzernmanagement (Professor Ernst Maug, Ph.D.) und Doktorand im Berlin Doctoral Program in Economics and Management Science
1999 - 2004	UNIVERSITÄT MANNHEIM
	Studium der BWL mit interkultureller Qualifikation (Englisch)
	- Abschluß als Diplom-Kaufmann
	- Preis der Prechel-Stiftung für das beste Diplom
	- Austauschstudent 2002 - 2003 an der Warwick Business School, GB
1998	Abitur, Ludwig-Uhland-Gymnasium, Kirchheim unter Teck