Risk Premia in Executive Compensation: A Life-Cycle Perspective

Florian S. Peters*

UC Berkeley, University of Zurich and SFI

November 2009

JOB MARKET PAPER

Abstract

How much of the rise in CEO pay can be explained by the increased risk that CEOs are exposed to? This paper employs a life-cycle model of consumption and saving to answer this question, and, more broadly, to study the relationship between the risk and level of CEO compensation. The model incorporates the main types of risk that executives of public corporations face: option- and stock-based pay, pay-performance sensitivity, dismissal risk, and stock return volatility. I use the model to compute risk premia in pay levels, and analyze how well they explain the observed variation in CEO pay. A calibration to a large panel of CEOs shows that, for realistic degrees of risk aversion, risk premia explain about 20 percent of the variation in CEO pay, both in the cross-section and the time-series. The model captures the higher moments of the cross-sectional pay distribution particularly well. The structural framework provided in this paper is robust to some sources of endogeneity typically encountered in reduced-form empirical research, and allows for the welfare analysis of policy interventions such pay limits for executives.

*University of California - Berkeley, University of Zurich and Swiss Finance Institute (SFI). Address: 508-1 Evans Hall #3880, Berkeley, CA 94720-3880. Email: fpeters@econ.berkeley.edu. I am especially grateful to Ulrike Malmendier and Stefano DellaVigna for their support, guidance, and the many valuable comments and suggestions. I thank Dirk Jenter and Fadi Kanaan for generously sharing their data. I thank Michel Habib, Yuriy Gorodnichenko, Pat Kline, Martin Lettau, David Sraer and Adam Szeidl for helpful discussions and participants at the Berkeley Financial Economics Seminar and Macro Lunch for comments. I am also grateful to Sebastian Stumpner and Johannes Wieland for discussing many details of the paper. Financial support from the Swiss Finance Institute and the University of Zurich is gratefully acknowledged.

1 Introduction

The extraordinary rise in executive compensation over the past decades has spurred an intense debate about the underlying causes of this phenomenon. Average CEO pay in S&P 500 firms has risen seven-fold between 1980 and 2000.[1] For the broader, more recent sample of S&P 1500 CEOs covered by Execucomp, total compensation increased by 150 percent between 1995 and 2000 alone. While declining for three years following the burst of the dot-com bubble, CEO pay started to rise again in 2003 until the break-out of the financial crisis, and remains exceptionally high by historical standards. The common perception of exaggerated pay levels has led to several regulatory changes in the recent past, from the sweeping disclosure rules imposed by the SEC in 1992, to the recently announced pay limits in firms receiving government bailout funds.

The discussion about the causes of the rise in executive pay has been shaped by three fundamentally different views: Proponents of the entrenchment view have attributed this development to poor governance and the ability of CEOs to set their own pay (Bebchuk and Fried, 2004). Others have advocated market-based explanations based on the rising marginal product of managerial talent due to increasing scale (Rosen, 1981; Gabaix and Landier, 2008) and/or technical change that favors highly skilled or generally skilled individuals (Murphy and Zabojnik, 2004; Frydman, 2005; Kaplan and Rauh, 2009). In this paper, I focus on the third explanation, the widespread adoption of high-powered incentives, and the resulting increased exposure of CEOs to firm-specific risk (Hall and Liebman, 1998; Hermalin, 2005).

In the presence of risk aversion, increasing risk exposure of CEOs requires a rise in the expected value of compensation. Suggestive evidence for such a link comes from aggregate time-series data. Many of the risks that CEOs face show a non-monotonic development similar to that of CEO compensation. Forced CEO turnover rates increase from two to four percent between 1995 and 2001, then decline until 2003, to rise again thereafter. The fraction of option-based pay also doubles from 20 to 40 percent until 2000 and then declines. This declining exposure to option-induced risk is partially offset by the shift towards stock-based compensation after 2000. Finally, stock return volatility more than doubles between 1993 and 2002, and then falls back to the levels of the early 1990s. The time-series pattern of CEO pay and the various risks is illustrated in Figures 1-6.[2]

While the existence of a "risk-return tradeoff" in executive compensation is unsurprising, quantifying it is a more difficult task. Ideally, one would estimate the risk premium

[1] These numbers are from Jensen, Murphy, and Wruck (2004) who calculate total compensation for S&P 500 CEOs based on data from Forbes and Execucomp.

[2] Contrary to all other types of risk, pay-performance sensitivity does not show a pronounced trend over the period 1995-2005.

in a regression of CEO pay on the various risks using exogenous variation in the risks as a source of identification. However, all of these risks are endogenous. For example, equity incentives are set by the board based on determinants that are, at least partly, unobservable by the researcher, and dismissal risk is likely correlated with unobservable CEO characteristics. So far, instrumental variable approaches have failed to provide a viable solution to the inference problem.

In this paper, I use a model-based calibration approach to study the relation between the risk and level of CEO pay. I develop a life-cycle model which incorporates, in a calibratable way, the various risks that CEOs face: Risks to wealth due to the option- and stock-based components of pay; risks to income due to dismissal and pay-performance sensitivity; and risks to both income and wealth that result from stock return volatility. The solution of the life-cycle model yields the lifetime utility, or value function, which depends on the exposure to the various types of risk. Using the value function, I compute the proportional risk premium on the level of pay that would make a CEO indifferent between the risks he actually faces and an alternative safe income stream.

With the methodology in hand, I then investigate how much of the observed variation in CEO pay can be explained by model-implied risk premia. To do this, I calibrate the model to a panel dataset of CEOs of the largest public US firms for the period 1995 to 2005 and compute risk premia for each CEO and year. The core of my analysis consists of the comparison of the model-implied, purely risk-based variation, with the observed variation in CEO pay.

I analyze a number of dimensions of the relationship between the risk and level of CEO pay: (i) various moments including the correlation between the risk and level of pay, the variance as well as the higher moments of total pay and risk premia; (ii) the contribution of the individual types of risk to the total variation in CEO pay; and (iii) cross-sectional and time-series relationships. The model-implied risk premia, and hence all sample statistics computed thereof, depend on the degree of relative risk aversion. For relative risk aversion of three and homogeneity of risk aversion across CEOs, I obtain the following results:

- *Correlation.* Empirical correlations appear to yield biased estimates of the causal effect of the various risks on CEO pay. The empirical estimates of option and stock risk premia are significantly larger than their model counterparts. Conversely, the empirical estimates of dismissal and volatility risk premia are much smaller than their model analogues. One interpretation of this finding is that equity-based pay is positively correlated with unobserved skill and effort, while dismissal risk and volatility are negatively correlated with these characteristics.

- *Variance.* Model-implied risk premia explain about 20 percent of the cross-sectional variation in CEO pay. They explain about 30 percent of the *residual* variation in CEO pay once firm size is accounted for. Due to the fact that risk aversion scales the relation between risk and risk premia, the explanatory power of risk premia for observed CEO pay rises with the degree of risk aversion.

- *Higher moments.* Similar to the distribution of observed CEO pay, model-implied risk premia are heavily right-skewed and fat-tailed. Skewness and kurtosis coefficients of risk premia are very close to their empirical counterparts. There are two reasons for this feature: First, the cross-sectional distributions of the some of the risks themselves are skewed. Second, with relative risk aversion risk premia are convex in the size of the risks. Hence even normally distributed risks generate skewed distributions of risk premia.

- *Relative contributions of the various risks.* Equity-based pay, in particular stock options, contribute by far the most to the sample variation in the total risk premium. While the premium for dismissal risk is large for *unit* changes in risk, the sample variation in dismissal risk is too small to explain a significant portion of the total variation in pay. Pay-performance sensitivity appears to be unimportant for explaining variation in CEO pay, as it exhibits both small unit risk premia and small sample variation.

- *Risk premia and the development of average pay levels.* Average risk premia rise steadily from 1995 to 2001 and revert thereafter. Risk premia can explain a 30 percent increase in pay levels from 1995 to 2001. The development over the entire 1995-2005 period can be well-explained by risk and firm size as the main determinants of CEO pay.

In short, the calibrations suggest that risk is an important factor in determining executive compensation. In terms of cross-sectional explanatory power, risk and firm size appear to be of similar importance.[3] In the time-series, risk explains a smaller part of the increase in pay levels than does size, although risk does help to explain the non-monotonic pattern over 1995-2005.[4] Finally, risk helps explain the pronounced inequality in the income distribution of executives of the largest US corporations, a fact which has led to a renewed focus on the superstar status of some CEOs (Rosen, 1981; Gabaix and Landier, 2008; Malmendier and Tate, 2009).

[3]Calibrations of the Gabaix and Landier (2008) model as well as cross-sectional regressions of log CEO pay on log firm size yield R-squared statistics between 20 and 40 percent.

[4]It is difficult to assess the quantitative importance of explanations based on entrenchment or the shift towards general skills since these models have not been calibrated.

Regarding the current policy debate about pay limits for executives, the results of this paper provide a quantitative estimate of the "fair", or utility-neutral, compensation for risk. Because the recently imposed pay limits pertain to cash compensation only, they will likely be accompanied by even larger stock grants than would be necessary without these restrictions in order to retain managerial talent.

The structural approach taken in this paper has the advantage of being more robust to endogeneity than reduced-form empirical methods, given the lack of instruments. For example, estimating the option risk premium in a regression of CEO pay on the fraction of options in total compensation will likely overstate the causal effect of risk on pay. Option-based pay might be correlated with partially unobservable skill, if more highly skilled executives select into jobs with high-powered incentives. In addition, stronger incentives induce higher effort, the costs of which also need to be compensated. In this case, the regression coefficient of option-based pay will include skill and effort premia, in addition to the risk premium. In contrast, a model that realistically captures how CEOs perceive risks can quantify the premium needed to compensate for risk alone, while being unaffected by omitted, model-external variables. On the other hand, endogeneity remains a concern if it is driven by model-external characteristics that affect the risk premium itself. For example, if highly skilled executives tend to select into jobs with high-powered incentives *and* are also less risk-averse, then the model, which does not capture such selection, would overstate the risk premium on average.

This paper makes several contributions to the economics and finance literature. First, it adds to the large body of research studying the determinants of CEO pay and the causes of the extraordinary rise of average pay levels. The most prominent explanations include rising firm size (Gabaix and Landier, 2008; Terviö, 2009), the shift towards general skills needed to manage large corporations (Murphy and Zabojnik, 2004; Frydman, 2005), managerial entrenchment (Bebchuk and Fried, 2004; Kuhnen and Zwiebel, 2008), and the greater risks imposed on CEOs due to the push towards stronger incentives (Hall and Liebman, 1998; Hermalin, 2005). My results provide a quantitative assessment of the role that risk plays in determining CEO pay. The risk premia calculated from the structural model can also be viewed as a benchmark for "justified" variation in CEO pay. In that sense, the framework may be useful to compute excess compensation, i.e. pay levels that cannot be explained by economically sound determinants.

Second, this paper introduces life-cycle models into the study of executive compensation. To date, life-cycle models have mainly been used in macroeconomics to study aggregate consumption and savings patterns (Zeldes, 1989; Deaton, 1991; Carroll, 1992; Hubbard, Skinner, and Zeldes, 1994; Carroll and Samwick, 1997) or to estimate preference parameters (Gourinchas and Parker, 2002; Laibson, Repetto, and Tobacman, 2007).

Rather than studying the choice of consumption and savings, this paper focuses on the utility implied by optimal choice under risk.

Besides the particular modeling framework, this paper extends the use of calibrated structural models to study the determinants of the level of CEO pay. The link between the risk and level of executive compensation under relative risk aversion has not been extensively calibrated.[5] Following the seminal contribution of Haubrich (1994), a number of recent papers employ calibrations of principal-agent models to study the *structure* of CEO pay, i.e. the optimal mix of stocks, options and cash salary (Armstrong, Larcker, and Su, 2007; Dittmann and Maug, 2007; Edmans, Gabaix, and Landier, 2008; Dittmann, Maug, and Spalt, 2009; Dittmann and Yu, 2009). This paper takes incentives as given and explores the implications of the cross-sectional and time-series variation in the risks implied by these incentives for the *level* of CEO pay.

Finally, this paper is related to the literature on the subjective valuation of executive stock options. A number of papers have used utility-based approaches to value executive stock options for undiversified executives (Lambert, Larcker, and Verrecchia, 1991; Huddart, 1994; Carpenter, 1998; Hall and Murphy, 2002; Ingersoll, 2006). These papers typically focus on the valuation of a particular option and the utility generated by the payoff at the expiry date, abstracting from any income generated between the grant and expiry dates of the option. The life-cycle model presented in this paper takes into account other sources of risk such as dismissal and pay-performance sensitivity, and allows the CEO to dynamically respond to these risks by adjusting consumption and savings according to the income and wealth realizations in each period.

The rest of the paper is organized as follows: The next section develops a life-cycle model for CEOs and describes the computation of risk premia. Section 3 describes the calibration procedure and the dataset used in this study. The results are presented in Sections 4, which contains the cross-sectional analysis, and Section 5, which presents the time-series results. Section 6 concludes.

2 A Life-cycle Model with Incentive Pay

In this section I present a life-cycle model that incorporates the main types of risk that CEOs of large public corporations are exposed to: Option- and stock-based pay, pay-performance sensitivity, stock return volatility, and dismissal. The model builds on Chapman, Evans, and Xu (2009), who study the effects of career concerns on fund managers'

[5]Gayle and Miller (2009) estimate a dynamic agency model of executive compensation assuming constant absolute risk aversion.

investment decisions, but differs from theirs in at least two respects: First, it includes constraints on CEOs' savings decisions in order to capture the effects of vesting periods of equity-based pay. Second, it focuses on the welfare implications of risk exposure rather than its effect on the manager's firm-related decisions. I first lay out the model, and then describe the computation of risk premia and the proposed relation between the risk and level of CEO pay. The technical details of the model solution are relegated to Appendix A.

2.1 The basic model

The starting point is the canonical, discrete-time, life-cycle model of consumption and saving behavior. Calendar time is indexed by subscript t, firms and CEOs are indexed by subscript i. The CEO lives for N periods and works for $T < N$ periods.

The income process. In every year of his working life, the CEO receives a stochastic income, Y_t, with the following dynamics:

$$Y_t = Y_{t-1} G_t \phi(r_t) \theta(r_t) \tag{1}$$

Income growth is determined by three components. First, it evolves according to an deterministic growth factor, G_t, which can be thought of as the expected premium for an additional year of professional experience. Second, conditional on the CEO keeping his job, income growth depends on the performance of the firm's stock price over the current year, r_t. This stochastic component, denoted by $\phi(r_t)$, reflects the pay-performance sensitivity typically observed in executive compensation. Third, the manager receives a discrete negative income shock if he is fired. The dismissal variable, $\theta(r_t)$, takes two values with the respective probabilities $q(r_t)$ and $1 - q(r_t)$:

$$\theta(r_t) = \begin{cases} \lambda < 1 & \text{dismissal, with probability } q(r_t) \\ 1 & \text{no dismissal, with probability } 1 - q(r_t) \end{cases} \tag{2}$$

Thus, the monetary risk of firing is reflected in both the probability of this event, $q(r_t)$, as well as the magnitude of the drop in income following dismissal, $1 - \lambda$.

The conditional probability of dismissal, $q(r_t)$, can be estimated by a probit regression of a dichotomous variable indicating dismissal on the stock return. The size of the shock, λ, can be separately estimated from observed income effects of dismissals in the CEO labor market. The exact empirical parametrization and estimation of all three components of income dynamics is described in detail in Section 3.

The return on savings. The savings process is modeled so as to explicitly account for CEOs' mandatory holdings of own-firm stocks and options. Such holdings result from the vesting periods on the equity-based part of the compensation package.[6] The dynamics of savings during the CEO's working life are specified as follows. Let R_{t+1} be the gross, discrete-time return on the manager's personal portfolio over the period t to $t+1$. Further, denote by η_f, η_s, and η_z the manager's portfolio shares in the risk free asset, own-firm stocks, and own-firm options, respectively, and let r_f, r_{t+1}, and $r_{z,t+1}$ denote the returns of the risk free asset, the firm's stock, and stock options over t to t+1. Then savings accumulate according to

$$\begin{aligned} R_{t+1} &= 1 + \eta_f r_f + \eta_s r_{t+1} + \eta_z r_{z,t+1} \\ &\approx 1 + \eta_f r_f + \eta_s r_{t+1} + \eta_z \varepsilon_{z,s} r_{t+1} \\ &\approx e^{\eta_f r_f + (\eta_s + \eta_z \varepsilon_{z,s}) r_{t+1}} \end{aligned} \tag{3}$$

The first line expresses the portfolio return as the weighted sum of the returns on three components: the risk-free asset, own-firm stocks, and own-firm options. The second line replaces the return of the stock options with its first order approximation, the stock return times the elasticity of the option price with respect to the stock price, $\varepsilon_{z,s}$.[7] The third line approximates discrete time returns by log returns. This last formulation ensures that gross returns cannot be negative. The term $\eta_s + \eta_z \varepsilon_{z,s}$ can be interpreted as the (portfolio share) weighted elasticity of total wealth with respect to the stock price: $\eta_f \varepsilon_{f,s} + \eta_s \varepsilon_{s,s} + \eta_z \varepsilon_{z,s} = \eta_f \cdot 0 + \eta_s \cdot 1 + \eta_z \cdot \varepsilon_{z,s} = \eta_s + \eta_z \varepsilon_{z,s}$. Thus the riskiness of the manager's personal portfolio depends on the shares held in own-firm stocks and options, the elasticity of the option price with respect to the stock price, and the stock return distribution. At retirement all company stocks and options vest and savings grow at the risk-free rate.

The budget constraint. The CEO's budget constraint reflects the risks of both labor income and the returns to his personal portfolio. Denote by M_t the CEO's financial wealth available for consumption at the beginning of period t (cash-on-hand), and by A_t the remaining wealth (=savings) at the end of period t, after income is realized and consumption

[6]In practice, CEOs often hold more equity than is required by vesting restrictions. This can be due to a number of reasons such as signalling, overconfidence, etc. In this study I use only the fraction of wealth a CEO is forced to hold in own-firm stock, since only the risk induced by that fraction should be compensated with higher levels of pay.

[7]The approximation can be made arbitrarily precise by including higher order terms of a Taylor expansion. The reason for using a Taylor approximation is to express the portfolio return as a function of one stochastic variable only.

is chosen. Financial wealth is the sum of accumulated savings and current-period income:

$$M_{t+1} = A_t R_{t+1} + Y_{t+1} \tag{4}$$

Substituting the expressions for Y_{t+1} and R_{t+1} from equations (1) and (3) yields

$$M_{t+1} = A_t e^{\eta_f r_f + (\eta_s + \eta_z \varepsilon_{z,s}) r_{t+1}} + Y_t G_{t+1} \phi(r_{t+1}) \theta(r_{t+1}) \tag{5}$$

Equation (5) shows that, in this model, all shocks to CEO wealth and income are correlated through the stock price. This leads to potentially bigger risks than result in the standard specification of the life-cycle model which assumes all shocks to be uncorrelated. In addition, in the above specification, the shock distribution is fat-tailed and left-skewed. This stems from the large negative shocks to income following dismissal. Together, these features imply much larger risk premia than would result from the standard specification of the life-cycle model.

Preferences. CEOs are assumed to derive instantaneous utility from the consumption of a single good according to the standard time-separable power utility with constant relative risk aversion (CRRA):

$$u(C_t) = \frac{C_t^{1-\rho}}{1-\rho} \tag{6}$$

where ρ is the coefficient of relative risk aversion.

CEOs' preferences over the life cycle are assumed to take the additively separable expected utility form with constant discount factor β:

$$E_t[\sum_{s=t}^{t+N} \beta^{s-t} u(C_s)] \tag{7}$$

The manager's problem. CEOs maximize expected lifetime utility given their initial wealth and income level, non-negativity constraints on consumption and terminal wealth, and the dynamic budget constraint. Note that in this setup CEOs will never choose to borrow against future income. This follows from (i) the fact that the income process has a lower bound of zero, and (ii) the Inada condition $\lim_{c \to 0} u'(c) = \infty$. To see this, suppose the CEO were to borrow in the next-to-last working period. Then, with a strictly positive probability he would be left without any wealth for the rest of his life. He would then have an infinite expected marginal utility. Backward induction then implies that it will never be optimal to borrow in any period. The CEO's optimization problem is then given by:

$$\max_{\{C_s\}_{s=t}^{t+N}} E_t[\sum_{s=t}^{t+N} \beta^{s-t} u(C_s)] \tag{8}$$

subject to

$$A_t = M_t - C_t \tag{9}$$

$$Y_{t+1} = Y_t \Gamma_{t+1} \tag{10}$$

$$\Gamma_{t+1} = \phi(r_{t+1})\theta(r_{t+1}) \tag{11}$$

$$R_{t+1} = e^{\eta_f r_f + (\eta_s + \eta_z \varepsilon_{z,s}) r_{t+1}} \tag{12}$$

$$M_{t+1} = A_t R_{t+1} + Y_{t+1} \tag{13}$$

The maximization problem has no closed form solution and has to be solved numerically and recursively. The details of the numerical solution are given in Appendix A. The solution consists of the optimal consumption as a function of wealth for each period. This, in turn, implies the value function which depends on the state variables current income, Y_t, and wealth, M_t, the CEO's years until retirement, T, and all other parameters of the model:

$$V(M_t, Y_t, T, \cdot) = \max_{\{C_s\}_{s=t}^{t+N}} E_t[\sum_{s=t}^{t+N} \beta^{s-t} u(C_s)] \tag{14}$$

2.2 The risk and level of CEO pay

The main question I address in this paper is how much CEOs should be compensated with higher expected levels of pay for bearing risks to income and wealth. To quantify this "risk-return" tradeoff in CEO compensation I compute the risk premium, π, as the percentage premium on expected income that makes a CEO indifferent between the risky income he receives and the alternative safe income, $\overline{Y}_i$. Technically, this is done by equating the value functions that result under the two different environments by scaling up the income under the risky environment:

$$V(M_i, \overline{Y}_i(1+\pi_i), T_i, ..., \sigma_i, ...) = V(M_i, \overline{Y}_i, T_i, ..., 0, ...) \tag{15}$$

The left hand side of the equation represents expected lifetime utility under the risky environment the CEO actually faces. Here, σ denotes the vector of parameters determining the risks to income and wealth. It includes (i) pay-performance sensitivity, (ii) the parameters determining the dismissal probability and the percentage drop in pay that dismissal entails, (iii) the fraction of unvested options and stocks the CEO expects to hold in his personal portfolio, and (iv) the mean and variance of the firm's stock return. The gross risk premium, $1 + \pi$, is the percentage premium that makes the CEO indifferent between this risky environment and the alternative safe environment, represented by the value function on the right hand side of equation (15).

The finance and economics literature often uses other measures to compare utility across risky environments such as certainty-equivalent income or certainty-equivalent consumption. While all of these metrics imply the same ranking of risky alternatives, the percentage premium on the pay level is more natural in the context of this study because it can readily be compared with its empirical counterpart, the actual pay level. In contrast, certainty-equivalent measures do not have observed sample analogues.

Equation (15) implicitly states that actual pay levels are determined as the product of the salary for a risk free job, $\overline{Y}_i$, and a proportional risk premium:

$$Y_i = \overline{Y}_i \times (1 + \pi(\sigma_i)) \tag{16}$$

Clearly, the level of pay varies across individuals depending on determinants other than risk, such as skill and experience, which are not captured by the life-cycle model. The effect of such determinants is subsumed in $\overline{Y}_i$, which is allowed to vary across individuals. Under CRRA utility, these other determinants do not affect the size of the risk premium to the extent that they are deterministic and uncorrelated with the risks. This feature implies that the model can separate the risk premium from other determinants of pay. Empirically, such model-external determinants can be controlled for in cross-sectional regressions. For instance, skill might be reflected in the size of the firm a CEO heads. In this case, $\overline{Y}_i$ can be expressed as a function of firm size. Sections 4 and 5 describe in more detail how I treat these model-external determinants empirically.

3 Data and calibration

3.1 Data

In order to calibrate the model and evaluate its performance I combine a number of data sources. I use the Execucomp database for the period 1993-2005 to obtain total CEO compensation, the option and stock-based components of pay, as well as CEO age. I then merge the compensation data with a hand-collected dataset on CEO turnovers. The turnover data covers all CEOs in Execucomp over the same period 1993-2005, and classifies turnovers as either forced or voluntary. It comprises a total of 639 forced and 1,855 voluntary CEO turnovers. Identifying a turnover as forced is not straightforward as firms almost never officially state a turnover as having occurred involuntarily. Classification thus requires hand-collection of data from multiple sources, in particular press releases.[8] Turnover data for 1993 to 2001 is from Jenter and Kanaan (2008). Data for 2002-2005 is from Peters and Wagner (2009).[9] Finally, I combine the resulting dataset with stock prices from CRSP.

Table 1 reports descriptive statistics for the final dataset. Panel A contains compensation statistics. Total compensation is truncated at the top and bottom 0.5 percent to reduce the influence of outliers. Total compensation is heavily skewed, with a mean that is about twice as large as the median ($4m and $2m, respectively). Option-based pay constitutes on average one third of the total market value of the pay package, while restricted stock represents only 6 percent and has a median of zero. The time-series pattern of total pay and its components is illustrated in Figure 1. The development of option- and stock-based pay over 1995-2005 is illustrated in Figures 3 and 4 using both equal and value-weights for computing the yearly means. Panel C of Table 1 reports basic firm characteristics. Panel D contains age and tenure statistics for CEOs.

[Table 1]

3.2 Calibration

I now turn to the calibration of the model parameters. The central inputs to the model are the parameters governing the various risks to income and wealth. I estimate these parameters separately for each CEO and year via regressions or simple sample moments using the relevant data over lagged two-year periods. This procedure contains the implicit assumption that CEOs form expectations about the future value of these parameters in the same way as the econometric procedure does. It is a statistical, adaptive concept of expectation formation broadly used in the modern macroeconomics literature (Evans and Honkapohja, 2001).

[8]The methodology follows (Parrino, 1997). All departures for which the press reports state that the CEO is fired, forced out, or retires or resigns due to policy differences or pressure, are classified as forced. Turnovers of CEOs below the age of 60 which have not been classified as forced by the press criterion, are classified as forced if the articles do not report the reason to be death, poor health, or acceptance of another position or the articles report that the CEO is retiring but does not announce the retirement date at least 6 months before the succession. For further details, see the original paper.

[9]I am grateful to Dirk Jenter and Fadi Kanaan for sharing their turnover data.

The various risks to CEO wealth operate through two main channels, income and savings. For ease of exposition, I restate the two main equations here:

$$Y_t = Y_{t-1} G_t \phi(r_t) \theta(r_t) \tag{17}$$

$$R_{t+1} \approx e^{\eta_f r_f + (\eta_s + \eta_z \varepsilon_{z,s}) r_{t+1}} \tag{18}$$

Pay-performance sensitivity and expected income growth rate. I specify both $\phi(r_t)$, the function governing pay-performance sensitivity, and G_t, the deterministic growth rate, as exponentials, and assume G_t to be constant over time:

$$\phi(r_t) = e^{\gamma_1 r_t}$$
$$G_t = e^{\gamma_0}$$

This is a standard specification in the empirical literature on pay-performance incentives (Gibbons and Murphy, 1992) and allows straightforward least-squares estimation. Conditional on the CEO keeping his job, i.e. $\theta(r_t) = 1$, equation (17) can be written as

$$\Delta \log(Y_t) = \gamma_0 + \gamma_1 r_t \tag{19}$$

The parameters γ_0 and γ_1 can be estimated by regressing the first differences of log CEO pay on the stock return, using the subsample of observations with no CEO turnovers.

In order to allow pay-performance sensitivity to vary across CEOs, I include interaction terms of the return with industry dummies, and the stock return variance. The latter choice is motivated by the theoretical result that boards should put less weight on the return as a signal of effort, the more noisy is the signal (Holmström and Milgrom, 1987). Instead of the stock return variance itself, I use its empirical cumulative distribution function, $\Phi(\sigma^2)$, as in Aggarwal and Samwick (1999). The rolling-regression procedure allows coefficients to also vary over time.

Table 2, Panel A, shows the results for the pay-performance regressions. The top rows contain estimates for two-year rolling window regressions, the bottom row contains the regression results for the entire period 1995-2005. Results are reported starting in 1995 due to the requirement of a two-year estimation window. The columns report the coefficients and standard errors for the three main regressors, the stock return, the variance and their interaction. Contrary to the Holmström-Milgrom prediction, the interaction of the variance with the stock return is statistically and economically insignificant in all rolling regressions, except for 2005, and it is only marginally significant when estimated over the entire sample period. Below, I show that the same interaction is highly significant in turnover-performance regressions. Thus, in order to keep regression specifications comparable, I use the complete specification including the interaction.

Figure 6 depicts the median cross-sectional pay-performance sensitivity over time. It is computed as the product of the return coefficient and 0.5 times the coefficient of the return-volatility interaction term. The figure illustrates that pay-performance sensitivity does not exhibit any particular trend over the sample period.

Conditional dismissal probability. Dismissal affects income through the term $\theta(r_t)$ in the income equation (17). Recall that

$$\theta(r_t) = \begin{cases} \lambda < 1 & \text{dismissal, with probability } q(r_t) \\ 1 & \text{no dismissal, with probability } 1 - q(r_t) \end{cases} \tag{20}$$

I estimate the conditional probability of dismissal using a the probit model:

$$q(r_{i,t}) = P[\theta_{i,t} = \lambda | r_{i,t}] = \Phi(\alpha_0 + \alpha_1 r_{i,t}). \tag{21}$$

As for the estimation of pay-performance sensitivities I also include the distribution function of the stock return variance, industry dummies as well as interactions of these variables with the stock return.

The results are shown in Table 2, Panel B. As expected, dismissal increases with poor performance. Column three indicates that, as predicted by agency theory, turnover-performance sensitivity decreases with the variance of the stock return.

Unconditional dismissal probability. The probit model described above models the dismissal probability *conditional* on the realization of the stock return. To obtain the unconditional probability of dismissal one needs to integrate over the stock return density:

$$P[\theta_{i,t} = \lambda] = \int \Phi(\alpha_0 + \alpha_1 r_{i,t}) \cdot f(r_{i,t}) dr_{i,t} \tag{22}$$

In addition to the parameters α_0 and α_1, the unconditional turnover probability also depends on the distribution of stock returns. Since, empirically, α_0 and α_1 are both negative, most probability mass of the argument $\alpha_0 + \alpha_1 r_{it-1}$ lies in the convex region of the distribution function $\Phi(\cdot)$. Thus, due to Jensen's inequality, higher variance increases the unconditional turnover probability for fixed α_0 and α_1.

Earnings losses of dismissed CEOs. Earnings consequences of dismissal for CEOs are difficult to measure, because SEC rules require income data to be disclosed for the five highest paid executives of *public* corporations only. Because it is difficult to track earnings of CEOs if they transferred to a private firm after being dismissed, sample attrition is potentially an issue. However, a number of related studies allow me to estimate the likely pay loss following turnover. Fee and Hadlock (2004) track the careers of CEOs around 205 turnovers. For the subsample of CEOs for which they are able to find post-turnover earnings, pay losses are about 20%. However, these are likely to be those CEOs who have gotten the best subsequent jobs. In the subsample for which the authors could identify the firm size of the subsequent employer, dismissed CEOs managed firms about one tenth the size of their former employers. If the coefficient of size of a regression of log compensation on firm size is used to predict the associated earnings loss, this implies earnings losses of about 75%. Another indication of the severe consequences of demotion comes from Chapman et al. (2009) who find that assets under management of mutual fund managers drop by 50% following performance-related demotions. Since compensation of mutual fund managers is typically proportionally linked to assets under management, this suggests that earnings drop by about the same amount. Finally, in a dataset free of sample attrition, Jacobson, LaLonde, and Sullivan (1993) find that workers displaced in mass-layoffs suffer earnings losses of 45-50% compared to their pre-displacement earnings. In the present analysis I use a value for this proportional drop of 50%, i.e. $\lambda = 0.5$.

Distribution of stock returns. I estimate the mean, μ_i, and the variance, σ_i^2, of stock returns using daily return data over the lagged two-year period preceding the year t. I use the maximum likelihood estimators

$$\hat{\mu}_{i,t} = \frac{1}{J} \sum_{s=j-J}^{j} r_{i,s} \tag{23}$$

$$\hat{\sigma}^2_{i,t} = \frac{1}{J} \sum_{s=j-J}^{j} (r_{i,s} - \hat{\mu}_{i,t})^2 \tag{24}$$

Here, j is the last trading day of the fiscal year prior to year t and J is the number of trading days in the previous two years. Figure 5 depicts the mean and median stock return variance over the sample period. The figure shows a substantial rise in volatility from 1993 to 2002, and the complete reversal thereafter.

Holdings of own-firm stocks and options. The model requires estimates of the fractions of CEOs' total wealth which are invested in own-firm stocks and options, η_s and η_z. More specifically, the forward-looking spirit of the model implies that η_s and η_z are the fractions invested in own-firm stocks and options the CEO *expects* to be required to hold over his future professional life. These quantities can only be approximated. I estimate η_s and η_z in the following way. Denote by w_z the fraction of options in current income, and assume the most common vesting schedule which provides that one third of the stock and option packages granted in a given year vests in each of the following three years. If the

CEO expects his income level and the fractions of income granted in the form of stocks and options to be constant, this means that in any given future year the CEO expects to hold $(1 + 2/3 + 1/3)w_z = 2w_z$ times the value of his current income in options. Using the wealth-to-income ratio, $m = M/Y$, this means that CEO expects to hold the fraction $\eta_z = 2w_z/m$ of his personal wealth in own-firm options. The analogous calculation applies to stock holdings.

Elasticity of the option portfolio with respect to the stock price. In the model, $\varepsilon_{z,s}$ measures the elasticity of the CEO's unvested option portfolio with respect to the stock price. This portfolio elasticity is the value-weighted average of the elasticities of the different unvested option grants. Consistent with a three year graded vesting schedule, the CEO expects to hold the entire current year option grant, two thirds of the prior year's grant, and one third of the grant made two years prior to the current year. To estimate the individual grants' elasticities I use the derivative of the Black-Scholes option value with respect to the stock price, multiplied by the ratio of the stock price to the option price:

$$\epsilon_{z,s} = e^{-dT}\Phi(D_1) \cdot \frac{S}{Z} = \frac{\Phi(D_1)}{\Phi(D_1) - X/S \cdot e^{d-r}\Phi(D_2)} \tag{25}$$

where S is the stock price, X is the strike price of the option grant, d is the dividend rate, $D_1 = [\ln(S/X) + (r - d + \sigma^2/2)T)]/\sqrt{\sigma^2 T}$, and $D_2 = D_1 - \sqrt{\sigma^2 T}$. The individual grant's elasticities depend on the price-to-strike ratio, S/X, the difference between the log dividend rate and the log risk free rate, $d - r$, the time to maturity, T, and volatility, σ^2. I assume that option grants have initial maturity of ten years, and are made at the money. The expected price-to-strike ratio of prior year grants is determined by the expected stock return, i.e. the prior year grant's price-to-strike ratio is $1 + \mu$.

CEO age and retirement age. The age of CEOs is available in Execucomp and can thus be inferred precisely. I use the 95th percentile of the CEO age distribution, 68 years, as the retirement age.

Table 3 summarizes the parameters of the model and reports their mean sample values. Table 4 reports sample statistics for the risks that the model captures.

[Table 3]

[Table 4]

4 The Cross-section of Risk and CEO Pay

In this section I analyze the relationship between the risk and level of CEO pay in the cross-section. I start by examining the correlation between risk and pay, and then proceed to evaluate the explanatory power of risk for the variance of CEO pay as well as for the higher moments of the pay distribution, that is, skew and kurtosis. For each of the moments I present the model-implied quantities for different degrees of risk aversion, and compare them with their empirical analogues. I further disentangle the contributions of the different types of risk to the various moments.

The analysis presented in this section rests on the proposition that CEO pay is determined in part by a compensating risk premium that makes CEOs indifferent between the risks inherent in the pay package they actually receive and an alternative safe income stream:

$$Y_i = \overline{Y}_i \times (1 + \pi(\sigma_i)) \tag{26}$$

where Y_i is CEO i's pay, $\overline{Y}_i$ is the safe income that CEO i could obtain, σ_i is the vector of the various risks that CEO i bears, and $\pi(\sigma_i)$ is the risk premium that compensates the CEO for the risks.

The risk premium, π, has somewhat complicated properties in that is it generally nonlinear in the individual risks $\sigma_{k,i}$, and depends on the interaction between the various risks as well as on other parameters such as CEO age and wealth. Thus, risk premia computed for a sample of CEOs with varying risks, age, and wealth will not lie on a single straight line when plotted against the risks. In order to facilitate comparison of model-implied and empirical risk premia, it is useful to log-linearize the risk premium by approximating equation (26) as

$$Y_i \approx e^{\alpha_i} \times e^{\alpha_1' \sigma_i} \tag{27}$$

where e^{α_i} corresponds to $\overline{Y}_i$, and the term $e^{\alpha_1' \sigma_i}$ approximates the gross risk premium $1 + \pi(\sigma_i)$. In the log-linear specification, α_1 is the vector of risk premia *per unit of risk*, and σ_i captures the amount of each type of risk the CEO bears.

While equations (26) and (27) are simple and intuitive, they omit other determinants of pay such as firm size. A simple extension of this model is one where pay scales with firm size and multiplicative risk premia for each of the risks:

$$Y_i \approx e^{\alpha_0} \times e^{\alpha_1' \sigma_i} \times Size_i^{\alpha_2} \tag{28}$$

One possible interpretation of this equation is that CEO pay is determined by two main factors: talent and risk. As in GL, CEOs are compensated for their marginal product of talent, which, in a competitive assignment model, ranks with firm size. In addition, they receive a multiplicative risk premium for each type of risk depending on the respective unit risk premium and amount of risk. The theoretical and empirical versions of equations (26), (27) and (28) form the basis for the following analysis.

4.1 A First Look at the Data

I start by presenting simple scatter plots of actual CEO pay against model-implied risk premia. This comparison is purely illustrative, since the relation between model-implied risk premia and actual pay is prone to endogeneity. That is, there is no reason to expect that risk premia vary one-to-one with actual pay, as the model predicts. Omitted variables that are correlated with both risk and pay may change the slope of the relationship between risk premia and actual pay.

The top graph of Figure 7 plots log CEO pay against the log of the total gross risk premium. This is an illustration of how well the simple model of equation (26) fits the data. CEO pay and risk premia are demeaned for each year separately so as to reflect cross-sectional variation only. The plot shows that, as predicted by the model, risk premia are positively associated with actual CEO pay, though the slope is smaller than one, and explanatory power is relatively low.

The bottom graph of Figure 7 plots the size-adjusted quantities of actual pay and risk premia. Size-adjusted pay is estimated as the residual of a panel regression of log CEO pay on log firm size and year dummies. Size-adjusted risk premia are computed using size-adjusted risks as model inputs, i.e. the residuals of a panel regression of the various risks on log firm size and year dummies. The graph illustrates how the residual variation in CEO pay is related to risk premia once size is accounted for. The figure shows that risk premia explain very well the size-adjusted variation in CEO pay. The slope of the fitted regression line is closer to one and explanatory power increases significantly. In the following subsections I turn to a more detailed statistical analysis of these relationships.

4.2 Correlation

In this section I compare model-implied and empirical correlations between risk and CEO pay. I distinguish between five types of risk: risks to wealth induced by option and stock holdings, risks to income due to dismissal and pay-performance sensitivity, and finally, stock return volatility which affects all of the above risks simultaneously.

Empirically, the vector of unit risk premia α_1 can be estimated by an ordinary least squares regression of log CEO pay on the vector of risks, i.e. by estimating the regression equation

$$\log(Y_i) = \alpha_0 + \alpha_1' \sigma_i + \alpha_2 \log(Size_i) + \varepsilon_i \tag{29}$$

where firm size is included as a control to avoid any bias due to the correlation between risks and firm size. To obtain cross-sectional relations for my dataset I estimate the above regression on the entire panel and include year fixed effects.

The model counterpart to the vector of empirical unit risk premia α_1, is the sample average of the derivative of $\log(1 + \pi(\sigma_i))$ with respect to σ_i. This derivative vector can be obtained, for each type of risk, as the coefficients of a regression of the log risk premium $\log(1 + \pi(\sigma_i))$ on the vector of risks σ_i:

$$\log(1 + \pi(\sigma_i)) = \alpha_0 + \alpha_1 \sigma_i + \varepsilon_i \tag{30}$$

To estimate the cross-sectional, model-implied unit risk premium on my dataset, I run regression (30) on the entire panel and include year fixed effects. Note that in order to estimate the log-linearized risk premium implied by the model, there is no need to account for firm size since the relationship if purely model-driven. Details of the estimation of empirical and model-implied risk premia is provided in Appendix B.1.

Empirical and model-implied unit risk premia are reported in Table 5. The empirical coefficient estimates are reported in column two, while their model-implied analogues are reported in columns three to six. Comparison of the model-implied with the empirical coefficients suggests that endogeneity is likely to be a concern. The theoretical estimates indicate that *exogenous* shifts in dismissal risk should command the highest risk premia followed by option holdings and volatility. In contrast, the empirical estimates suggest that risk premia for option and stock holdings are larger than for dismissal, and volatility premia are even negative.

One explanation for these differences is that the different risks are related to unobserved CEO skill. Stock and option holdings might be positively associated with CEO skill, perhaps because more skilled CEOs tend to take jobs that come with high-powered incentives. This would bias upwards the respective regression coefficients. On the other hand, volatility and dismissal might be negatively associated with CEO skill, thereby inducing a negative bias to the empirical correlations.

[Table 5]

4.3 Explained Variance

I now turn to the explanatory power of risk for CEO pay. Specifically, I examine what fraction of the cross-sectional variation in CEO pay can be explained by risk. As for the correlation, I compute both the empirical and the model-implied quantities of explained variance, for each risk individually as well as for all risks combined.

I compute three different measures of explained variance. First, I compute the fraction of total variance of CEO pay that can be explained by risk alone. The empirical version of this measure is the R-squared of a regression of log CEO pay on the various risks. Hence I call this measure R-squared. Second, I compute the partial R-squared, i.e. the fraction of total variance that can be explained by the size-orthogonal component of risk. Its empirical version is the partial R-squared of a regression of CEO pay on risk and firm size. Finally, I compute the fraction of the residual variance of CEO pay that is explained by risk once firm size is accounted for. This is the fraction of the variance of the size-orthogonal component of pay that is explained by the size-orthogonal component of risk. I call this statistic residual R-squared. The empirical R-squared statistics are computed from regressions, while their model analogues are computed from the model-implied risk premia (using raw risk as inputs for R-squared and size-adjusted risks for partial R-squared and residual R-squared). The details of the computations of the R-squared statistics are provided in Appendix B.

Panel A of Table 6 presents the results for R-squared. Each entry in the table represents the time-series average of cross-sectional R-squares statistics computed for each year in 1995-2005. The empirical R-squared statistics provide a benchmark for their model counterparts, though they might either over- or understate the true explanatory power of risk depending on the direction of the bias of the coefficients. Upward-bias (in absolute value) implies that the empirical R-squared will overstate the true R-squared, while bias towards zero will downward-bias empirical R-squared. Since the coefficients of options and stock holdings, which contribute most to the explanatory power of all risks, are likely to be upward biased, the empirical R-squared of all risks (shown in the bottom row of Panel A) can be regarded as an upper bound of the true R-squared implied by the causal effect of risk on pay.

Panel A of Table 6 shows that the model can explain a sizeable fraction of the total variance of CEO pay for plausible levels of risk aversion. For a risk aversion of $\rho = 3$ all risks combined explain about 20 percent of the cross-sectional variation in CEO pay. This compares to the 42 percent explained by the empirical model which, as explained above, likely overstates the true effect of risk.

Looking at the individual risks separately, it is apparent that equity-based pay and volatility contribute the most to the overall explanatory power of risk for CEO pay. This is because these risks exhibit both significant unit risk premia (as shown in Table 5) as well as large sample variation. In contrast, dismissal risk contributes only modest explanatory power to the cross-sectional variance of CEO pay. Despite the large unit risk premia commanded by turnover risk, its sample variation is simply too low to generate significant variation in risk premia. Pay-performance sensitivity appears to be essentially irrelevant for explaining the cross-section of CEO pay, because both unit risk premia are sample variation are modest.

Panel B contains partial R-squared statistics. The empirical statistics are substantially lower than the raw R-squared statistics of Panel A, indicating that omission of size leads to overstating the explanatory power of risk. While even after controlling for size some endogeneity remains, partial R-squared is a much more credible benchmark than raw R-squared. The partial R-squared statistics generated by the model are only slightly lower than the model-implied raw R-squared. And, for $\rho = 3$, the explanatory power of model quantities falls only slightly short of the empirical one (18 percent for the model versus 25 percent empirically). Panel C shows that the model can explain 30 percent of the residual variation in CEO pay (implying that size explains about 40 percent).

Taken together the results on explanatory power show that the model can explain a significant portion of the total variation in CEO pay and, for risk aversion of three, this variation is not much lower than the explained variance obtained from reduced form regressions once firm size is accounted for. Risk aversion between three and five is needed for the model to generate the same explanatory power as reduced form regressions.

4.4 Higher Moments

Another important question is whether risk premia can explain part of the asymmetry and fat-tailedness of the CEO pay distribution. It is well-known that CEO pay is heavily right-skewed with a small number of CEOs receiving extraordinary pay checks while the median CEO only makes a small fraction of that. The median CEO in my dataset earns around 2 million dollars per year while average pay is about 4 million, and CEOs in the top percentile of the pay distribution make more than 29 million. Several explanations for this phenomenon have been proposed in the literature, in particular the superstar status of some CEOs (Gabaix and Landier, 2008; Malmendier and Tate, 2009). In this section I explore whether the asymmetry, and in particular the high incidence of extreme CEO pay realizations can, in part, be explained by risk.[10]

[10]Maybe the most suggestive example of this possibility is the compensation Steve Jobs, the CEO of Apple Inc., received in 2000. In that year, Job's compensation was exclusively in stock options. The Black-Scholes value of his option package was about 600 million dollars, but the options were out of the money in the following year, i.e. had an intrinsic value of zero.

I restrict attention to the cross-sectional higher moments of actual CEO pay and gross risk premia. That is, I evaluate how well the simple model given in equation (26) fits the data. (Results on size-adjusted pay and risk premia are available from the author upon request.) As for the R-squared statistics, I compute skewness and kurtosis coefficients for each cross-section in 1995-2005, and report the observation-weighted average of the time-series of the cross-sectional statistics. Also as above, I compute the model statistics for each type of risk separately as well as for all risks combined, and for different degrees of risk aversion.

Table 7 presents the results. Panel A contains the skewness statistics, Panel B the results on kurtosis. The average cross-sectional skewness of CEO pay is about 3.7, documenting the pronounced inequality in CEO compensation resulting primarily from the right tail of the distribution. The skewness of model-implied risk premia is very similar in magnitude. The statistics range between 3.2 and 3.5 for all risks combined. Looking at individual risks separately, it is apparent that all risks generate positively skewed risk premia. This is due to the fact that, under CRRA utility, risk premia are convex in the size of the risk, so that even symmetrically distributed risks would translate into positively skewed risk premia (Arrow, 1971). Here again, it is equity-based pay and volatility that contribute most to the skewness of the total risk premium.

The kurtosis statistics, given in Panel B, show that model quantities exhibit very high, above-normal, values (between 14.3 and 16.5), though they are somewhat lower than in the data (22.7). In the next section I show that this deviation is entirely due to an extremely high kurtosis of actual pay in the year 2000 which the model does not capture. Notably, skewness and kurtosis coefficients of risk premia are unsensitive to the risk aversion parameter. For risk aversion between 2 and 10, skewness and kurtosis statistics are almost identical. This is due to the fact that risk aversion scales risk premia approximately linearly leaving the normalized higher moments unaffected.

Overall, the analysis of higher moments shows that risk premia capture very well the extreme observations of CEO pay. From a theoretical perspective this is due to the fact that risk premia are nonlinearly increasing in the size of the risks, so that large risks imply disproportionately large premia. Thus, some of the seemingly outrageous pay packages observed in recent years may be justified to the extent that they compensate for more risk. In this discussion it should be kept in mind that the assumptions made in the model do not allow for any kind of actions which, ex post, reduce or eliminate the effective risk that CEOs bear. Such actions include backdating or repricing of options, earnings management.

5 Risk and CEO Pay over Time

While the above section examines the average cross-sectional relations of risk and CEO pay, this section explores time-series patterns of moments of actual pay and model-implied risk premia over the period 1995 to 2005. Of particular interest is the question of whether the extreme characteristics of CEO pay during the height of the dot-com bubble in 2000-2001 is to some extent captured by the model. I start by comparing average levels of actual pay with model-impied risk premia and then proceed by analyzing variance and higher moments over time.

Figure 8 depicts yearly sample averages of actual CEO pay and gross risk premia. All moments are normalized to one in 1995. The top solid line plots average CEO pay and shows the steep increase from 1995 to 2000, and the non-monotonic pattern over the entire sample period. The lower lines represent the development of average risk premia. From 1995 to 2003, the pattern of risk premia is very similar to that of CEO pay. Risk premia also increase for a number of years, and then revert after the peak of the dot-com bubble. In terms of magnitude, the figure shows that the actual increase in pay between 1995 and 2000 is much steeper than could be justified by risk alone, even for risk aversion of ten. For a more plausible risk aversion of $\rho = 3$, risk premia would justify an increase in pay of about 30 percent. Following 2003, risk premia and actual pay diverge. While risk premia tend to decline through 2005, actual pay starts to rise again after 2003. A plausible explanation for this divergence is the effect of firm size (not shown in the figure). Firm size increases almost steadily through the entire sample period (monotonically if measured in total assets and with a slight dip in 2001 if measured in market capitalization). Rising firm size may then explain why, from 1995 to 2000, CEO pay increases more steeply than do risk premia and starts to rise again after 2003 when the decline of risk premia moderates.

[Figure 8]

The time pattern of cross-sectional variation is depicted in Figure 9. The figure plots the coefficient of variation instead of the variance in order to avoid the mechanical effect of the changing mean. The model appears to capture well the time-series pattern of the cross-sectional variation, rising until 2001, and then falling back to the level of 1995. Apparently, the cross-sectional variation in risk premia is sensitive to the level of risk aversion. For risk aversion of three, the coefficient of variation rises from 8 percent to 25 percent between 1995 and 2001. For risk aversion of ten it rises from 60 percent to 260 percent, far exceeding the variation in actual pay. Taken together, the time patterns of cross-sectional mean and variation in risk premia suggests that risk *alone* is unable to match both first and second moments of CEO pay. The high level of risk aversion needed to match the time-series pattern of mean pay would lead to a counterfactual time pattern of the cross-sectional variation.

[Figure 9]

Finally, Figures 10 and 11 depict the time-series of skewness and kurtosis. As noted above in Section 4.4, the higher sample moments of risk premia do not exhibit much sensitivity to risk aversion. Figures 10 and 11 illustrate that this is also true in the time-series. While the model fits very well the overall level of higher moments of the pay distribution it fails to capture the peak levels at the height of the dot-com bubble.

[Figures 10 and 11]

Taken together, the time-series pattern of cross-sectional moments suggests that risk helps to explain a number features of the aggregate development of CEO pay, in particular the rise and fall of average pay levels and cross-sectional variation over 1995-2005. However, the extreme magnitudes of average pay levels and higher moments around the peak of the dot-com bubble appear not be rationalizable by risk alone.

6 Conclusion

This paper employs a life-cycle model of consumption and saving to analyze the relationship between the risk and level of CEO compensation. Under the assumption that CEOs have the outside option of receiving a safe salary, the model generates quantitative predictions about the variation in CEO pay that can be attributed to risk. The model incorporates the various types of risk that executives of public corporations are exposed to, and thus allows for an assessment of the contribution of these individual risks to the total variation in pay.

A calibration of the model shows that, for typical levels of relative risk aversion, 20 percent of the cross-sectional variation and 30 percent of the observed increase in average pay levels from 1995 to 2000 can be explained by risk alone. Risk premia also help explain the pronounced inequality in CEO pay among the largest US corporations. In terms of the importance of the different types of risk, equity-based pay and stock return volatility contribute by far the most to the total variation in pay.

In interpreting these results it should be kept in mind that the model does not provide for the possibility of actions that, ex post, reduce the effective risks that CEOs bear. Such actions include backdating or repricing of options, earnings management, and excessive discretionary severance pay. In this sense, the risk premia calculated from the model can be viewed as fair compensation for risk from an *ex ante* perspective, that is, assuming compliance with the contractual agreements in place at the time of the grant.

This paper contributes to the debate about justified pay levels for corporate executives and has important implications for the design of compensation packages. First, the model provides a quantitative benchmark for variation in pay levels that keeps CEOs indifferent to receiving a safe income. By implication, any observed, additional variation in pay must affect CEO utility. The actual development of CEO pay levels over the past decades indicates that nowadays CEOs are, indeed, much better off than they used to be. Second, by quantifying the cost of the various types of risk, the model can be helpful for identifying the least costly mix of incentives.

A Numerical Solution of the Incentive Pay Life Cycle Model

The maximization problem given by equations (8)-(13) has no closed form solution and has to be solved numerically and recursively. It is well known that the problem can be written in the recursive form

$$V_t(M_t, Y_t) = \max_{\{C_s\}_{s=t}^{t+N}} E_t[\sum_{s=t}^{t+N} \beta^{s-t} u(C_s)] = \max_{C_t} \{u(C_t) + \beta E_t[V_{t+1}(M_{t+1}, Y_{t+1})]\} \tag{A-1}$$

subject to the same constraints.

A.1 Homogeneity of the Value Function

To simplify the problem we can exploit the fact that the value function is homogenous of degree $1-\rho$ in cash-on-hand and labor income. By normalizing all variables by Y_t, we can formulate the problem in terms of the single state variable $m_t = M_t/Y_t$, the wealth-to-income ratio. It is easy to show that the original value function, $V_t(M_t, Y_t)$ can then be written as a function of a newly defined value function, v_t, and scaled current income:

$$V_t(M_t, Y_t) = v_t(m_t) Y_t^{1-\rho} \tag{A-2}$$

where

$$v_t(m_t) = \max_{c_t} \{u(c_t) + \beta E_t[\Gamma_{t+1}^{1-\rho} v_{t+1}(m_{t+1})]\} \tag{A-3}$$

Thus one only needs to solve the simpler problem

$$\max_{c_t} \{u(c_t) + \beta E_t[\Gamma_{t+1}^{1-\rho} v_{t+1}(m_{t+1})]\} \tag{A-4}$$

subject to

$$a_t = m_t - c_t \tag{A-5}$$

$$m_{t+1} = a_t \frac{R_{t+1}}{\Gamma_{t+1}} + 1 \tag{A-6}$$

$$R_{t+1} = e^{\eta_f r_f + (\eta_t + \eta_z \varepsilon_{z,s}) r_{t+1}} \tag{A-7}$$

$$\Gamma_{t+1} = G_t \phi(r_{t+1}) \theta(r_{t+1}) \tag{A-8}$$

A.2 Computing the Consumption Function

The first order conditions for all periods prior to $T-1$ are given by the Euler equation

$$u'(c_t) = \beta E_t[\Gamma_{t+1}^{-\rho} u'(c_{t+1}) R_{t+1}]$$

where $c_{t+1} = c_{t+1}(m_{t+1})$ is the consumption function. Substituting for m_{t+1} and assuming CRRA utility yields

$$c_t^{-\rho} = \beta E_t[\Gamma_{t+1}^{-\rho}[c_{t+1}(a_s\frac{R_{t+1}}{\Gamma_{t+1}} + 1)]^{-\rho} R_{t+1}]$$

The solution to this equation involves the computation of the right hand side expectation. This expectation is a function of two random variables: the stock return, r_{t+1}, and the shock due to dismissal, θ_{t+1}, which is itself a function of the stock return. Using the law of iterated expectations one can split the expectation into two parts corresponding to the events of dismissal and no job change:

$$\begin{aligned} c_t^{-\rho} =& \beta E_t[\Gamma_{t+1}^{-\rho}[c_{t+1}(a_t\frac{R_{t+1}}{\Gamma_{t+1}} + 1)]^{-\rho} R_{t+1}|\theta_{t+1} = \lambda]P[\theta_{t+1} = \lambda] \\ &+\beta E_t[\Gamma_{t+1}^{-\rho}[c_{t+1}(a_t\frac{R_{t+1}}{\Gamma_{t+1}} + 1)]^{-\rho} R_{t+1}|\theta_{t+1} = 1]P[\theta_{t+1} = 1] \end{aligned}$$

Substituting for Γ_{t+1}, and writing the conditional expectation as the integral over the appropriate conditional density, yields

$$\begin{aligned} c_t^{-\rho} = \beta \int [G_t\phi(r_{t+1})\lambda]^{-\rho}\big[c_{t+1}(a_t\frac{R_{t+1}}{G_t\phi(r_{t+1})\lambda} + 1)\big]^{-\rho} R(r_{t+1}) f(r_{t+1}|\theta = \lambda) dr_{t+1} \cdot P[\theta = \lambda] \\ + \beta \int [G_t\phi(r_{t+1})]^{-\rho}\big[c_{t+1}(a_t\frac{R(r_{t+1})}{G_t\phi(r_{t+1})} + 1)\big]^{-\rho} R(r_{t+1}) f(r_{t+1}|\theta = 1) dr_{t+1} \cdot P[\theta = 1] \end{aligned} \tag{A-9}$$

The first order condition for the next-to-last working period $T-1$ is different. Due to the deterministic nature of the problem for all periods from T to N, the consumption function at period T is linear in wealth:

$$c_T = \frac{1 - \beta^{1/\rho}(1 + r_f)^{1/\rho - 1}}{1 - (\beta^{1/\rho}(1 + r_f)^{1/\rho-1})^{N-T}} \cdot m_T \tag{A-10}$$

The first order condition for T-1 is then given by

$$u'(c_{T-1}) = \beta E_{T-1}[\Gamma_T^{-\rho} u'(\frac{1 - \beta^{1/\rho}(1 + r_f)^{1/\rho - 1}}{1 - (\beta^{1/\rho}(1 + r_f)^{1/\rho-1})^{N-T}} \cdot m_T) R_T]$$

Again, substituting for m_T and assuming CRRA utility gives

$$c_{T-1}^{-\rho} = \beta E_{T-1}[\Gamma_T^{-\rho}\big[(\frac{1 - \beta^{1/\rho}(1 + r_f)^{1/\rho - 1}}{1 - (\beta^{1/\rho}(1 + r_f)^{1/\rho-1})^{N-T}} \cdot (a_{T-1}\frac{R_T}{\Gamma_T} + 1))\big]^{-\rho} R_T]$$

A.3 Computing the Value Function

The (normalized) value function also needs to be computed numerically and recursively. Substituting optimal consumption into the value function gives

$$v_t(m_t) = \frac{c_t^{*1-\rho}}{1-\rho} + \beta E_t\Big\{\Gamma_{t+1}^{1-\rho} v_{t+1}(m_{t+1})\Big\}$$

and replacing m_{t+1} and Γ_{t+1} yields

$$v_t(m_t) = \frac{c_t^{*1-\rho}}{1-\rho} + \beta E_t\Big\{[G_t\phi(r_{t+1})\theta_{t+1}]^{1-\rho} v_{t+1}(a_t \frac{R_{t+1}}{[G_t\phi(r_{t+1})\theta_{t+1}]} + 1)\Big\}$$

Again, we can split up the expectation into two parts and write the expectation as the integral over the conditional densities:

$$\begin{aligned} v_t(m_t) = {} & \frac{c_t^{*1-\rho}}{1-\rho} \\ & + \beta \int [G_t\phi(r_{t+1})\lambda]^{1-\rho} v_{t+1}(a_t \frac{R_{t+1}}{G_t\phi(r_{t+1})\lambda} + 1) f(r_{t+1}|\theta_{t+1} = \lambda) dr_{t+1} \cdot P[\theta_{t+1} = \lambda] \\ & + \beta \int [G_t\phi(r_{t+1})]^{1-\rho} v_{t+1}(a_t \frac{R_{t+1}}{G_t\phi(r_{t+1})} + 1) f(r_{t+1}|\theta_{t+1} = 1) dr_{t+1} \cdot P[\theta_{t+1} = 1] \end{aligned} \tag{A-11}$$

A.4 The conditional distribution of stock returns

The solution of the life-cycle model requires integration over the stock return distribution conditional on dismissal and no dismissal. Using Bayes' rule for probability densities, these conditional densities are given by

$$f(r_t|\theta_t = \lambda) = f(r_t) \cdot \frac{P[\theta_t = \lambda|r_t]}{P[\theta_t = \lambda]} \tag{A-12}$$

$$f(r_t|\theta_t = 1) = f(r_t) \cdot \frac{P[\theta_t = 1|r_t]}{P[\theta_t = 1]} \tag{A-13}$$

The empirical estimation of the conditional densities is detailed in Section 3.

B Comparing model-implied with empirical moments

B.1 Correlation

This paper explores the proposition that part of the variation in CEO pay can be described by a compensating risk premium that makes CEOs indifferent between the risks inherent in the pay package they actually receive and a pay package implying average risks. For the cross-section the model predicts that CEO pay depends on risk in the following way:

$$Y_i = \overline{Y}_i \times (1 + \pi(\sigma_i)) \tag{B-1}$$

where Y_i is CEO i's pay, $\overline{Y}_i$ is the safe income that CEO i could obtain (his safe outside option), σ_i is the vector of risks that CEO i faces, and π is the risk premium that compensates the CEO for these risk. The risk premium, π, is generally nonlinear in each of the risks $\sigma_{k,i}$, and, in addition, depends on other parameters such as CEO age and the wealth-to-income ratio. Thus, sample risk premia will not lie on a single line when plotted against the individual risks. In order to facilitate comparison of model-implied and empirical risk premia, it is useful to log-linearize the risk premium by writing equation (B-1) as

$$Y_i \approx e^{\alpha_0} \times e^{\alpha_1' \sigma_i} \tag{B-2}$$

where e^{α_0} corresponds to $\overline{Y}_i$, and the term $e^{\alpha_1'\sigma_i}$ approximates the gross risk premium $1+\pi(\sigma_i)$. In the log-linear specification, α_1 is the vector of risk premia *per unit of risk* and σ_i is the vector of the quantities of each risk that the CEO of firm i bears.

Empirically, the unit risk premium α_1 can be estimated by a simple ordinary least squares regression of log CEO pay on the risks, i.e. by estimating the regression equation

$$\log(Y_i) = \alpha_0 + \alpha_1'\sigma_i + \varepsilon_i \tag{B-3}$$

When estimating empirical risk premia via regression equation (B-3) one might be worried that firm size, the most important determinant of CEO pay, is omitted. Thus, in the empirical specification, I include log firm size as an explanatory variable:

$$\log(Y_i) = \alpha_0 + \alpha_1'\sigma_i + \alpha_2 \log(Size_i) + \varepsilon_i \tag{B-4}$$

This is in an intuitive extension of the simple model (B-2), in which pay scales with firm size according to a power law and a multiplicative risk premium:

$$Y_i \approx e^{\alpha_0} \times e^{\alpha_1'\sigma_i} \times Size_i^{\alpha_2} \tag{B-5}$$

One possible interpretation of this equation is that CEO pay is determined by two factors: talent and risk. As in GL, CEOs are compensated for their marginal product of talent, which, in a competitive assignment model, ranks with firm size. In addition, they receive a multiplicative risk premium depending on the unit risk premium and the amount of risk they bear.

The model-implied counterpart to the vector empirical unit risk premia α_1, is the sample average of the derivative of $\log(\overline{Y}_i \times (1+\pi(\sigma_i)))$ with respect to σ_i. This derivative can be obtained as the coefficient of a regression of the log risk premium $\log(1+\pi(\sigma_i)$ on the vector of risks. Note that in order to estimate the log-linearized risk premium implied by the model, there is no need to account for firm size since the relationship if purely model-driven.

B.2 Explained Variance

R-squared. I denote by R-squared the fraction of the variance of log CEO pay that is explained by risk alone - either by one particular type of risk or alternatively by all risks combined. The empirical R-squared is simply the explained sum of squares from regression (29), denoted by ESS_1 divided by the total sum of squares, denoted by TSS:

$$R^2 = \frac{ESS_1}{TSS} \tag{B-6}$$

The model-implied counterpart to the empirical R-squared is given by the fraction of the model-implied explained sum of squares and the total sum of squares. The model-implied explained sum of squares for risk k is given by

$$ESS_1 = (N-1) \cdot Var(\log(1+\pi(\sigma_i))) \tag{B-7}$$

where N is the number of observations in the cross-section.

Partial R-squared. I denote by partial R-squared the fraction of variance of CEO pay that is explained by the component of risk that is orthogonal to firm size. For the empirical partial R-squared this can be done by

1. regressing log pay on log firm size and taking the residual $\varepsilon_{1,i}$:

$$\log(Y_i) = \alpha_0 + \alpha_1 \log(Size_i) + \varepsilon_{1,i} \tag{B-8}$$

   I denote the residual sum of squares of this regression by RSS.

2. regressing each of the risks, σ_k, on log firm size and taking the residual $\varepsilon_{2,i}$:

$$\sigma_{k,i} = \alpha_0 + \alpha_1 \log(Size_i) + \varepsilon_{2,i} \tag{B-9}$$

3. regressing $\varepsilon_{1,i}$ on $\varepsilon_{2,i}$:

$$\varepsilon_{1,i} = \alpha_0 + \alpha_1 \varepsilon_{2,i} \tag{B-10}$$

4. computing the explained sum of squares of regression (B-10), denote it by ESS_2.

5. computing partial R-squared as:

$$R^2_{partial} = \frac{ESS_2}{TSS} \tag{B-11}$$

To obtain the model-implied counterpart to the empirical partial R-squared I obtain the residuals from regression (B-9) and use these size-adjusted risks as inputs for the model. The model-implied explained sum of squares is then given by

$$ESS_2 = (N-1) \cdot Var(\log(1 + \pi(\sigma_i^{adj.}))) \tag{B-12}$$

and the partial R-squared is obtained as the fraction of the model-implied ESS_2 and the total sum of squares, TSS.

Residual R-squared. Finally, I compute the fraction of the residual variance (after controlling for firm size) that is explained by the component of risk orthogonal to size. Residual R-squared is given by

$$R^2_{residual} = \frac{ESS_2}{RSS} \tag{B-13}$$

B.3 Higher moments

For space reasons I report skew and kurtosis statistics only for total risk, not for the individual risks separately. Also, I simply compare model-implied higher moments with the actual higher moments of CEO pay, not the higher moments generated by the empirical model. I compute the skew and kurtosis as

$$Skew = \frac{\mu^3}{\sigma^3} \tag{B-14}$$

$$Kurtosis = \frac{\mu^4}{\sigma^4} \tag{B-15}$$

where μ^3 is the third central moment and σ is the standard deviation of the variable of interest. The empirical statistics are simply the skew and kurtosis of total pay. The model-implied statistics are computed from the gross risk premium on all risks, $1 + \pi(\sigma_i)$.

Table 1: Descriptive statistics. This table reports descriptive statistics for CEOs and firms in the Execucomp database over the period 1993-2005. Panel A contains compensation data. Total compensation is the sum of salary, bonus, the Black-Scholes value of options granted, the market value of restricted stocks granted, and other compensation. Panel B reports turnover statistics. Panel C contains firm characteristics. Market capitalization is the market value of equity plus the book value of debt. Volatility is computed as the sample variance of daily stock returns over a given year. Total compensation, mean and volatility of stock returns are winsorized at the top and bottom 0.5 percent. All quantities are inflation-adjusted to constant 2000 dollars using the Bureau of Economic Analysis GDP deflator.

Panel A: CEO compensation

	Mean	Median	Std	Observations
Total compensation [$1000]	4011.50	2078.02	6036.03	21160
Salary [$1000]	622.35	566.58	333.44	21160
Bonus [$1000]	689.28	358.20	1263.50	21160
Black-Scholes value of option grant [$1000]	1910.17	552.18	4547.29	21160
Value of restricted stock grant [$1000]	393.56	0.00	1551.40	21160
Other compensation [$1000]	212.47	37.67	990.08	21160
Options-based pay [%]	0.33	0.29	0.29	21160
Stock-based pay [%]	0.06	0.00	0.14	21160

Panel B: CEO turnover

	# turnovers	# involuntary	# voluntary
Turnover statistics	2426	639	1855

Panel C: Firm characteristics

	Mean	Median	Std	Observations
Market capitalization [$m]	12082.95	1949.32	50334.41	18425
Total assets [$m]	10578.76	1380.03	48393.26	21352
Market-to-book	2.09	1.50	2.51	18425
Mean stock return	0.19	0.12	0.59	21103
Volatility	0.40	0.36	0.18	20875

Panel D: CEO characteristics

	Mean	Median	Std	Observations
Age	55.63	56.00	7.60	21146
Tenure	7.80	5.00	7.29	19731

Table 2: Performance regressions. This table reports pay-performance (Panel A) and turnover-performance regressions (Panel B). Row one to eleven of each panel contain estimates for two-year rolling window regressions. The last row of each panel reports estimates for the entire sample period 1995-2005. The regression equation for the pay-performance regressions is $\Delta \log(Y_{i,t}) = \alpha_0 + \alpha_1 r_{i,t} + \alpha_2 \Phi(\sigma^2_{i,t}) + \alpha_3 r_{i,t} \times \Phi(\sigma^2_{i,t}) + \nu_j + r_{i,t} \times \nu_j + \varepsilon_{i,t}$, where Y is CEO pay. The probit model for the turnover-performance regressions is $\text{Prob}(D_{i,t} = 1|\cdot) = \Phi(\alpha_0 + \alpha_1 r_{i,t} + \alpha_2 \Phi(\sigma^2_{i,t}) + \alpha_3 r_{i,t} \times \Phi(\sigma^2_{i,t}) + \nu_j + r_{i,t} \times \nu_j)$, where $D_{i,t}$ is the dismissal indicator, $r_{i,t}$ is the stock return, $\Phi(\sigma^2_{i,t})$ is the sample cumulative distribution function of stock return volatility, and ν_j is a vector of industry dummies. The industry dummies follow the Fama-French (1997) classification into 12 industries. The coefficients of the industry dummies and their interactions with the stock return are omitted to save space. Standard errors are clustered at the firm level and reported in brackets. *** significant at the 1% level, ** significant at 5% level, * significant at 10% level.

Panel A: Pay-performance regressions

	Return		$\Phi(\sigma^2)$		Return $\times\Phi(\sigma^2)$		Constant		Observations
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	
1995	0.23	[0.17]	-0.18***	[0.07]	-0.19	[0.17]	0.11**	[0.05]	2168
1996	-0.01	[0.17]	-0.13**	[0.06]	0.03	[0.18]	0.17***	[0.05]	2554
1997	-0.07	[0.19]	0.00	[0.06]	0.02	[0.16]	0.14**	[0.06]	2550
1998	0.23	[0.14]	0.02	[0.07]	-0.13	[0.15]	0.04	[0.06]	2488
1999	0.13	[0.13]	-0.02	[0.07]	0.18	[0.14]	0.06	[0.06]	2543
2000	0.04	[0.16]	-0.17**	[0.07]	-0.07	[0.17]	0.10*	[0.06]	2583
2001	0.18	[0.19]	-0.19**	[0.08]	-0.29	[0.18]	0.15**	[0.07]	2571
2002	0.15	[0.19]	-0.16**	[0.08]	-0.23	[0.19]	0.16**	[0.07]	2628
2003	-0.05	[0.18]	-0.22***	[0.06]	0.03	[0.19]	0.11*	[0.06]	2672
2004	0.39***	[0.13]	-0.07	[0.06]	-0.23	[0.14]	0.07	[0.06]	2710
2005	0.40***	[0.12]	0.19***	[0.06]	-0.38***	[0.12]	0.02	[0.06]	2719
1995-2005	0.13**	[0.05]	-0.11***	[0.03]	-0.10*	[0.05]	0.12***	[0.03]	14298

Panel B: Turnover-performance regressions

	Return		$\Phi(\sigma^2)$		Return $\times\Phi(\sigma^2)$		Constant		Observations
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	
1995	-2.05***	[0.56]	0.57**	[0.28]	1.41*	[0.85]	-2.57***	[0.27]	1975
1996	-1.01*	[0.59]	0.89***	[0.26]	0.67	[0.76]	-2.39***	[0.23]	2389
1997	-0.39	[0.45]	0.58**	[0.26]	0.13	[0.60]	-2.00***	[0.20]	2510
1998	-1.66***	[0.56]	0.20	[0.24]	0.80	[0.61]	-1.58***	[0.19]	2545
1999	-1.99***	[0.49]	0.32	[0.22]	0.77	[0.52]	-1.82***	[0.20]	2590
2000	-1.78***	[0.67]	0.48*	[0.25]	0.93	[0.65]	-2.21***	[0.26]	2647
2001	-3.06***	[0.85]	0.49*	[0.27]	1.85***	[0.65]	-2.74***	[0.35]	2528
2002	-2.25**	[0.94]	0.25	[0.31]	1.18	[0.93]	-2.48***	[0.32]	2673
2003	-2.82**	[1.18]	0.47	[0.30]	0.37	[1.00]	-2.79***	[0.42]	2720
2004	-0.85	[0.81]	0.41	[0.27]	0.61	[0.54]	-2.11***	[0.24]	2657
2005	-1.46***	[0.40]	-0.09	[0.26]	1.81***	[0.49]	-1.64***	[0.21]	2593
1995-2005	-1.52***	[0.28]	0.47***	[0.10]	1.10***	[0.28]	-2.06***	[0.17]	15804

Table 3: Parameter description and baseline values. This table summarizes the parameters of the model as well as their average sample values. See Section 3 for details on the calibration of the parameters.

Parameter	Description	Average sample values
m	Wealth-to-income ratio	4
γ_0	Expected income growth rate	0.03
γ_1	Pay-performance sensitivity	0.11
T	Years until retirement	12
α_0	Intercept of probit model	-1.81
α_1	Turnover-performance sensitivity	-0.45
η_s	Percentage of wealth invested in company stock	3.09%
η_z	Percentage of wealth invested in company options	16.32%
μ	Mean stock return	12.00%
σ^2	Variance of stock return	19.55%
r_f	Risk free rate	6.00%
ε_{zs}	Elasticity of the option value with respect to the stock price	1.5
β	Discount factor	0.98
$1-\lambda$	Percentage drop in income following dismissal	0.5
ρ	Coefficient of relative risk aversion	varying

Table 4: Sample distribution of risks to CEO income and wealth. This table reports sample statistics of the various risks to CEO income and wealth.

Risk	Model parameter / quantity	Mean	Std	Q25	Q50	Q75
Option holdings	η_z	0.163	0.145	0.000	0.147	0.273
Stock holdings	η_s	0.031	0.071	0.000	0.000	0.000
Dismissal	$\Phi(\alpha_0 + \alpha_1 r) f(r) dr$	0.035	0.017	0.022	0.032	0.047
Pay-performance	γ_1	0.123	0.089	0.049	0.120	0.184
Volatility	σ^2	0.196	0.168	0.078	0.140	0.259

Table 5: Cross-sectional correlations. This table presents empirical and model-implied cross-sectional correlations between the various risks and log CEO pay. The rows report statistics for the different types of risk. The second column reports the coefficients and standard errors of a panel regression of log CEO pay on the various risks, log firm size and year dummies. The regression equation is $\log(Y_{i,t}) = \alpha + \beta\sigma_{i,t} + \log(Size_{i,t}) + \psi_t + \varepsilon_{i,t}$, where Y is total CEO pay, σ is the vector of risks, $Size$ is firm size measured as market capitalization, ψ_t is a vector of year dummies, and $\varepsilon_{i,t}$ is an error term. Columns three to six contain the model-analogues to the empirical coefficients. They are estimated via panel regressions with year fixed effects of the log gross risk premium, $\log(1 + \pi(\sigma_i)$ on the risks, σ_i. See Appendix B.1 for details on the computation of empirical and model-implied correlations. Standard errors are clustered at the firm level and reported in brackets. *** significant at the 1% level, ** significant at 5% level, * significant at 10% level.

	Empirical	Model			
Risk		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
Option holdings	3.46***	1.70***	2.13***	3.52***	9.09***
	[0.08]	[0.02]	[0.02]	[0.02]	[0.06]
Stock holdings	3.69***	1.16***	1.48***	2.51***	7.52***
	[0.13]	[0.02]	[0.02]	[0.03]	[0.10]
Dismissal	2.33**	4.49***	5.49***	7.62***	11.14***
	[1.03]	[0.04]	[0.04]	[0.06]	[0.10]
Pay-performance	0.75***	0.42***	0.64***	1.11***	2.09***
	[0.17]	[0.00]	[0.01]	[0.01]	[0.02]
Volatility	-0.20**	1.21***	1.64***	2.84***	7.02***
	[0.09]	[0.01]	[0.01]	[0.03]	[0.09]

Table 6: Cross-sectional R-squared. This table presents empirical and model-implied cross-sectional R-squared statistics. Panel A reports R-squared, computed as the fraction of the total variance of log CEO pay explained by the various risks. Panel B reports partial R-squared, computed as the fraction of the total variance of log CEO pay explained by the size-orthogonal components of the risks. Panel C reports residual R-squared, computed as the fraction of the size-orthogonal component of log CEO pay explained by the size-orthogonal component of the risks. The top five rows in each panel report R-squared statistics for each type of risk separately, the bottom row contains R-squared for all risks combined. The second column in each panel contains empirical R-squared statistics obtained from panel regressions with year fixed effects of log CEO pay on the respective risks (Panel A), or the size-orthogonal components of these variables (Panel B and C). Columns three to six contain the model-analogues to the empirical R-squared statistics, computed as the observation-weighted time-series average of the cross-sectional R-squared statistics. The theoretical R-squared uses model-implied risk premia to compute explained variance, computed from total (Panel A) or the size-orthogonal components of the risks (Panel B and C). See Appendix B for details on the computation of the various R-squared statistics.

Panel A: R-squared

	Empirical	Model			
Risk		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
Option holdings	0.24	0.06	0.10	0.27	1.11
Stock holdings	0.06	0.06	0.08	0.18	0.66
Dismissal	0.03	0.00	0.01	0.01	0.03
Pay-performance	0.01	0.00	0.00	0.00	0.02
Volatility	0.03	0.03	0.06	0.18	0.83
All risks	**0.42**	**0.15**	**0.19**	**0.33**	**0.96**

Panel B: Partial R-squared

	Empirical	Model			
Risk		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
Option holdings	0.19	0.06	0.09	0.25	1.03
Stock holdings	0.02	0.01	0.01	0.03	0.20
Dismissal	0.01	0.00	0.01	0.01	0.02
Pay-performance	0.01	0.00	0.00	0.00	0.02
Volatility	0.01	0.03	0.05	0.14	0.62
All risks	**0.25**	**0.14**	**0.18**	**0.31**	**0.94**

Panel C: Residual R-squared

	Empirical	Model			
Risk		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
Option holdings	0.31	0.10	0.16	0.43	1.75
Stock holdings	0.03	0.01	0.02	0.05	0.35
Dismissal	0.02	0.01	0.01	0.02	0.04
Pay-performance	0.02	0.00	0.00	0.01	0.03
Volatility	0.02	0.04	0.08	0.23	1.06
All risks	**0.40**	**0.24**	**0.30**	**0.53**	**1.60**

Table 7: Cross-sectional skewness and kurtosis. This table presents empirical and model-implied cross-sectional skewness and kurtosis statistics. The coefficient of skewness, reported in Panel A, is calculated as the ratio of the third centered moment and the standard deviation raised to the third power. The kurtosis, reported in Panel B, is calculated as the ratio of the fourth centered moment and the standard deviation raised to the fourth power. The top part of each panel contains statistics for gross risk premia generated by the individual risks (rows one to five) and all risks combined (row six). The columns contain the respective statistics by degree of relative risk aversion, ρ. The bottom part of each panel reports the skewness and kurtosis of total CEO pay. All statistics are computed as the observation-weighted mean of the cross-sectional statistics for each year in 1995-2005. See Appendix B for details on the computation of the skewness and kurtosis statistics.

Panel A: Skewness

	Model			
Risk	$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
Option holdings	1.84	1.88	2.27	2.99
Stock holdings	3.96	3.87	4.13	5.12
Dismissal	0.40	0.44	0.34	0.10
Pay-performance	0.59	0.59	0.57	0.58
Volatility	1.61	2.06	2.83	3.61
All risks	**3.41**	**3.29**	**3.22**	**3.49**
	Data			
Skewness of CEO pay:	**3.66**			

Panel B: Kurtosis

	Model			
Risk	$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
Option holdings	7.21	7.00	8.70	12.88
Stock holdings	22.21	21.47	24.12	33.69
Dismissal	2.97	2.99	2.83	2.40
Pay-performance	3.53	3.53	3.50	3.32
Volatility	5.83	8.29	13.05	18.15
All risks	**16.49**	**15.29**	**14.32**	**16.07**
	Data			
Kurtosis of CEO pay:	**22.70**			

Table 8: Mean and variation in CEO pay and risk premia over time. This table reports the sample mean (Panel A) and coefficient of variation (Panel B) of CEO pay and gross risk premia for each year in 1995-2005. The yearly means are expressed as a fraction of the 1995 mean. The second column in each panel contains the respective cross-sectional statistics for total CEO pay. Columns three to six contains sample statistics for gross risk premia calculated for different degrees of risk aversion.

Panel A: Mean (1995=1)

	Data	Model			
Year		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
1995	1.00	1.00	1.00	1.00	1.00
1996	1.27	1.04	1.04	1.03	1.05
1997	1.54	1.07	1.08	1.10	1.15
1998	1.64	1.15	1.19	1.29	1.39
1999	1.99	1.14	1.19	1.33	1.54
2000	2.36	1.23	1.28	1.46	1.81
2001	2.22	1.25	1.30	1.49	1.84
2002	1.83	1.26	1.31	1.49	1.80
2003	1.68	1.13	1.18	1.33	1.52
2004	1.85	1.14	1.16	1.23	1.41
2005	1.89	1.12	1.16	1.24	1.38

Panel B: Coefficient of variation

	Data	Model			
Year		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
1995	0.82	0.05	0.08	0.22	0.58
1996	0.92	0.06	0.09	0.23	0.62
1997	1.00	0.07	0.12	0.32	0.86
1998	1.04	0.10	0.17	0.47	1.40
1999	1.26	0.11	0.18	0.51	1.66
2000	1.34	0.15	0.25	0.76	2.20
2001	1.41	0.19	0.31	0.86	2.66
2002	1.14	0.19	0.30	0.84	2.37
2003	1.07	0.12	0.20	0.57	1.87
2004	0.98	0.09	0.13	0.34	1.18
2005	1.01	0.08	0.14	0.38	1.18

(Table 8 continued)

Panel C: Skewness

	Data	Model			
Year		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
1995	3.53	3.66	3.85	3.70	3.64
1996	3.50	3.37	3.43	3.89	4.27
1997	3.65	3.44	3.60	3.56	3.66
1998	3.90	3.38	3.38	3.29	3.06
1999	4.09	3.30	2.92	3.10	3.31
2000	6.22	3.18	3.18	3.09	3.48
2001	4.85	3.00	2.79	2.65	3.71
2002	2.69	3.25	2.91	2.76	3.72
2003	2.63	3.31	3.28	3.10	3.47
2004	2.65	3.85	3.25	2.93	3.04
2005	2.86	3.72	3.56	3.27	3.06

Panel D: Kurtosis

	Data	Model			
Year		$\rho = 2$	$\rho = 3$	$\rho = 5$	$\rho = 10$
1995	19.59	18.79	20.25	18.53	17.24
1996	21.26	16.60	16.36	19.79	23.57
1997	22.34	16.46	17.19	16.55	17.47
1998	25.17	15.92	15.76	14.66	12.29
1999	26.06	16.03	12.05	13.48	14.45
2000	65.56	13.82	13.89	13.23	15.90
2001	34.52	12.91	11.34	10.02	18.00
2002	12.72	15.27	12.26	10.70	17.70
2003	11.87	14.65	14.96	13.19	15.68
2004	12.08	20.38	15.04	12.24	12.62
2005	13.74	19.98	18.69	14.89	12.25

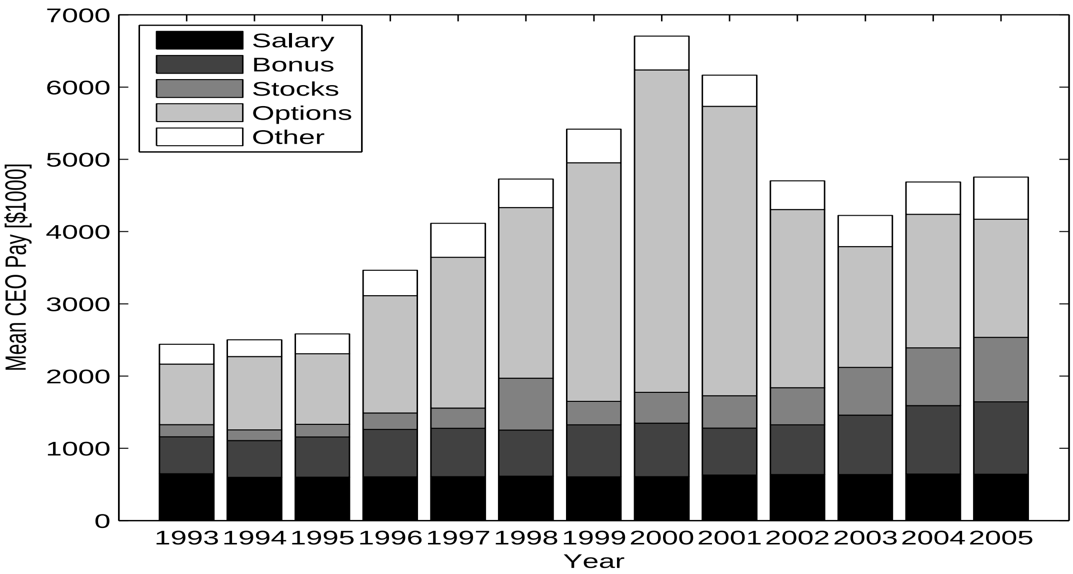


Figure 1: Average CEO pay. This figure depicts average CEO pay of Execucomp firms (indicated by bar height), and its components over 1993-2005. Total pay includes salary, bonus, restricted stock, option grants (valued using the Black-Scholes formula), and other compensation. Total CEO pay is truncated at the 0.5 percent level and inflation-adjusted to constant 2000 dollars using the Bureau of Economic Analysis GDP deflator.

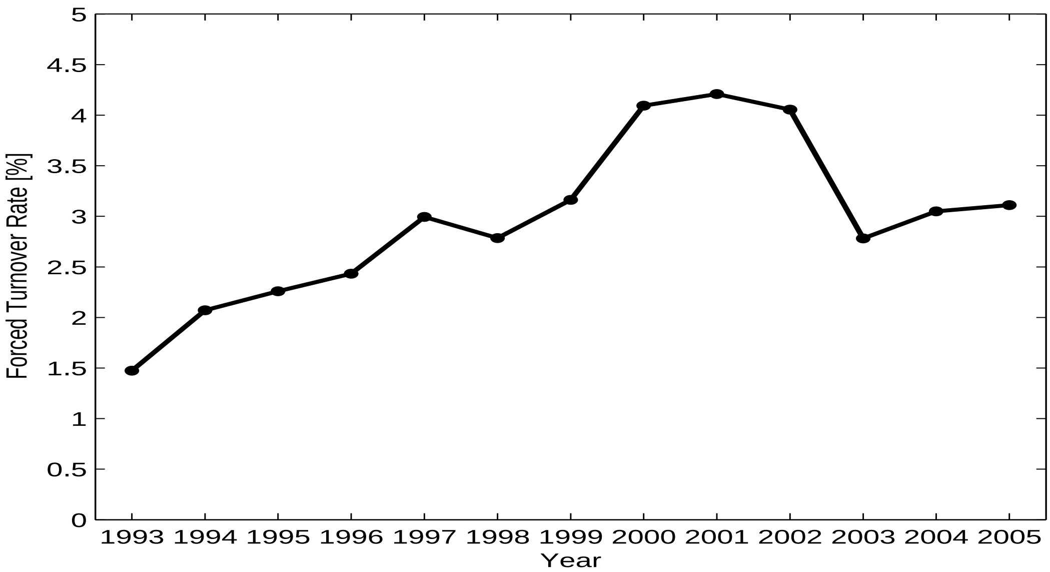


Figure 2: Dismissal rates. This figure depicts forced CEO turnover rates over 1993-2005 for Execucomp firms. Underlying data is from Jenter and Kanaan (2008) for 1993-2000 and Peters and Wagner (2009) for 2001-2005.

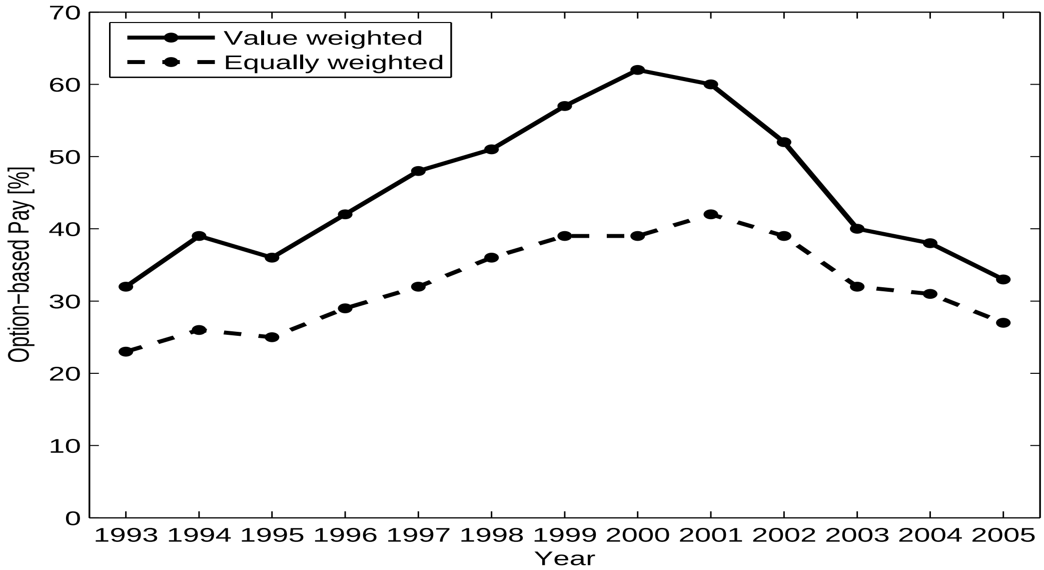


Figure 3: Fraction of option-based pay. This figure shows value- and equally weighted fractions of option-based pay over 1993-2005 for the CEOs of Execucomp firms.

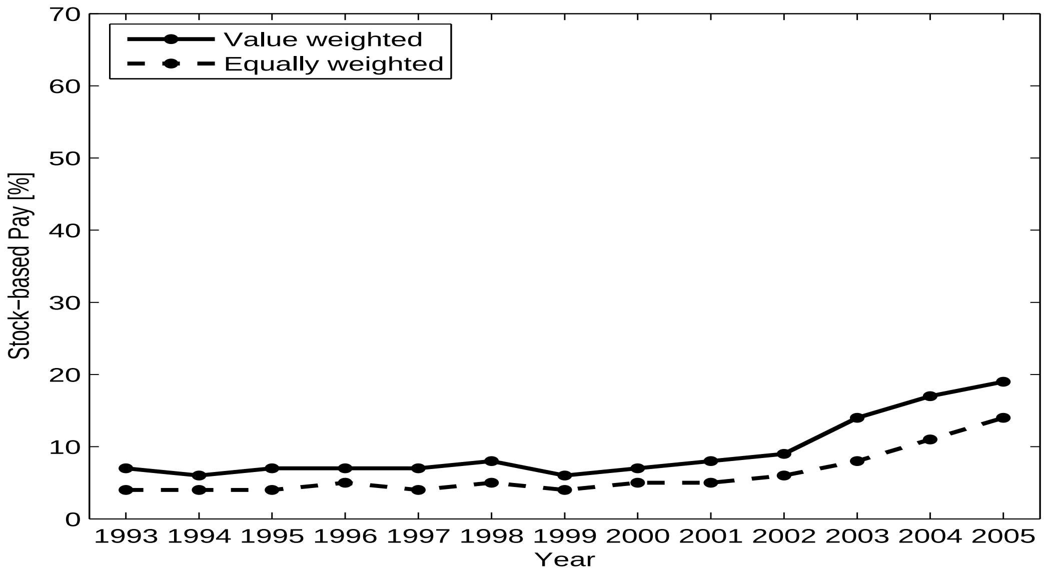


Figure 4: Fraction of stock-based pay. This figure shows value- and equally weighted fractions of stock-based pay over 1993-2005 for the CEOs of Execucomp firms.

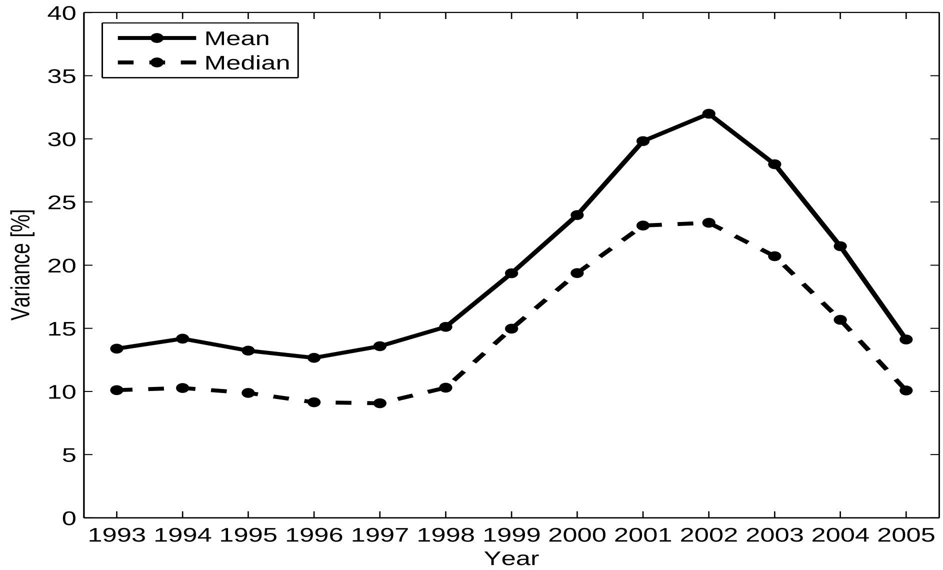


Figure 5: Stock return volatility over time. This figure depicts the yearly mean and median stock return variances over 1993-2005 for Execucomp firms. Variances are computed using daily stock returns over the two years preceding the respective fiscal year. The yearly means are computed as equally weighted averages.

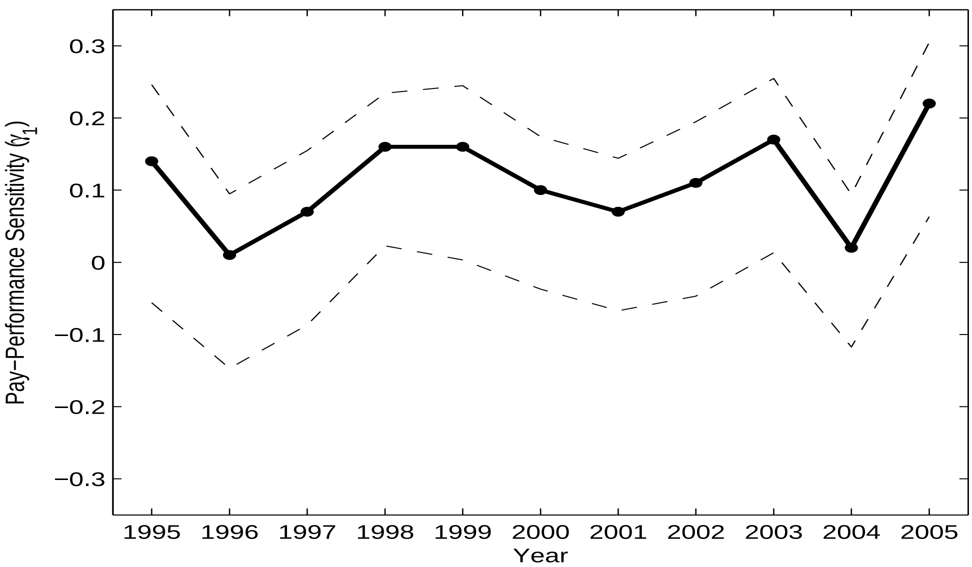
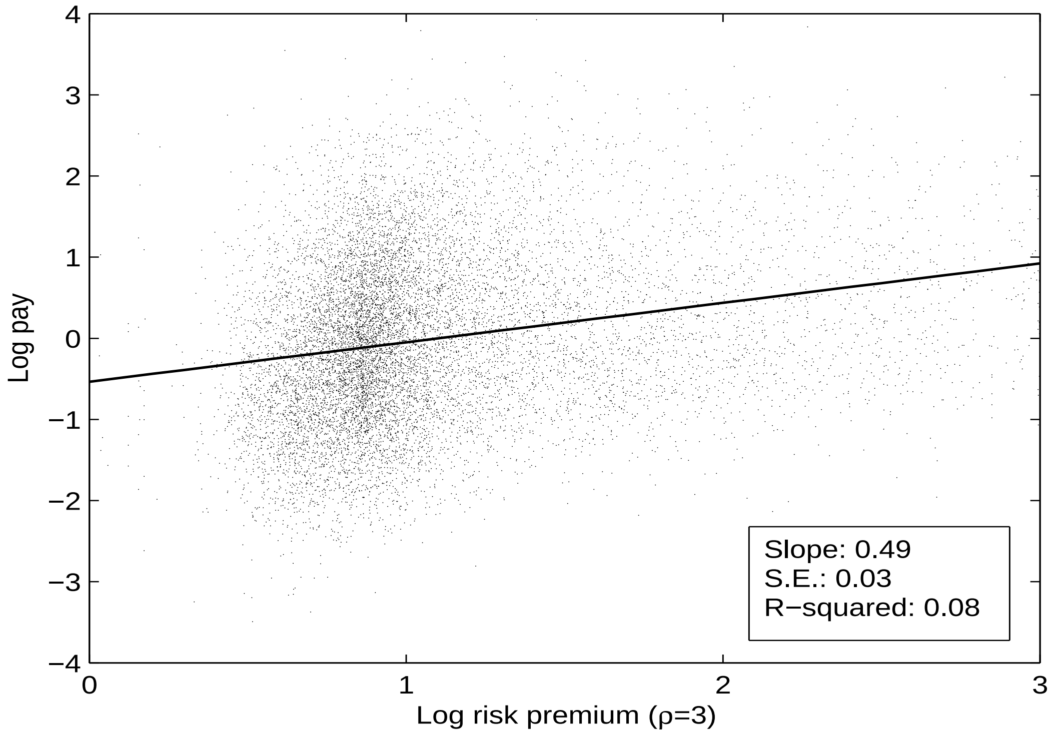


Figure 6: Pay-performance sensitivities over time. This figure depicts pay-performance sensitivities over time. The regression equation is $\Delta \log(Y_{i,t}) = \gamma_0 + \gamma_1 r_{i,t} + \gamma_2 \Phi(\sigma^2_{i,t}) + \gamma_3 \Phi(\sigma^2_{i,t}) \times r_{i,t} + \nu_j + \nu_j \times r_{i,t} + \varepsilon_{i,t}$. ϑ_t is a vector of year dummies, ν_j is a vector of industry dummies corresponding to 12 Fama-French industry groups. The regression is run over two-year rolling windows, and the figure plots the cross-sectional median pay-performance sensitivity, $\hat{\gamma}_1 + \hat{\gamma}_3 \times 0.5$, for each year. The two dashed lines represent 95% confidence bounds.



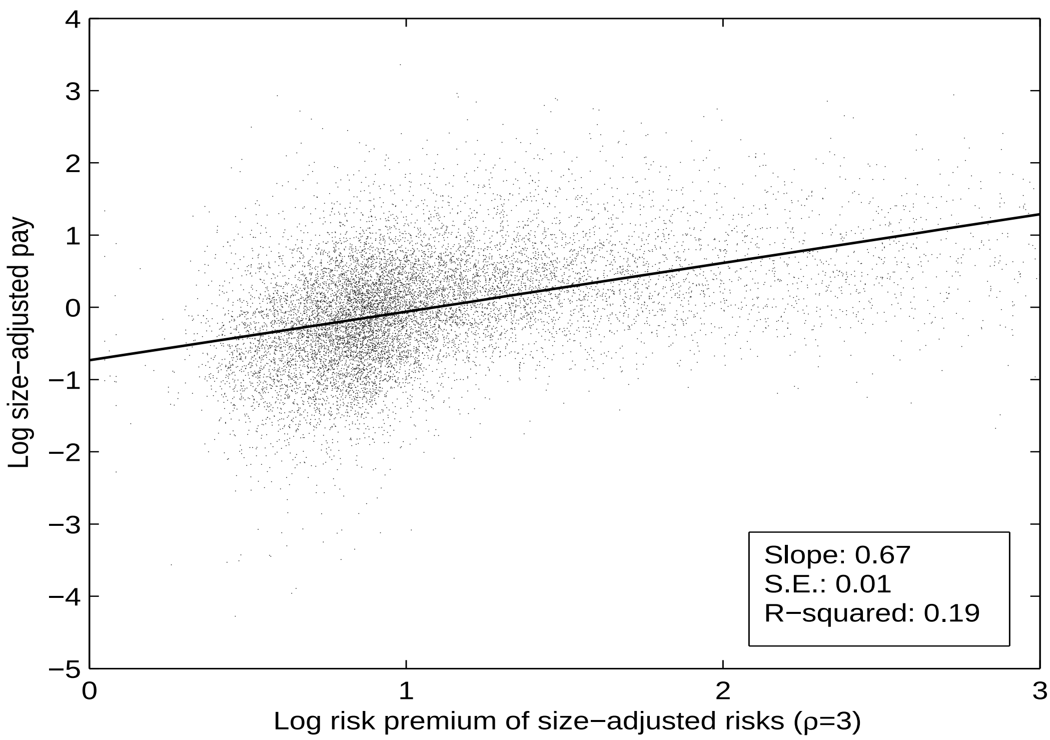


Figure 7: CEO pay and risk premia in the cross-section. The top figure shows a scatter plot of log CEO pay, against the log gross risk premium, $\log(1 + \pi(\sigma_i))$ for risk aversion of three. Both variables are adjusted by their yearly means to reflect cross-sectional variation only. For the bottom figure both CEO pay and risks are, in addition, adjusted for firm size. The solid lines represent fitted linear regressions of the data points.

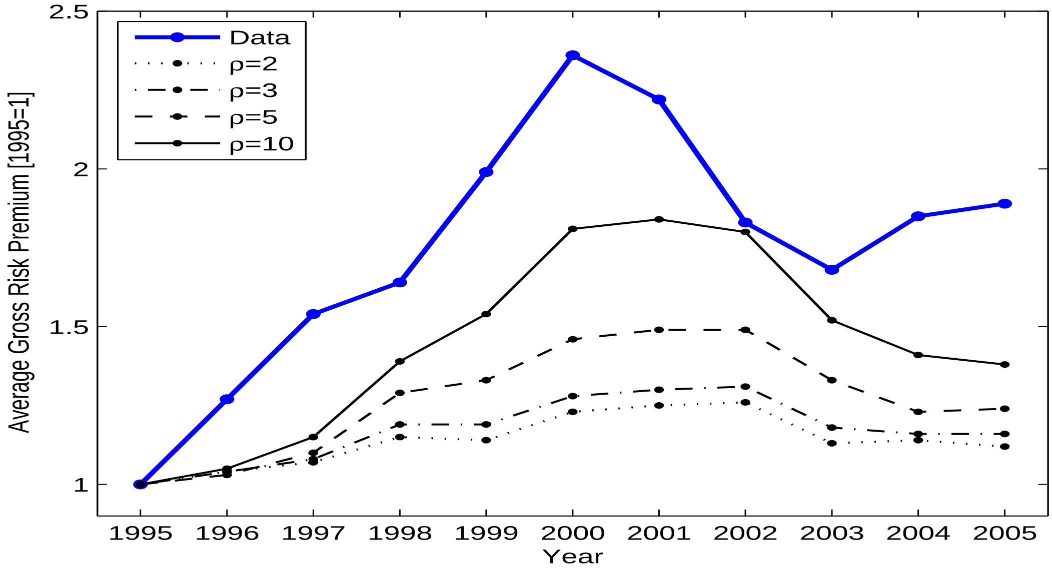


Figure 8: Average CEO pay and risk premia for different degrees of risk aversion. This figure depicts average CEO pay (thick, solid line) and cross-sectional risk premia over 1995-2005 for different degrees of relative risk aversion, ρ.

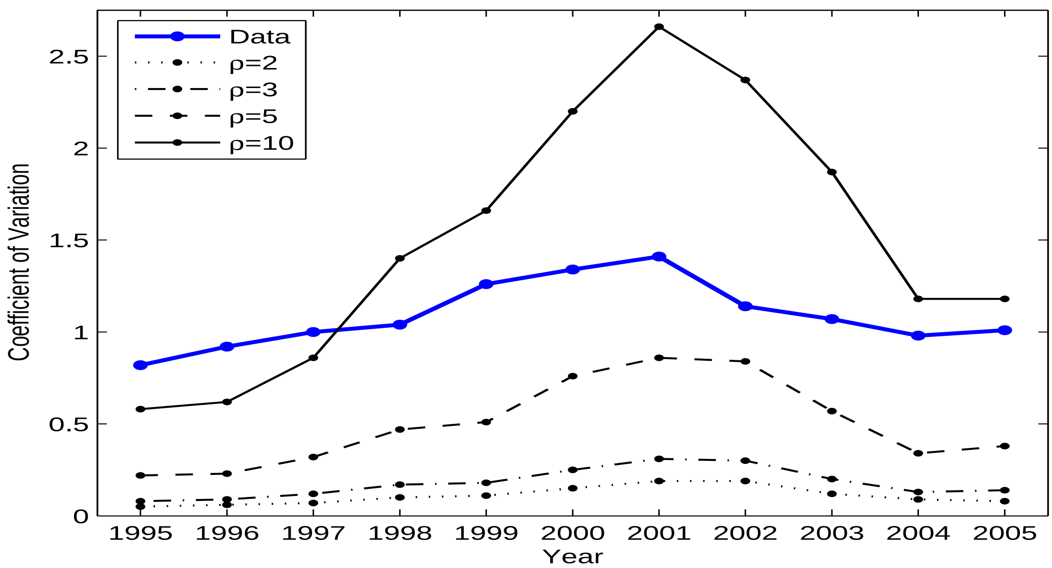


Figure 9: Cross-sectional variation in risk premia and CEO pay. This figure depicts the coefficient of variation of CEO pay (thick, solid line) and gross risk premia for different degrees of relative risk aversion, ρ, over 1995-2005.

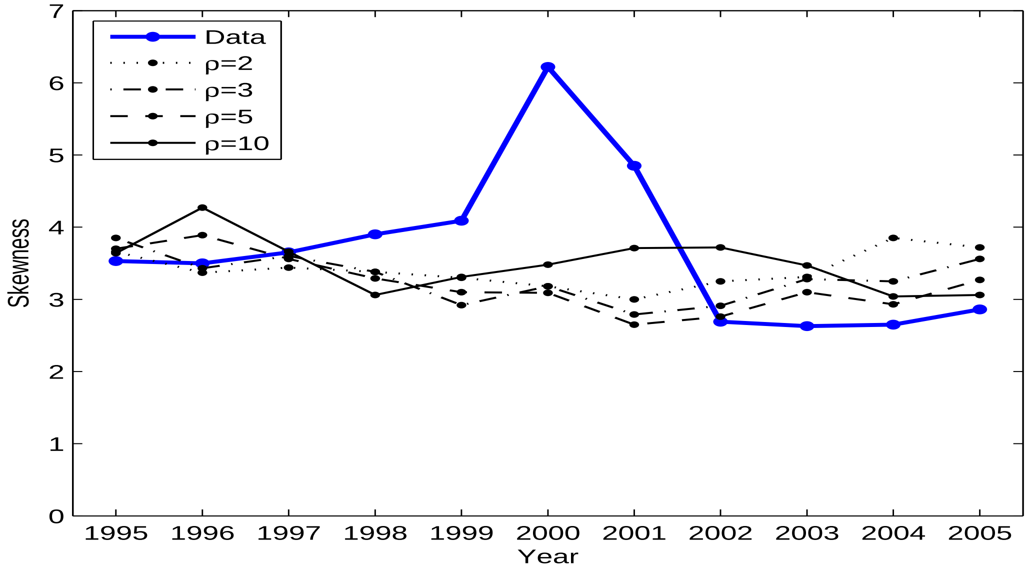


Figure 10: Skewness of CEO pay and risk premia for different degrees of risk aversion. This figure depicts skewness of CEO pay (thick, solid line) and gross risk premia for different degrees of relative risk aversion, ρ, over 1995-2005.

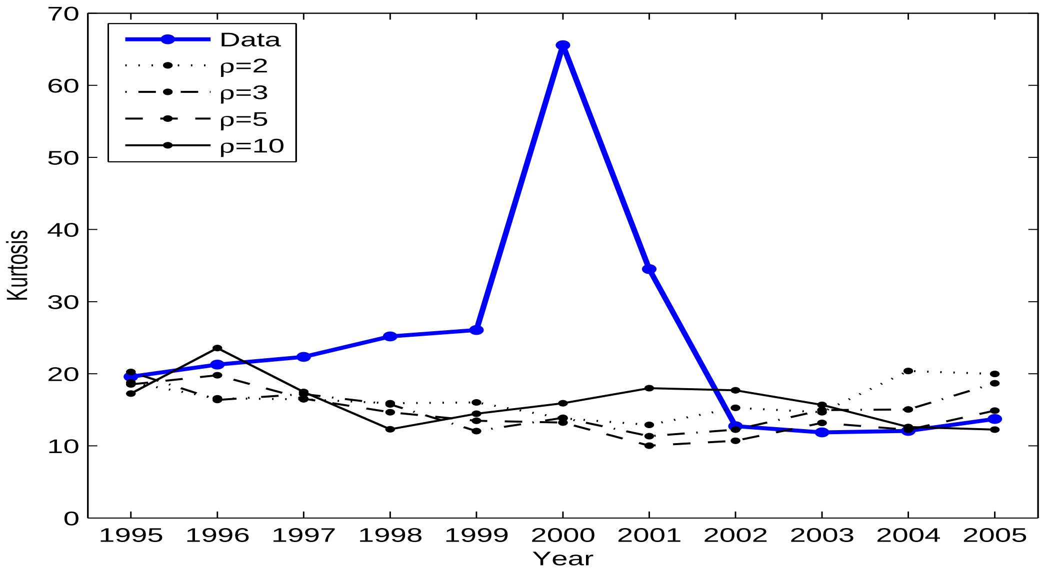


Figure 11: Kurtosis of CEO pay and risk premia for different degrees of risk aversion. This figure depicts kurtosis of CEO pay (thick, solid line) and gross risk premia for different degrees of relative risk aversion, ρ, over 1995-2005.

References

Aggarwal, R. K. and A. A. Samwick (1999): "The Other Side of the Trade-off: The Impact of Risk on Executive Compensation," *Journal of Political Economy*, 107, 65–105.

Armstrong, C. S., D. F. Larcker, and C.-L. Su (2007): "Stock Options and Chief Executive Officer Compensation," *Working Paper, Stanford University.*

Arrow, K. (1971): *Essays in the Theory of Risk-Bearing*, Markham Publishing Company.

Bebchuk, L. and J. Fried (2004): *Pay without performance*, Cambridge: Harvard University Press.

Carpenter, J. N. (1998): "The exercise and valuation of executive stock options," *Journal of Financial Economics*, 127–158.

Carroll, C. D. (1992): "The Buffer Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 2, 61–135.

Carroll, C. D. and A. A. Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40, 41–71.

Chapman, D. A., R. B. Evans, and Z. Xu (2009): "Career Concerns and the Active Manager's Problem," *Working Paper, Boston College.*

Deaton, A. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59, 1121–1142.

Dittmann, I. and E. Maug (2007): "Lower Salaries and No Options? On the Optimal Structure of Executive Pay," *Journal of Finance*, 62, 303–343.

Dittmann, I., E. Maug, and O. Spalt (2009): "Sticks or Carrots? Optimal CEO Compensation when Managers are Loss Averse," *Journal of Finance, forthcoming.*

Dittmann, I. and K.-C. Yu (2009): "How Important are Risk-Taking Incentives in Executive Compensation?" *Working paper, Erasmus University.*

Edmans, A., X. Gabaix, and A. Landier (2008): "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *forthcoming, Review of Financial Studies.*

Evans, G. W. and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics*, Princeton: Princeton University Press.

Fee, E. C. and C. J. Hadlock (2004): "Management Turnover across the Corporate Hierarchy," *Journal of Accounting and Economics*, 37, 3–38.

Frydman, C. (2005): "Rising through the Ranks: Th Evolution of the Market for Corporate Executives," *MIT Working Paper.*

Gabaix, X. and A. Landier (2008): "Why has CEO pay increased so much?" *Quarterly Journal of Economics*, 123, 49–100.

Gayle, G.-L. and R. A. Miller (2009): "Has Moral Hazard Become a More Important Factor in Managerial Compensation?" *The American Economic Review, forthcoming.*

Gibbons, R. and K. J. Murphy (1992): "Optimal Incentive Contracts in the Presence of Career Concers," *Journal of Political Economy*, 100, 468–505.

Gourinchas, P.-O. and J. A. Parker (2002): "Consumption over the Life Cycle," *Econometrica*, 70, 47–89.

Hall, B. and J. Liebman (1998): "Are CEOs Really Paid Like Bureaucrats?" *Quarterly Journal of Economics*, 113, 653–691.

Hall, B. J. and K. J. Murphy (2002): "Stock Options for Undiversified Executives," *Journal of Accounting and Economics*, 33, 3–42.

Haubrich, J. G. (1994): "Risk Aversion, Performance Pay, and the Principal-Agent Problem," *Journal of Political Economy*, 102, 258–276.

Hermalin, B. E. (2005): "Trends in Corporate Governance," *The Journal of Finance*, 60, 2351–2384.

Holmström, B. and P. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55, 303–328.

Hubbard, G., J. S. Skinner, and S. P. Zeldes (1994): "The Importance of Precautionary Motives for Explaining Individual and Aggregate Saving," in *Carnegie-Rochester Conference Series on Public Policy*, ed. by A. H. Meltzer and C. I. Plosser, North Holland, Amsterdam, vol. 40, 59–125.

Huddart, S. (1994): "Employee Stock Options," *Journal of Accounting and Economics*, 18, 207–231.

Ingersoll, J. E. (2006): "The Subjective and Objective Evaluation of Incentive Stock Options," *Journal of Business*, 79.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993): "Earnings Losses of Displaced Workers," *American Economic Review*, 83, 685–709.

Jensen, M. C., K. J. Murphy, and E. G. Wruck (2004): "Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them," *Working paper, European Corporate Governance Institute.*

Jenter, D. and F. Kanaan (2008): "CEO turnover and relative performance evaluation," *Working Paper, MIT.*

Kaplan, S. N. and J. Rauh (2009): "Wall Street and Main Street: What Contributes to the Rise in the Highest Incomes?" *Review of Financial Studies, forthcoming.*

Kuhnen, C. M. and J. Zwiebel (2008): "Executive pay, hidden compensation and managerial entrenchment," *Working paper.*

Laibson, D., A. Repetto, and J. Tobacman (2007): "Estimating Discount Functions with Consumption Choices over the Lifecycle," *Working Paper, Harvard University.*

Lambert, R. A., D. F. Larcker, and R. E. Verrecchia (1991): "Portfolio Considerations in Valuing Executive Compensation," *Journal of Accounting Research*, 29, 129–149.

Malmendier, U. and G. Tate (2009): "Superstar CEOs," *Quarterly Journal of Economics, forthcoming.*

Murphy, K. J. and J. Zabojnik (2004): "CEO pay and appointments: A market-based explanation for recent trends," *American Economic Review*, 94, 192–196.

Parrino, R. (1997): "CEO turnover and outside succession: A cross-sectional analysis," *Journal of Financial Economics*, 46, 165–197.

Peters, F. S. and A. F. Wagner (2009): "The Executive Turnover Risk Premium," *Working Paper, Swiss Finance Institute.*

Rosen, S. (1981): "The Economics of Superstars," *American Economic Review*, 71, 845–858.

Terviö, M. (2009): "The Difference that CEOs Make: An Assignment Model Approach," *American Economic Review*, 98, 642–668.

Zeldes, S. P. (1989): "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *Quarterly Journal of Economics*, 104, 275–289.