Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 15 (2008) 145 – 166
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Assessing the role of option grants to CEOs: How important is heterogeneity?

Nina Baranchuk [a,*], Siddhartha Chib [b]

[a] School of Management, University of Texas — Dallas, P.O. BOX 830688 SM31, Richardson, TX 75083-0688, United States
[b] John M. Olin School of Business, Washington University, 1 Brookings Dr., Campus Box 1133, St. Louis, MO 63130, United States

Accepted 23 October 2006
Available online 23 May 2007

Abstract

This paper revisits the question of whether CEO compensation practices are in keeping with those justified by agency theory. We develop and analyze a new panel Tobit model, estimated by modern Bayesian methods, in which the heterogeneity of covariate effects across firms is modeled in a hierarchical way. We find that our specification of heterogeneity provides a significantly improved fit to the data. Our results show support for the hypothesis that companies increase option awards to their CEOs when agency problems become more pronounced. We also find that liquidity constraints matter in defining the cash–option mix of CEO compensation.

JEL classification: G34; M52
Keywords: Executive compensation; Incentive pay; Bayesian estimation

1. Introduction

The topic of CEO compensation has generated a large literature in economics and finance. One important question is the extent to which option grants, which are part of most CEO compensation packages, serve to align the interests of CEOs to those of the shareholders, thus providing the means to mitigate the principal-agent problem. In one strand of the literature, the derivative of the Black–Scholes value of option grants with respect to the price of the underlying stock (delta on option grants) is used to measure the incentives given to the CEO through options. In this paper, we study the use of option grants by analyzing the link between this response, referred to as the pay-per-performance sensitivity (PPS), and the covariates that might indicate the extent of the principal-agent problem, variables such as CEO share ownership, dollar return volatility, Tobin's Q, relative noise in accounting measures of performance, and variables related to tenure and career concerns of the CEOs.

One complication that has been largely ignored in the literature (and which is the focus of this paper) is the large cross-sectional variation in CEO compensation packages. This has the potential to disguise the link between the

* Corresponding author. Tel.: +1 972 883 4771.
E-mail addresses: nina.baranchuk@utdallas.edu (N. Baranchuk), chib@olin.wustl.edu (S. Chib).

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2006.10.004

covariates mentioned above and the PPS of option grants. Although a number of papers in the current corporate governance literature recognize the importance of controlling for heterogeneity in the relationship between agency covariates and the PPS of option grants, the statistical analysis of the implications has been rather limited (see Core et al., 2003 for a recent survey). Yermack (1995) and Anderson and Bizjak (2003), for example, utilize panel data methods in conjunction with a Tobit censored regression model but limit the heterogeneity to just the intercept of their model.

An important exception is Hermalin and Wallace (2001), who utilize a model that allows two covariates in addition to the constant term to have heterogeneous coefficients. They concentrate on the relationship between the size of CEO compensation and firm performance characteristics and limit their analysis to a sample of 104 firms from one industry (thrift). Chen and Jiang (2003) indirectly also discuss this complication but not in the context of models with coefficient heterogeneity. They mainly focus on the dollar return volatility, and test whether sales growth and firm size are related to the PPS of CEOs total compensation.

On the whole, the results in the existing literature that deal with the relationship between PPS of option grants and agency covariates have been mixed. We conjecture that the differences in the existing results are due to the heterogeneity present in the data. The estimated relationship between firm characteristics and incentive pay may be significantly affected if omitted variables influence the relationship between incentive pay and firm characteristics (for a detailed discussion of the effect that failing to account for heterogeneity may have on parameter estimates in the context of pay-for-performance sensitivity of executive compensation see Himmelberg et al., 1999; Hermalin and Wallace, 2001). Additionally, in some firms, the contracting process may be affected by agency problems which in turn would lead to heterogenous compensation practices.

In this paper we adopt more flexible heterogeneity models (than those based on intercept heterogeneity) to understand these data. Specifically, with the help of new hierarchical Bayesian Tobit panel data model and a data set that covers 1476 firms for the years 1992–2002, we show that a richer form of heterogeneity is prevalent in the data. In other words, we find support for firm-specific variability in covariate effects (beyond that captured by variation in the intercept). We also test by formal Bayesian marginal likelihood/Bayes factor methods whether our heterogeneity assumptions, built into the hierarchical Tobit panel model, are supported by the data by comparing the performance of our model with a number of alternative models that embed less heterogeneity. We find that our proposed model produces the smallest out-of-sample prediction error, and is more likely to correctly predict the censored outcome.

Moreover, we also find that accounting for heterogeneity has an important effect on significance and signs of the estimated model parameters. Our parameter estimates are generally consistent with agency theory prescriptions. Specifically, we find that CEOs with higher share ownership tend to receive smaller option grants. It is also interesting to note that the cross-firm variation in the estimates of the coefficient on CEO share ownership is rather large and robust to various model specifications. Therefore, while CEO share ownership may not be a significant factor in some firms, it plays an important role in other firms. We also find that firms with less noise in return on equity tend to rely less on option grants, which is likely due to the relatively larger amount of information such firms can obtain from accounting measures of CEOs performance. We find that the employment cycle matters: new CEOs receive a bigger portion of their compensation in terms of option awards than CEOs continuing their tenure, while departing CEOs receive significantly fewer option awards. We find support for the theoretically-prescribed negative relationship between pay-for-performance sensitivity of option grants and dollar volatility of the underlying stock returns. Our results are robust to various reasonable prior specifications.

The variables that we use to control for financial flexibility issues have significant coefficients with the expected signs. The indicator variable for dividend payout, which measures the firm's need for cash, has the expected positive and significant coefficient, indicating that firms that are more in need of cash tend to reduce cash-based compensation in favor of option grants. We also find limited support for a negative relationship between interest coverage and use of option grants in CEO compensation.

We also examine if there is support for the view that option grants are used to compensate for temporary deviations from the optimal level of incentives. We check this by fitting a model in which the agency related variables appear as year-to-year differences rather than in levels. We find that the coefficients on almost all the differenced variables are insignificant with the exception of the coefficients on the changes in CEO share and option ownership (which are positive and significant). Thus we find limited support if any for the view that option grants are a mechanism to correct for deviations from the optimal level of incentives.

The rest of the paper is organized as follows. In Section 2, we describe the source of our data and present motivating evidence of heterogeneity in coefficient effects. The general model we fit is discussed in Section 3. Section 4 contains a brief outline of the estimation method which is based on Markov Chain Monte Carlo (MCMC) methods and the framework developed in Chib (1992) and Chib and Carlin (1999). The results of the fitting are presented in Section 5. Section 6 analyzes the reasons for the divergence between our results and those in the existing empirical literature. Section 6 concludes. An Appendix contains further details on the econometrics in the paper.

2. Data

The data analyzed in this paper is obtained from the following sources: ExecuComp was the source of information on executive compensation, the data on stock prices was obtained from the Center of Research in Security Prices (CRSP), and the data on financial characteristics of the firms was obtained from the CRSP/Compustat merged database. Finally we also use the governance index developed in Gompers, Ishii, and Metrick (2003). After removing outliers (as will be specified later) and the firms that had less then 4 (not necessarily consecutive) years of observations, we obtained a data set that contains 1530 firms (10859 firm-year observations).

It has been argued that the Black–Scholes formula does not correctly represent the value of options written on the firm's stock to the insiders (e.g. Core and Guay, 2002). Nonetheless, the Black–Scholes valuation formula is the predominant method of choice in the empirical literature on executive compensation and we remain faithful to that tradition.

We focus on the following response variable:

$$\frac{\partial(\text{Black} - \text{Scholes value of option award})}{\partial P}\left(\frac{\text{shares in option award}}{\text{shares outstanding}}\right), \tag{1}$$

where P is the stock price. This variable reflects the increase in pay-for-performance sensitivity of the CEO wealth that can be attributed to the current year option award.

We now turn to describing some features of the data that motivates our empirical modeling strategy. First, there is considerable censoring of the response variables with about 28% of the responses censored at zero. This feature of the data suggests the use of the Tobit family of models. Second, as can be seen from the histogram in Fig. 1(on the left), the response variable appears to be highly skewed. This indicates that the assumption of normality is not likely to be reasonable for these data. To address this issue, we consider two transformations of this variable: a square root transformation and a natural log transformation $\ln(y+1)$. The histograms for the variable and its square root transformation are shown in Fig. 1. The histogram for the data with the log transformation is similar. As can be seen from the Fig. 1, the transformed variable is closer to being symmetric, and has much thinner tails. Since both

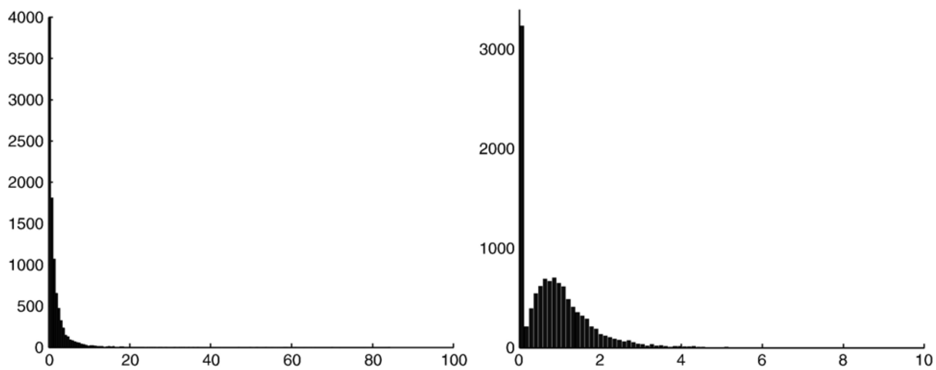

Fig. 1. The histogram on the left shows pay-for-performance sensitivity of last year option grants to CEOs. The histogram on the right shows square root of the same variable.

Table 1
List of variables

Variable name	Year	Calculation method
Pay-for-performance sensitivity	Current	Option delta multiplied by % shares underlying the award (see Eq. (1))
Dollar volatility of returns	Current	Volatility of percentage returns based on 60 months multiplied by the market value of the shares outstanding
CEO stock ownership	Current	% of outstanding stock currently owned by the CEO
Tobin's Q	Current	Book value of long-term debt+market value of stock divided by book value of total assets
Financial leverage	Current	Book value of debt divided by book value of total assets
Relative noise in ROE	Current	Volatility of ROE divided by stock return volatility
Zero dividends	Current	1 if paid no dividends
Tax loss c.f.	Current	1 if the firm has tax loss carry-forward
Interest coverage	Current	Earnings before interest and taxes divided by interest spending (absolute value capped at 50)
ln(total assets) last year	Current lagged	Percentage return on common stock earned
Stock return	1 year	In the previous year
GI	Current	N/A
Age [a]	Current	CEO's age
Tenure [a]	Current	Number of years the CEO spent with the firm

The total sample consists of 1530 firms, 10859 firm-year observations for years 1992–2002. The source for the variables is Compustat (unless mentioned otherwise).

[a] 4656 firm-year observations (773 firms).

transformations produce similar results, only the results with the square root transformation are reported in the paper. From now on, we will use the following notation:

$$y = \left(\frac{\partial\ (\text{Black} - \text{Scholes value of option award})}{\partial P}\left(\frac{\text{shares in option award}}{\text{shares outstanding}}\right)\right)^{0.5} \tag{2}$$

Our explanatory covariates (summarized in Table 1 can be divided into two main groups: the variables included to capture some aspects of the firm's agency relationship and control variables. The first group contains the following seven variables: 'CEO share ownership' — CEO beneficial ownership of the firm's shares; 'Tobin's Q' — Tobin's Q is approximated by the ratio of the firm's market value to the book value of its assets; 'relative 'noise' in ROE' — ratio of the variance in return on equity to the annual volatility of the stock returns; 'financial leverage' — ratio of the book value of debt to the book value of assets; 'vol($ returns)' — volatility of dollar returns on the firm's shares; 'new CEO dummy' — dummy variable, equal to 1 when a new CEO accepts the position; 'departing CEO dummy' — dummy variable, equal to 1 when a the current CEO leaves the position. Our proxies for agency relationships are imperfect in that some of the above variables reflect firm characteristics unrelated to agency problems. We discuss this issue in more detail in Section 5, when we discuss our estimation results.

Our control variables include three variables that proxy for financial flexibility: 'zero-dividend dummy' — dummy variable, equal to 1 when the company paid no dividends during the year; 'tax loss c.-f. dummy' — dummy variable that is equal to 1 when there is a tax loss carry-forward; and 'interest coverage' — ratio of operating income before interest and depreciation to the interest charges. In addition, we include 'GI' — governance index, constructed in Gompers et al. (2003),[1] 'ln(total assets)' — natural log of total assets, to control for firm size; 'last year stock return' — return on buy-and-hold strategy for the year preceding the option grant year; time dummies and industry dummies. Each industry dummy represents one of the 30 industry groups that were formed based on two-digit SIC codes. For the list of industries, see Table 6.

Table 2 shows a correlation matrix for some of the variables (we omitted the variables that have correlation less then 0.2 in absolute value with all other variables). Most of the covariates have correlation close to zero, except for high correlation values of −0.58 between financial leverage and interest coverage, and −0.42 between firm size and zero

[1] We acknowledge that it would be desirable to include additional variables such as ownership and the number of large outside shareholders and characteristics of the boards of directors. Unfortunately, the data on such variables is not available for most firms in our data set.

Table 2
Variance–covariance matrix

	var ($ returns)	var (% returns)	CEO stock ownership	Financial leverage	Zero dividends dummy	Interest coverage
var($ returns) CEO stock	1.00					
Ownership financial	−0.12	0.13	1.00			
Leverage zero dividends	−0.03	−0.07	−0.10	1.00		
Dummy interest	−0.15	0.55	0.12	−0.01	1.00	
Coverage	0.02	0.05	0.16	−0.58	0.04	1.00
ln(total assets)	0.62	−0.44	−0.22	0.13	−0.42	−0.09
Variance inflation factors	1.41	1.82	1.10	1.50	1.32	2.08

The sample consists of 1530 firms, 10859 firm-year observations for years 1992–2002.

dividends dummy. All Variance Inflation Factors (also reported in Table 2), however, are smaller than 3, which indicates that collinearity is not a problem.

To explore the panel data structure of the covariates, we calculate the average within-firm variance for each variable, as well as cross-sectional variance, and report the results in Table 3. Although for most variables, cross-sectional variance is larger than within-firm variance, the size of the within-firm variance is still noticeably larger than zero for all variables, suggesting that the time series data is also capable of providing information about firm compensation policies.

One important empirical feature that we seek to model (that has not been as carefully analyzed in the literature) relates to the nature of the heterogeneity in the coefficient values across firms. Simple intuitive reasoning suggests that the large variability in compensation observed in the data will be difficult to explain without allowing for considerable heterogeneity in the coefficient values.

To illustrate the kind of heterogeneity that needs to be modeled we estimate the effect of a change in CEO share ownership, firm by firm, on the response variable. We run the following set of regressions:

$$y_{it} = \alpha_i + \beta_i(\text{CEO share ownership})_{it} + \varepsilon_{it}, i = 1, 2, \ldots, n; \tag{3}$$

where the response variable y is given by Eq. (2) and n is the number of firms in the sample. The error terms are assumed to be i.i.d. normal. As can be seen from Fig. 2, showing the estimates of β_i for each firm i, there is considerable variability in the distribution of coefficient estimates across the firms in our sample. The histogram seems to indicate that the set of firms can be divided into two groups, with one group being more sensitive to changes in CEO share ownership than the other.

In addition to CEO share ownership, arguably a number of other variables may be expected to have a heterogenous effect on compensation. Whether this is the case can only be determined by specifying more general models than have

Table 3
Cross-sectional and within-firm variances of the covariates

Parameter	t^{cs}	mean(t^f)	min(t^f)	max(t^f)
Response variable y	0.32	0.52	0.00	4.41
Dollar volatility of returns	138.62	3.12	−0.00	1585.52
Percent volatility of returns	0.02	0.00	−0.00	1.19
CEO share ownership	0.13	0.01	0.00	0.43
Tobin's Q	1.47	0.31	0.00	15.13
Financial leverage	0.12	0.04	−0.04	0.76
Interest coverage	15.86	2.85	0.00	631.20
ln(total assets)	0.38	0.03	−0.00	0.48
Last year stock return	0.05	0.25	0.00	21.99

The table displays both cross-sectional and within-firm variances of the variables. The estimates are based on the sample of 1530 firms, 10859 firm-year observations for years 1992–2002. The following notation is used: t^{cs} denotes the ratio of the cross-sectional variance to the mean, calculated for each variable, where cross-sectional variance refers to the variance of the corresponding variable, averaged over time for each firm; t^f denotes the ratio of the within-firm variance to the within-firm mean, calculated separately for each firm.

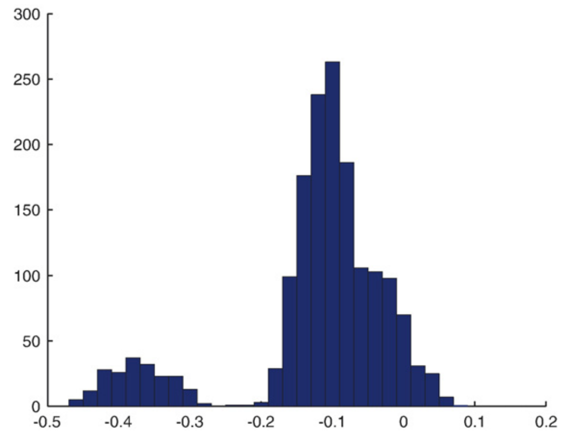

Fig. 2. Histogram of the coefficients on CEO share ownership obtained by fitting Tobit regression (3) to the time series data on each firm in the sample.

been done so far. If heterogeneity is important, then modeling and controlling for that heterogeneity should lead to a better understanding of the information in the data.

3. Model specification

To formulate our expanded Tobit hierarchical model, let x_{1it} denote a k vector of covariates (excluding the constant) on the ith firm in year t, whose effect on the outcome is not heterogenous across firms, and let w_{it} be a q vector of other covariates whose effect on the outcome is modeled heterogeneously. Now let

$$y_{it} = \begin{cases} x_{1it}'\beta_1 + w_{it}'\tilde{\beta}_i + \varepsilon_{it}, & \text{if } x_{1it}'\beta_1 + w_{it}'\tilde{\beta}_i + \varepsilon_{it} > 0, \\ 0, & \text{if } x_{1it}'\beta_1 + w_{it}'\tilde{\beta}_i + \varepsilon_{it} \le 0; \end{cases} \tag{4}$$

where $\tilde{\beta}_i$ is a q dimensional random vector whose distribution is assumed to depend on a set of firm-specific covariates A_i. In particular, we assume that

$$\tilde{\beta}_i = A_i\beta_2 + b_i$$

where A_i is a $q\times r$ matrix and b_i is a $q\times 1$ random vector distributed as multivariate normal with mean vector 0 and full covariance matrix D.

In our model, we let four agency-related covariates have firm-specific coefficients. We assume that the time-averaged value of GI for each firm explains the heterogeneity in $\tilde{\beta}_i$. To be more specific, w_{it} is defined as follows:

$$w_{it} = (1, \text{CEO share ownership, relative `noise' in ROE,} \;\; \text{Tobin's}Q, \quad \text{financial leverage})',$$

whereas x_{1it} contains the rest of the covariates described in the previous section. In addition, the model of the heterogeneity is

$$\begin{pmatrix} \tilde{\beta}_i \\ \tilde{\beta}_i \\ \vdots \\ \tilde{\beta}_i \end{pmatrix} = \begin{pmatrix} 1 & \text{GI} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \text{GI} & & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & \text{GI} \end{pmatrix} \begin{pmatrix} \beta_{211} \\ \beta_{212} \\ \beta_{221} \\ \beta_{222} \\ \vdots \\ \beta_{251} \\ \beta_{252} \end{pmatrix} + \begin{pmatrix} b_{i1} \\ b_{i2} \\ \vdots \\ b_{i5} \end{pmatrix}.$$

Therefore, the first equation in this specification models the heterogeneity of the coefficient $\tilde{\beta}_{i1}$ in terms of the covariate GI, the second equation models the heterogeneity of the coefficient $\tilde{\beta}_{i2}$ whereas the last equation models $\tilde{\beta}_{i5}$. Other specifications of the second-stage of our model are discussed below.

In our analysis we are interested in the parameters $\beta=(\beta_1, \beta_2)$ and in the matrix D which measures the extent of firm-specific heterogeneity in compensation practices above and beyond what is explained by the covariates. By modeling heterogeneity in a richer way than in previous work, and by allowing firm-specific covariates in A_i to partially explain the heterogeneity of the coefficients, our model can be expected to lead to smaller residual heterogeneity (equivalently, smaller diagonal elements of D) and, therefore, more precise estimates of the parameters β. This is borne out in our empirical findings. Because we let D be a full matrix, the off-diagonal elements of D capture the covariance in the firm-specific effects induced perhaps by some common unobserved factors that affect $\tilde{\beta}_i$.

We refer to the model we have specified as Model H, where 'H' stands for hierarchical. After shortly describing the estimation method, we turn to discussing the results of fitting this model to the data.

4. Model fitting

The hierarchical model that we have specified in the preceding section is difficult to analyze by frequentist methods. The likelihood function is not available directly and even with simulation methods finding the maximum likelihood estimates is a challenge. For this reason we turn to Bayesian methods which over the past 10 years have gone through a period of almost revolutionary growth and development (see for example Chib and Greenberg (1995, 1996) and Chib, 2001) in large part due to the emergence of Markov Chain Monte Carlo (MCMC) simulation methods and new ways of dealing with models with latent variables (Chib, 1992; Albert and Chib, 1993). In the Bayesian approach, the parameters are treated as random variables, and given a prior distribution, the focus is on the posterior distribution of the parameters, which by Bayes' theorem is proportional to the prior distribution times the likelihood function (standard reference texts of the Bayesian approach to statistical inference include Zellner (1971) and Bernardo and Smith (1994).

In our model, this prior-posterior calculation can be conducted by Markov Chain Monte Carlo methods even though the likelihood function is largely intractable. This is achieved by enlarging the parameter space with suitable latent variables for the responses that are censored. MCMC can be applied to sample this augmented posterior distribution in an iterative process: sampling latent variables conditioned on the parameters and then sampling parameters conditioned on the simulated latent variables. In particular, within one cycle of the MCMC algorithm, conditioned on the current values of the parameters, the latent censored responses are simulated from truncated normal distributions (according to the prescription outlined in Chib, 1992). Given values of the latent variables, the model resembles a continuous data hierarchical model and values of the parameters and the heterogenous coefficients are sampled from relevant conditional posterior distributions, as discussed in Chib and Carlin (1999). It can be shown that iterations of these cycles produce draws from the joint posterior distribution of the parameters and latent variables.

The marginal posterior distribution of the various parameters can be summarized in terms of the sample means, standard deviations and quantiles of the sampled draws. Full details of the MCMC algorithm that we have developed to fit our proposed model are given in Appendix A. In that appendix we also provide the prior distributions that we use in our analysis. The priors we have chosen are non-restrictive and allow the data to play the major role in determining the posterior distribution. We center the prior on the covariate coefficients at zero to reflect the currently prevalent view in the empirical literature that the covariates have no impact on option grants to CEOs. The marginal prior variances on these covariates are each set at ten. The results we present are robust to reasonable changes in our prior hyperparameters, as we have confirmed.

In order to cast light on the importance of heterogeneity in the data, we compare the various models through the model marginal likelihood and the pairwise Bayes factors (ratios of marginal likelihoods). The marginal likelihood of each of our model specifications is computed by the method of Chib (1995), based on the output of the MCMC simulations. The Chib method is also outlined in Appendix A. Higher values of the marginal likelihood indicate greater support for the model over the alternatives.

5. Estimation results

In this section, we discuss the estimation results obtained by fitting Model H. Table 4 shows the posterior means for the coefficients $\beta=(\beta_1, \beta_2)$, and the posterior means for the elements of matrix D are reported in Table 5.

Table 4
Model H, coefficients β_1 and β_2

Covariates in x_i	Mean β_1	Covariates in w_i	Covariates in A_i	Coef. name	Mean β_2
vol($ returns)	−0.006*	**Constant**	Constant	β_{211}	1.663*
	(0.002)				(0.210)
Relative noise in ROE	0.006*		GI	β_{212}	0.008
	(0.002)				(0.019)
Restricted stock	0.005	**CEO share ownership**	Constant	β_{221}	−4.128*
	(0.008)				
Option holdings($)	0.046		GI	β_{222}	(1.729)
	(0.393)				−0.020
New CEO	1.233*				(0.214)
	(0.084)				
Departing CEO	−0.557*	**Tobin's Q**	Constant	β_{241}	−0.030
	(0.061)				(0.047)
Zero-dividend dummy	0.248*		GI	β_{242}	−0.000
	(0.034)				(0.005)
Tax loss c.f. dummy	−0.030	**Financial leverage**	Constant		
	(0.028)			β_{251}	0.386
Interest coverage	−0.001		GI	β_{252}	(0.318)
	(0.001)				−0.037
GI	0.008				(0.034)
	(0.013)				
ln(total assets)	−0.128*				
	(0.011)				
$\log_{10}$ marginal likelihood		−5752.85			

This is our main model with full heterogeneity specification (constant, CEO share ownership Tobin's Q, and financial leverage have heterogeneous coefficients, linear in GI; Table 5 reports their variance–covariance matrix). The response is our pay-for-performance sensitivity measure (1). The covariates in bold are agency-related. The sample consists of 1530 firms for years 1992–2002. The estimates are marked with '*' if the posterior distribution implies that a 95% confidence interval does not contain zero. The model also includes last year stock return and time and industry dummies. Table 6 reports the coefficients on the industry dummies.

5.1. Agency problems and option grants

We derive our predictions about the expected signs of the coefficients on agency-related covariates from the main hypothesis that firms use option grants to efficiently manipulate incentives of their CEOs (for more on this hypothesis see, for example, Core and Guay, 1999). We also assume that the benefit of increasing CEOs PPS through option grants decreases when the gap between the current and the optimal levels of PPS of the CEOs portfolio becomes smaller. In other words, we expect PPS of option grants to be negatively related to the current level of PPS, and positively related to the optimal level of the PPS of the CEO portfolio.

To measure the current level of PPS of the CEO portfolio, we include the following three covariates: CEO share ownership, restricted stock ownership, and the dollar value of the options currently held by the CEO, which is calculated using the Black–Scholes method. Since the optimal level of the CEO PPS is mainly determined by the agency factors in the firm, we include the following six variables that the prior literature has identified as related to different aspects of this relationship: volatility of dollar returns, Tobin's Q, relative noise in ROE, financial leverage, and indicator variables for new and departing CEO (the relationship between agency and these covariates has been

Table 5
Model H, Matrix D

	Constant	CEO stock ownership	Tobin's Q	Financial leverage
Constant	0.284*	2.462*	−0.050*	−0.470*
CEO stock ownership	2.462*	83.116*	−0.376*	−5.772*
Tobin's Q	−0.050*	−0.376*	0.034*	0.055*
Financial leverage	−0.470*	−5.772*	0.055*	1.701*

This table continues reporting the results of the same model as in Table 4. It reports the variance–covariance matrix for the heterogeneous coefficients on CEO share ownership, Tobin's Q, financial leverage and the constant.

extensively discussed in the literature, see for example Yermack, 1995; Aggrawal and Samwick, 1999). Note that, in our model, CEO share ownership, Tobin's Q, and leverage have firm-specific coefficients. Moreover, to analyze heterogeneity, we model each firm-specific coefficient as a linear combination of a constant term and governance index (GI). Therefore, for each variable that has a firm-specific coefficient, we discuss two estimates (two corresponding elements of vector β_2): the constant term and the coefficient on GI.

We now turn to discussing the coefficients on each agency-related variable in more detail.

5.1.1. PPS of the CEO current portfolio

We estimate the coefficients on CEO share ownership β_{221} (constant) and β_{222} (coefficient on ln(GI)) to be negative. Fig. 3 shows the posterior densities for these coefficients. The posterior density for β_{222} implies that this coefficient is insignificantly different from zero. The constant term, on the other hand, is significantly negative: the mean of the posterior is equal to −4.13, and more than 97.5% of the posterior density lies to the left of zero.

The negative coefficient on CEO share ownership indicates that, on average, the relationship is consistent with agency theory recommendations. In particular, CEOs whose wealth is already closely tied to shareholders' wealth receive less option grants. It is worth noting that the absolute value of our coefficient estimate may be biased upwards if the CEOs who own a large percentage of the firm's stock are more likely to sell their shares in response to new option grants, as suggested by Ofek and Yermack (2000). This argument, however, does not affect the implication that firms use option grants to improve incentives. It may additionally be argued that CEOs with high share ownership are likely to have more influence on the firm's compensation policies. Therefore, a negative relationship between CEO share ownership and option grants may arise, if CEOs prefer cash compensation to option grants. While a number of papers find support for this hypothesis (e.g. Anderson and Bizjak, 2003; Yermack, 1996), Bertrand and Mullainathan (2001) present evidence that implies the opposite: CEOs with more discretion over their compensation tend to receive more option grants. We revisit this hypothesis in our discussion of the coefficient on governance index, where we present additional support for the evidence presented in Bertrand and Mullainathan (2001).

The economic significance of the coefficients on CEO share ownership is difficult to infer due to non-linear structure of the model: not only is the coefficient on CEO share ownership dependent on firm size, it is also correlated with other firm-specific coefficients. To analyze the implied effect of a change in CEO share ownership on option grants to CEO, we compute predicted values for the two response variables first using the observed share ownership levels, and next using the ownership levels increased by 5%. We find that on average, our measure of pay-for-performance sensitivity (response variable y_1) decreases by 14.6% with the standard deviation of the percentage decrease equal to 6.9%. The probability of observing zero for y1 on average increases by 7.9%, with the standard deviation equal to 3.0%. Our proxy for compensation mix (response variable y_2) decreases by 17.6% on average, with the standard deviation of the percentage decrease equal to 7.2%. The probability of observing zero for y_2 on average increases by 9.0%, with the standard deviation equal to 2.9%.

We find the coefficients on both restricted stock holdings and dollar value of CEO option holdings to be positive but insignificant. Although we hypothesized a negative relationship, finding a positive (although insignificant) coefficient

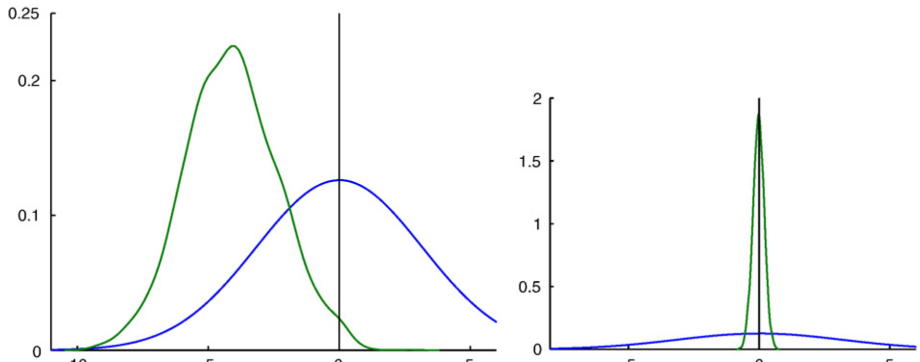

Fig. 3. Posterior distributions of the constant term (left) and the coefficient on GI (right) corresponding to CEO share ownership obtained by fitting Model H.

may not be surprising: firms typically achieve significant increases in PPS by granting options to their CEOs during the course of several years.

5.1.2. Tobin's Q

Following Smith and Watts (1992) and Gaver and Gaver (1993), we interpret Tobin's Q as a proxy for growth opportunities and relative importance of intangible assets in the firm. Given this interpretation, we expect Tobin's Q to be positively related to option grants to CEO. Our results, however, do not support this hypothesis. Both coefficient estimates on Tobin's Q, the constant term and the coefficient on GI, are negative and insignificant.

The lack of support for our hypothesis concerning the coefficient on Tobin's Q, may partly be due to the endogeneity of this variable. Himmelberg et al. (1999) argue that firm performance and managerial compensation are endogenously determined by such factors as firm's monitoring technology, intangible assets, firm's market power. Since Tobin's Q can also be interpreted as a measure of firm performance, our estimates of the coefficients may be biased. In addition, the value of Tobin's Q may depend to some extent on the actions of the CEO, which in turn depend on the form of the CEOs compensation. Since option grants provide incentives to the CEO to increase the market value of the firm, and therefore, increase the value of Tobin's Q, a positive bias may be present in our estimates of the coefficients corresponding to Tobin's Q. Since our coefficient estimates are negative, this argument only reinforces our findings.

5.1.3. Financial leverage

We expect to observe negative coefficients on leverage for the following three reasons. First, firms with high leverage are closely monitored by the debt holders, and therefore do not have to depend on option grants for providing incentives. Second, option grants align CEOs' interests with those of stockholders, but not debt holders. Third, the probability of bankruptcy increases with leverage. Thus, the effect of option grants on incentives is less favorable for firms with large debt.

Somewhat surprisingly, both coefficients corresponding to leverage are insignificant. Moreover, the constant term is positive, with slightly more than 88% of the posterior distribution lying to the right from zero, as can be seen from Fig. 4.

5.1.4. Relative noise in ROE

Relative noise in ROE reflects the relative informativeness of accounting information. We expect a positive relationship between this variable and option grants to CEO. Indeed, the posterior mean of the coefficient on this variable is 0.006, which is significant at the 5% level. To the extent that noise in ROE can be manipulated by the CEO, however, one can argue that the coefficient on this variable is subject to an endogeneity bias. It is not clear, though, how to determine the direction of this bias, since it depends on the effect of option grants on the volatility of both stock returns and return on equity. While option grants are likely to encourage risk-taking on behalf of the CEO, and therefore, lead to an increase in stock return volatility, it is not clear how option grants may affect the relative volatility of an accounting variable such as ROE.

5.1.5. Volatility of stock returns

A class of agency models represented by Holmstrom and Milgrom (1987) predicts a negative relationship between PPS of managerial portfolio and volatility of the corresponding performance measure. Assuming that firms use option

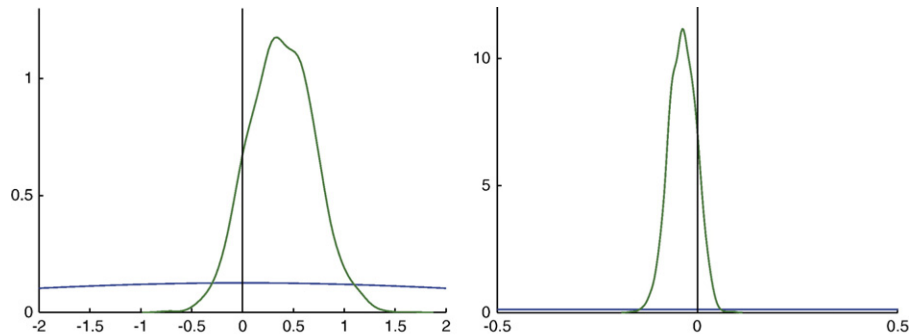

Fig. 4. Posterior distributions of the constant term (left) and the coefficient on GI (right) corresponding to leverage obtained by fitting Model H.

grants to close the gap between the current and the optimal level of PPS, we hypothesize that, after controlling for the current level of PPS, we should find a negative relationship between PPS of current-year option grants and stock return volatility.

Existing empirical results on the relationship between PPS of CEOs wealth and variance of stock returns are controversial. Aggarwal and Samwick (1999) find a significant negative relationship between volatility of stock returns and sensitivity of total managerial compensation to firm performance. Although they use two measures of stock return volatility – volatility of dollar returns and volatility of percent returns – they argue that volatility of dollar returns is more appropriate since it accounts for firm size. Contrary to Aggarwal and Samwick (1999), Kaplan and Stromberg (2002) find a positive relationship between managerial PPS and risk levels in start-up firms. Additionally, Chen and Jiang (2003) argue that firms with more growth opportunities and with CEOs that face strong career concerns are likely to exhibit a positive relationship between PPS and volatility. Several reasons have been offered for the divergence in the empirical results. First, volatility of the firm's stock can be at least partially manipulated by the CEO. Moreover, the extent to which the CEO is able to control volatility may depend on the firm's environment. If stock return volatility is indeed endogenous, then the coefficient on this covariate may be hard to interpret in line with Holmstrom and Milgrom (1987). Second, option grants to CEOs may have an ambiguous effect on CEOs' actions, which in particular depends on the current level of stock return volatility (Ross, 2004; Ju et al., 2003). Third, high stock return volatility may affect CEOs' career concerns, and in particular, may cause the CEO to be excessively cautious, which may be against the owners' interests. The optimal level of PPS of the CEO portfolio would then be positively related to stock return volatility, which in turn may lead to a positive bias in the coefficient on this variable (Chen and Jiang, 2003).

We find that, consistent with Holmstrom and Milgrom (1987), volatility of dollar returns has a significant negative coefficient equal to −0.006. It is interesting to note that the models presented in both Chen and Jiang (2003) and Ju, Leland and Senbet (2003) imply a heterogeneous relationship between volatility and the PPS. We find, however, that altering Model H by allowing this coefficient to be heterogeneous adversely affects model performance: the log marginal likelihood drops from −5752.85 to −5850.3.

5.1.6. CEO employment cycle

Both indicator variables for new and departing CEO, representing CEO employment cycle, have significant coefficients with the expected signs. The coefficient on new CEO dummy is positive, equal to 1.23. These estimates indicate that during their first years at the position, CEOs receive substantial option grants, probably since they typically have a low initial ownership level. If the CEO is expected to leave the firm, option grants may not be efficient anymore in providing long-term incentives, since the CEO is likely to exercise the options upon departure. This hypothesis is supported by negative significant estimate for the departing CEO dummy, equal to −0.56.

5.2. Control variables

As control variables, we use variables that proxy for financial flexibility of the firm, as well as ln(total assets), last year stock return, and time and industry dummies. Table 4 reports the posterior means for the coefficients on these control variables.

5.2.1. Financial flexibility and option grants

Of the three control variables that capture financial flexibility – zero-dividend dummy, tax-loss carry forward dummy, and interest coverage – the first one is significant and has the expected positive sign. The coefficients on the other two variables are insignificant. Positive and significant coefficient on the zero-dividends dummy, equal to 0.25, supports the hypothesis that firms, which are likely to have scarce cash reserves, provide a smaller share of the compensation in the form of cash. In particular, for firms that do not distribute options, the value of the response variable is on average 25% larger.

5.2.2. Governance index

The governance index, developed in Gompers et al. (2003), reflects the strength of shareholder rights in the firm. Smaller values of the index correspond to weaker shareholder rights (i.e. weaker governance). There are several factors that may affect the relationship between the governance index and option grants to CEOs. First, if an improvement in monitoring of the CEO that accompanies stronger shareholder rights is a substitute for the incentives provided through option grants, then firms with low governance index (strong shareholder rights) will choose to grant less options.

Second, in better governed firms, CEOs are likely to have less control over compensation decisions. Therefore, firms with low governance index will be less likely to deviate from optimal compensation practices. The implication of this argument on the relationship between the governance index and option grants to CEOs is ambiguous: although one would expect the dollar value of CEO compensation to increase when governance is weak, it may be difficult to predict whether the increase will come in the form of cash or options. Bertrand and Mullainathan (2001) argue that, for political reasons, it is easier for CEOs to obtain a raise in their total compensation by choosing large option grants. Anderson and Bizjak (2003), Core et al. (1999), and Yermack (1996), however, find that firms with weaker governance structures choose to provide larger cash compensation to their CEOs. Moreover, CEOs in firm with weak governance structures also typically have portfolios with low PPS.

We find the coefficient on GI to be positive, which is consistent with both the substitutability between direct monitoring and option grants to CEOs and the argument of Bertrand and Mullainathan (2001). This coefficient, however, is insignificantly different from zero.

5.2.3. Time and industry dummies

It is interesting to note that the time dummy coefficients are increasing with years for both response variables, as can be seen from the box plot shown in Fig. 5. The estimated time trends for both response variables confirm the common view, seen in both popular and academic literature, that the popularity of option grants for top executive officers has been increasing during the boom years of the '90s. Surprisingly, it also implies that the trend continued until the year 2001, after the stock market has already started to decline: the coefficient for year 2001 is larger than that for year 2000. Our estimates imply, however, that in year 2002 there was a significant decline in option grants: the coefficient for year 2001 is positive and significant.

The posterior means for the coefficients on industry dummies are reported in Table 6 (the omitted dummy represents a firms with non-identified industry classification, which corresponds to SIC code 99). We find that the only industry that has a negative and significant coefficient for both dependent variables is utilities. This was anticipated, since utilities industry is subject to numerous regulations and restrictions. The estimates of the coefficients on industry dummies provide evidence on heterogeneity of the compensation practices across industries. The estimated values range from −0.37 to more than 0.3 for both response variables, which implies a more than 60% variation in the average values of the response variables across industries, all else equal.

5.2.4. Heterogeneity

As discussed above, the coefficients on industry dummies reflect heterogeneity in firms' compensation practices. Stronger evidence of heterogeneity, however, is provided by the estimates of variance–covariance matrix D for the random effects b_i, which is reported in Table 5.

The large size of the diagonal elements in matrix D, showing variances of firm-specific coefficients, indicate a significant diversity in firms' sensitivity to the corresponding covariates, beyond what was captured by the firm size variable (included in matrix A_i). While all the diagonal elements in D are comparable or even larger than the mean

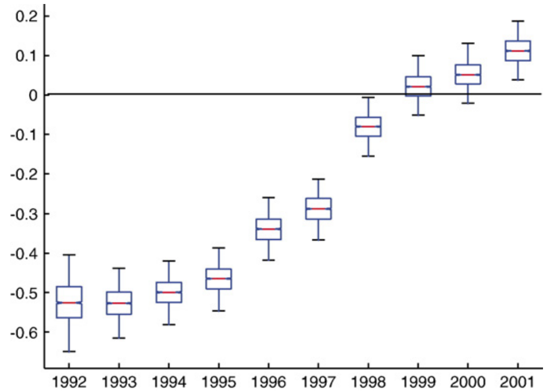

Fig. 5. Box plot for the coefficients on time dummies obtained by fitting Model H.

Table 6
Model H (cont.): coefficients on industry dummies

Industry name	SIC code	# of firms	Mean (β)
Food	1,2,7,20,21	51	0.0032
Mining and extraction	10–14	54	0.1165
Construction	15,17	7	0.3527
Manufacturing	22–25,31	69	0.1336
Paper	26	30	0.0375
Printing and publishing	27	32	0.0758
Chemicals	28	121	0.1777
Petroleum refining	29	18	0.2081
Petroleum refining	29	17	0.2081
Rubber and plastics	30	11	0.1163
Stone, clay, and glass	32	40	0.1277
Primary metals	33	24	0.2072
Fabricated metals	34	98	0.3232
Industrial machinery	35	110	0.1971
Electrical equipment	36	46	0.2135
Transport equipment	37	67	0.0986
Instruments	38	29	0.1589
Rail, ground, and other transport	40,44,46–47	14	0.1596
Airline transportation	45	33	0.3901
Communications	48	124	0.0824
Utilities	49	54	−0.3657*
Wholesale trade	50,51	133	0.2279
Retail trade	52–59	2	0.1053
Depository institutions	60	46	0.5904
Other financial	61,62,64–67	50	0.2769
Insurance carriers	63	183	0.3389*
Services	70,72,73,75,76,78–84,86–89	67	0.1906

This table continues the results presented in Tables 4 and 5. The omitted dummy represents non-identified firms with SIC code equal to 99. The estimates are marked with '*' if the posterior distribution implies that a 95% confidence interval does not contain zero.

values for the corresponding coefficients, the variance of the coefficient on CEO share ownership stands out: the posterior mean for this variance is equal to 83.1. This number is several times larger than the absolute mean value of the constant term in the equation for the coefficient on CEO share ownership, equal to −4.13.

Off-diagonal terms in matrix D show significant correlation between all the covariates. The coefficient on CEO share ownership is negatively correlated with both the coefficient on Tobin's Q and the coefficient on leverage, with correlations equal to −0.22 and −0.49, respectively. Coefficients on Tobin's Q and leverage, on the other hand, are positively correlated, with correlation equal to 0.23. In order to illustrate an interpretation of these estimates, consider the negative correlation between CEO share ownership and Tobin's Q. It implies that firms, which are more sensitive to CEO share ownership (i.e. which have a smaller coefficient on CEO share ownership), are more likely to increase option grants in response to a rise in Tobin's Q.

5.3. Option grants and temporary changes in agency

The model specifications we have used embed the view that the current pay-for-performance sensitivity levels of the CEO's portfolio are significantly below the optimal level, and thus option grants should be used to (possibly slowly) bring the level of pay-for-performance sensitivity closer to the optimum. A number of studies, however, argue that the current pay-for-performance sensitivity in many firms is close to the equilibrium level; see, for example, Palia (2001). Probably the most relevant to our study is Core and Guay (1999) who argue that option grants are used to correct temporary deviations from the optimal level of pay-for-performance sensitivity of the CEO's portfolio.[2]

[2] While Core and Guay (1999) interpret their results as indicating that option grants correct for deviations from the optimum, their empirical model specification in fact relates option grants to levels of agency-related variables, and not changes in these variables.

To see whether option grants are given to correct for temporary deviations from the optimal level of incentives, we construct lagged first differences of the agency-related variables. We use the lagged instead of the current differences to account for the time needed to implement a revision to compensation in response to changes in firm and CEO characteristics. As a result, our sample is reduced to 8751 observations on 1158 firms. We next estimate Model Diff with heterogeneous coefficients on these four lagged and differenced variables. The first column of Table 7 reports the results of the analysis. The coefficients on almost all the variables are insignificant, with the exception of the coefficients on the changes in CEO share and option ownership. The positive sign of these coefficients, however, contradicts the hypothesis that an increase in CEO share ownership reduces the need for option grants. Including additional control variables as well as agency-related covariates used in Model H only further weakens the results. The results, therefore, do not support the hypothesis that option grants are used to address recent changes in CEO or firm characteristics.

The second column of Table 7 shows that the lack of significance is unlikely to be due to model specification or a smaller size of the data set. Indeed, when levels instead of first differences are used as covariates in the same model specification, we find significant coefficient estimates on most of the variables. More importantly, this model is preferred on the marginal likelihood metric.

5.4. Importance of heterogeneity assumptions

The discussion in the earlier sections shows that Model H produces substantial evidence of unexplained heterogeneity in coefficients on CEO share ownership, Tobin's Q and leverage. In this section, we test whether the heterogeneity assumptions present in Model H contributes to explaining CEO option award practices. In order to do that, we estimate a model with no heterogeneous coefficients, which we call Model NoH, and two models: Model B1 and Model B2, where 'B' stands for basic, where only constant term is heterogeneous. In Model NoH, there is not matrix w, and in Models B1 and B2, matrix w contains only a constant term and matrix A_i is an identity matrix. Models NoH and B1 use the same covariates as Model H, and therefore, matrix x is the same as in Model H. In Model B2, matrix x additionally contains products of governance index and CEO share ownership, Tobin's Q, and leverage (the variables that comprise w in Model H). Comparing the results and performances of Models H, NoH, B1 and B2 will show the extent to which our heterogeneity assumptions are capable of explaining CEO compensation practices.

Table 8 reports the results of estimating models NoH, B1 and B2. As in the case of Model H, the coefficient estimates produced by models NoH, B1 and B2 show support for the hypotheses concerning coefficients on CEO share

Table 7
Models D and B, coefficients β

	Model diff	Model B
Constant	1.863*	2.080*
Δ (CEO share ownership)	0.010*	
Δ(restricted stock)	−2.390	
Δ(option holdings($))	0.966*	
Δ(Tobin's Q)	0.012	
Δ(financial leverage)	−0.073	
Δ(vol($ returns)	−0.690	
Δ(relative 'noise'in ROE)	−7.400	
CEO share ownership		−2.453*
Restricted stock		−1.740
Option holdings($)		1.260*
Tobin's Q		−0.032*
Financial leverage		0.077
vol($ returns)		−0.420*
Relative 'noise'in ROE		0.005*
New CEO dummy	1.134*	1.181*
Departing CEO dummy	−0.432*	−0.474*
$\log_{10}$ marginal likelihood	−4762.6	−4468.2

Model D tests whether option grants respond to changes in the firm or CEO characteristics that may affect the optimal level of incentives. The response is the pay-for-performance sensitivity measure (1). The models also include last year stock return, firm size, and time dummies. Model B is the same as Model D except that levels of agency-related variables are used instead of first differences. The covariates in bold are agency-related. The sample contains 8751 observations on 1158 firms for years 1994–2002. The estimates are marked with '*' if the posterior distribution implies that a 95% confidence interval does not contain zero.

Table 8
Models B1 and B2, coefficients β

	Model NoH	Model B1	Model B2
Constant	1.805*	1.788*	1.642*
CEO share ownership	−3.135*	−3.200*	−4.087*
Restricted stock	0.000	0.013	0.013
Option holdings ($)	1.681*	0.847*	0.832*
Tobin's Q	−0.025*	−0.011	0.025
Financial leverage	0.016	−0.130	0.536*
vol($ returns)	−0.006*	−0.001*	−0.006*
Relative 'noise'in ROE	0.005*	0.005*	0.005*
New CEO dummy	1.283*	1.310*	1.306*
Departing CEO dummy	−0.569*	−0.563*	0.563*
GI	0.009*	0.011*	0.026*
GI * CEO share ownership			0.122
GI * Tobin's Q			−0.005
GI * leverage			−0.047
Zero-dividend dummy	0.290*	0.252*	0.251*
Tax loss c.f. dummy	−0.011	−0.033	−0.030
Interest coverage	−0.003*	−0.002*	−0.002*
ln(assets)	−0.127*	−0.133*	−0.132*
$\log_{10}$ marginal likelihood	−5973.3	−5803.2	−5797.7

These models are used to test for the importance of model flexibility concerning coefficient heterogeneity. Model NoH has no heterogeneous coefficients, and Models B1 and B2 allow for heterogeneity only the constant term. Additionally, unlike Model H, these models do not use the observed variables to explain the heterogeneity in coefficients. Model B2 differs from Model B1 only in that Model B2 includes interaction terms (GI * CEO share ownership), (GI * Tobin's Q), and (GI * leverage). The response is the pay-for-performance sensitivity measure (1). The covariates in bold are agency-related. The sample consists of 1530 firms for years 1992–2002. The estimates are marked with '*' if the posterior distribution implies that a 95% confidence interval does not contain zero. The models also include last year stock return and time and industry dummies.

ownership, dollar return volatility, relative noise in ROE, and control variables for departing and new CEOs. In addition, these models also produce positive and significant coefficients on CEO option holdings and governance index. More importantly, however, all three alternative models NoH, B1 and B2 have a significantly smaller marginal likelihood, providing evidence in favor of the hierarchical model specification (Model H).

As an additional test of model performance, we compare Model H to Models B1 and B2 in terms of their ability to predict observations out-of-sample. We first randomly select 100 firms from the sample. Next, we apply the models of interest to the rest of the sample, and calculate predicted expected value for the two response variables for the selected sample of 100 firms. Since both response variables have many observations clustered at zero, we also estimate the predicted probability of the zero outcome for each dependent variable. We assume that the predicted outcome is zero if the estimated probability of observing zero is bigger than 0.5. We define the average prediction error as a ratio of the average difference between the predicted and the true values, divided by the average true value of the dependent variable. To find the probability of observing zero when the predicted outcome is zero, $\Pr(y=0|y^{pr}=0)$, we divide the number of the correctly predicted zero outcomes by the total number of predicted zero outcomes for the dependent variable. The probability $\Pr(y>0|y^{pr}>0)$ is defined similarly.

Table 9
Out-of-sample prediction for Model H

	Model B1	Model B2	Model H
Average prediction error (as % of y)	1.18	53.8	55.3
$\Pr(y=0\|y^{pr}=0)$	25.3	48.4	55.8
$\Pr(y>0\|y^{pr}>0)$	1.0	76.9	79.3

This table reports statistics on out-of-sample prediction performance of our main model, Model H, which has full heterogeneity specification. The coefficients estimated using this model are reported in Tables 4–6. Average prediction error is calculated as the average of the absolute value of the difference between the predicted and the realized values of the dependent variable. y^{pr} denotes the predicted value of the dependent variable. The total sample consists of 1530 firms, 10859 firm-year observations for years 1992–2002. Out-of-sample prediction statistics was obtained by randomly removing 100 firms from the sample, estimating the model on the remaining sample, and using the estimates to obtain predicted values for the response variables.

Table 10
Model T, coefficients β_1 and β_2

Covariates in x_i	Mean β_1	Covariates in w_i	Covariates in A_i	Coef. Name	Mean β_2
vol($ returns)	−0.005	**Constant**	Constant	β_{211}	2.176*
Relative noise in ROE	0.233		GI	β_{212}	−0.018
Restricted stock	0.002		Age	β_{213}	−0.004
Option holdings($)	0.009		Tenure	β_{214}	−0.000
New CEO	1.225*				
Departing CEO	−0.566	**CEO share ownership**	Constant	β_{211}	−2.274
Zero-dividend dummy	0.250*		GI	β_{212}	0.136
Tax loss c.f. dummy	−0.055		Age	β_{213}	−0.055
Interest coverage	−0.001		Tenure	β_{214}	0.032
GI	0.026				
Age	−0.007	**Tobin's Q**	Constant	β_{211}	−0.092
Tenure	0.003		GI	β_{212}	0.006
ln(assets)	−0.135*		Age	β_{213}	0.000
			Tenure	β_{214}	−0.001
		Financial leverage	Constant	β_{211}	0.534
			GI	β_{212}	−0.016
			Age	β_{213}	−0.003
			Tenure	β_{214}	−0.003
$\log_{10}$ marginal likelihood	−2497.84				

This model has full heterogeneity specification. Constant, CEO share ownership Tobin's Q, and financial leverage have heterogeneous coefficients. These coefficients are modeled as linear in GI, age, and tenure. Table 11 reports their variance–covariance matrix. The response is our pay-for-performance sensitivity measure (1). The covariates in bold are agency-related. The sample consists of 773 firms, 4656 firm-year observations for years 1992–2002. The estimates are marked with '*' if the posterior distribution implies that a 95% confidence interval does not contain zero. The model also includes ln(total assets), last year stock return, time and industry dummies.

Table 9 shows the results of our prediction analysis for Model H and Models B1 and B2. Model B1 produces an average prediction error that is more than 60% larger than the one obtained for Model H, while model B2 produces an error that is statistically indistinguishable from the error obtained for Model H. Probability of a correct prediction of a zero outcome is larger for Model H than for both Model B1 and Model B2, with the difference of more than 30% for Model B1. Model H also has a higher probability of predicting a positive outcome than Model B2. It appears that Model B1 does better than Model H in predicting a positive outcome: according to our tests, if Model B1 predicts the outcome to be positive, the true outcome is in fact positive in 100% of cases, while for Model H the percentage of correct predictions is equal to 79.3. It is important to note, however, that Model B1 is very unlikely to predict an occurrence of a positive outcome: out of 752 observations, Model B1 predicts only 4 to be positive, Model H predicts 657 to be positive, while the true number of positive outcomes is 563. Not surprisingly, in all 4 instances when Model B1 predicted a positive outcome, the prediction turned out to be correct.

Note that, in specifying Model H, we chose to omit the data on tenure and age of the CEO because a large number observations were missing. In order to see, however, if our results are robust to specification, we next estimate the following version of Model H. In this version, the matrices X and A include two additional variables, CEO age and CEO tenure. We will refer to this model as Model T. Since the data on CEO age is incomplete, Model T is estimated on a subset of 713 firms, for which the data on CEO age is available. Next, we compare performance of Model T to that of Model H.

The results of fitting Model T are shown in Tables 10 and 11.

Table 11
Model T, Matrix D

	Constant	CEO stock ownership	Tobin's Q	Financial leverage
Constant	0.234*	0.273	−0.048*	−0.171
CEO stock ownership	0.273	17.125*	−0.102	0.248
Tobin's Q	−0.048*	−0.102	0.046*	0.002
Financial leverage	−0.171	0.248	0.002	0.756*

This table continues reporting the results of the same model as in Table 10. It reports the variance–covariance matrix for the heterogeneous coefficients on CEO share ownership, Tobin's Q, financial leverage and the constant.

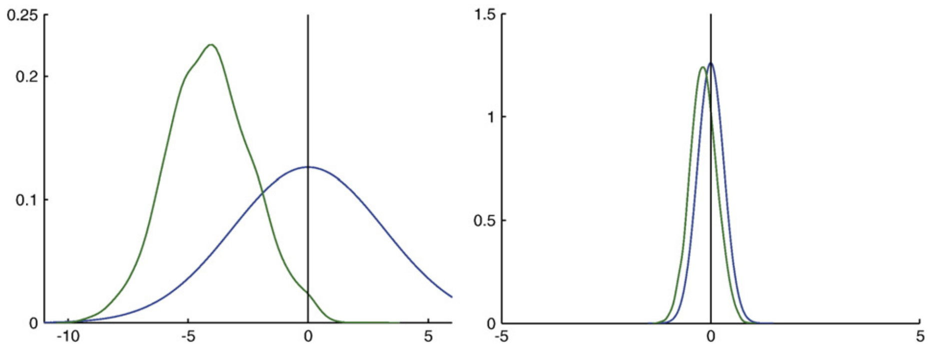

Fig. 6. Prior and Posterior distributions of the constant term corresponding to CEO share ownership obtained by fitting Model H with prior variance of 10 (left) and 0.1 (right).

The estimates produced by this model are also similar to those presented for Models H and B on the larger data set. The estimates mainly differ in that several of the covariates in Model T have coefficients significant only at *the* 10% level, while they were significant at the 5% level in Models H and B. The results on the matrix D, reported in Table 11, show that adding age and tenure does not significantly reduce the amount of unexplained heterogeneity in firm-specific coefficients. The largest variance, that of the coefficient on CEO share ownership is equal to 17.1 when age and tenure are included, while it equals 10.4 when the two variables are omitted. Including CEO age and tenure does not significantly affect the rest of the variances. Finally, marginal likelihood values are almost identical for Models T and H, with the marginal likelihood for Model T equal to −2497.84 and for model H (on the same set of 713 firms) equal to −2497.95.

While using only information on tenure allows us to retain a larger sub-sample, the model produces coefficients that are very similar to those obtained from fitting Model H. The marginal likelihood estimates, however, show that Model H outperforms Model T (while the marginal likelihood for Model H equals −5752.85, for Model T it equals −6044.58). We therefore, do not report the detailed estimates for this model.

5.5. Robustness to prior specifications

To investigate the robustness of our results to the prior, we re-estimate Model H with different prior variances for all the coefficients. Although a virtually uninformative prior variance of 100 does not change the estimated coefficients and their significance, a more restrictive prior variance of 0.1 may have a substantial impact on some variables. Figs. 6–8 compare the effects of setting the prior variance to 10 and 0.1 on the intercept in the second-stage regression for the two coefficients corresponding to CEO share ownership and the coefficient on dollar return volatility. Fig. 8 shows almost no difference in the posterior density of the coefficient on volatility for the two prior specifications. Since neither prior is in conflict with the data, the shape of the posterior density is mainly determined by the data, not the prior. Figs. 6 and 7 show that, when the prior variance is set to 0.1, the posterior density for the coefficients on CEO share ownership changes somewhat. This prior,

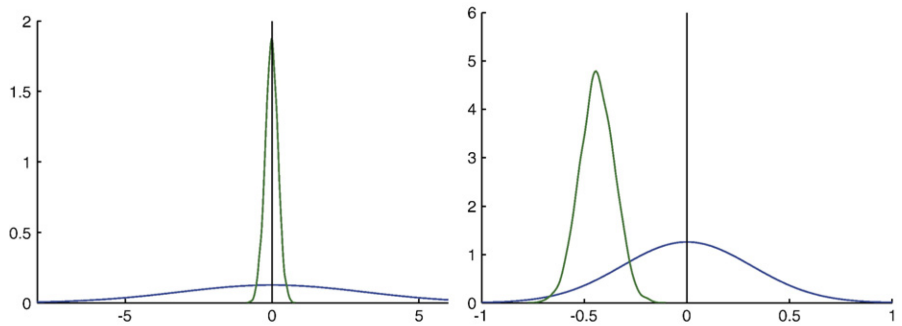

Fig. 7. Prior and Posterior distributions of the coefficient on GI corresponding to CEO share ownership obtained by fitting Model H with prior variance of 10 (left) and 0.1 (right).

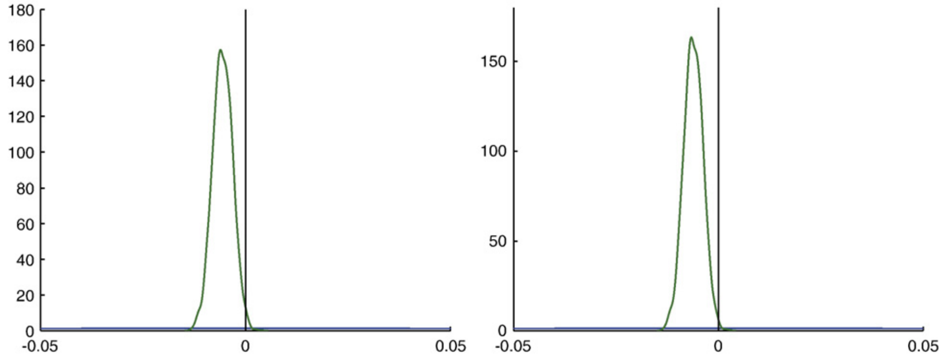

Fig. 8. Prior and Posterior distributions of the coefficient on dollar return volatility obtained by fitting Model H with prior variance of 10 (left) and 0.1 (right).

however, is too restrictive for the coefficients, since it places nearly zero weight on the main part of the posterior for the constant term, obtained with a prior variance of 10.

6. Conclusion

The empirical literature on executive compensation offers ample evidence that CEO compensation practices differ widely across firms and identifies a number of firm and CEO characteristics that explain part of this variation (Milbourn, 2003; Hermalin and Wallace, 2001). We argue that the diversity and inconsistency of empirical results regarding the role of agency theory variables in observed executive compensation practices mainly stems from an incomplete treatment of this heterogeneity in the formulation of the statistical models. The hierarchical Bayesian panel Tobit model that we have specified and compared with alternative models provides evidence that executive compensation policies of large US public corporations are in line with a number of agency theory recommendations. As prescribed by agency theory, in response to increased agency problems – as reflected by CEO share ownership, stock return volatility, relative noise in accounting measures of performance, and CEO employment cycle parameters – firms tend to grant more options to their CEOs. We also find that controlling for the quality of governance helps account for heterogeneity, although a significant portion of it remains unexplained.

It is important to point out that this paper does not test whether larger option awards to CEOs lead to better firm performance. Our results simply indicate that firms seem to respond to agency problems by extending larger option grants to their CEOs. Standard principal-agent models claim that if owners increase option grants to CEOs, it means that they expect option grants to improve firm performance. In practice, however, compensation packages are normally developed by the board of directors, which may also be subject to agency problems. Boards of directors may, for example, decide to increase option grants for strategic reasons, if option grants seem to be popular with shareholders. Additionally, the empirical analysis is complicated by the possible endogeneity of some of the covariates. In this paper, we follow the standard practice of not directly addressing this issue and focus instead on the large heterogeneity in compensation practices across firms.

Acknowledgements

We are grateful to Franz Palm and the anonymous referees for providing valuable comments.

Appendix A

A.1. Estimation

To fit the proposed hierarchical Tobit model we follow Chib (1992) and Albert and Chib (1993) and introduce the latent variables

$$z_{it} = x_{it}'\beta_1 + w_{it}'\tilde{\beta}_i + \varepsilon_{it}$$

where $\epsilon_{it} \sim N(0, \sigma^2)$ and let $y_{it}=\max(z_{it}, 0)$. Now substituting the second-stage equation for $\tilde{\beta}_i$ into Eq. (4) yields

$$z_{it} = x_{1it}'\beta_1 + w_{it}'(A_i\beta_2 + b_i) + \varepsilon_{it} = x_{it}'\beta + w_{it}'b_i + \varepsilon_{it} \tag{5}$$

where $\beta=(\beta_1, \beta_2)$ and $x'_{it}=(x_{1it}', w_{it}' A_i)$.

Let n be the total number of firms in the sample, and n_i be the number of observations on firm i. For the ith firm, let y_{iz} be a $n_i \times 1$ vector with ith component y_{it} if that observation is not censored and z_{it} if it is censored. Then the preceding equation can be written as

$$y_{iz} = X_i\beta + W_ib_i + \varepsilon_i$$

where $X_i=(x_{i1},\ldots,X_{ni})'$, $W_i=(w_{i1},\ldots,w_{ini})'$, and $\epsilon_i=(\epsilon_{i1},\ldots,\epsilon_{ini})'$.

The parameters in this regression are $\psi=(\beta, D, \sigma^2)$. The likelihood function of the parameters is

$$l(\beta, \sigma^2, D) = \prod_{i=1}^{n} (2\pi)^{-n_i/2}|V_i|^{-1/2} \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} e^{(y_{iz}-X_i\beta)V_i^{-1}(y_{iz}-X_i\beta)} \mathrm{d}z_i \tag{6}$$

where $V_i=W_iDW_i'+\sigma^2 I_{ni}$, z_i is a vector composed of all the negative elements of y_{iz}, and I_{ni} denotes $n_i \times n_i$ dimensional identity matrix. The complexity of the likelihood impedes inference by frequentist methods.

In our Bayesian framework, analysis is done by enlarging the parameter space to include both the random effects $\{b_i\}$ and the variables $\{z_i\}$. More formally, we place a prior density $\pi(\psi)$ on the parameters ψ and utilize MCMC methods to summarize the posterior distribution

$$\pi(\psi, \{z_i\}, \{b_i\}|y) \sim \pi(\psi) \prod_{i=1}^{n} N(y_{iz}|X_i\beta + W_ib_i, \sigma^2 I_{n_i}) \prod_{t \in C_i} I\{z_{it}<0\}$$

where y is the observed data, N is the normal density function, $I\{\cdot\}$ is the indicator function and C_i is the index of observations for firm i that are censored.

For the prior density we assume that

$$\begin{aligned} &\beta \backsim N_{k+r}(\beta_0, \beta_0), \\ &\sigma^2 \backsim IG\left(\frac{v_0}{2}, \frac{\delta_0}{2}\right) \\ &D^{-1} \backsim \mathrm{Wishart}_q(\rho_0, R_0), \end{aligned} \tag{7}$$

where β_0 is set to zero, β_0 to a diagonal matrix with 10 on the diagonal, v_0 and δ_0 are calculated from the maximum and minimum of the observed outcomes, ρ_0 is set to equal $q+4$ and R_0 is a $q \times q$-dimensional identity matrix.

To sample the posterior distribution we utilize MCMC methods. Our MCMC algorithm consists of the following steps.

1. Sample

$$\beta|\mathbf{y}_z, \sigma^2, D \backsim N_k(\hat{\beta}, B)$$

where

$$\hat{\beta} = B(B_0^{-1}\beta_0 + \sum_{i=1}^{n} X_i'V_i^{-1}y_{iz}) \qquad B = B = (B_0^{-1} + \sum_{i=1}^{n} X_i'V_i^{-1}X_i)^{-1}$$

2. Sample

$$\begin{aligned} &z_{it}|y_{iz(-t)}, y_{it}, \beta, D \backsim TN_{(-\infty,0])}(\mu_{it}, v_{it}) \text{ if } = y_{it} = 0 \\ &\mu_{it} = E(z_{it}|y_{iz(-t)}, \beta, D) \\ &v_{it} = \mathrm{Var}(z_{it}|y_{it(-t)}, \beta, D) \end{aligned}$$

where TN is the truncated normal density, $y_{iz(-t)}$ is the vector y_{iz} excluding the tth component

3. Sample

$$b_i|y_z,\beta,\sigma^2,D \backsim N_q(\hat{b}_i,D_i)$$

where

$$\hat{b}_i = D_i W_i'(y_{iz} - X_i\beta) \text{ and } \quad D_i = (D^{-1} + W_i'W_i)^{-1}$$

4. Sample

$$D^{-1}|y_z,\beta,\{b_i\},\sigma^2 \backsim W_q\{\rho_0+n,R_n\}$$

where

$$R_n = \left(R_0^{-1} + \sum_{i=1}^{n} b_i b_i'\right)^{-1}$$

5. Sample

$$\sigma^2|y_z,\beta,\{b_i\},D \backsim \mathcal{IG}\left(\frac{v_0+\Sigma n_i}{2},\frac{\delta_0+\delta_n}{2}\right)$$

where

$$\delta_n = \sum_{i=1}^{n} (y_i - X_i\beta - W_i b_i)'(y_i - X_i\beta - W_i b_i)$$

6. Goto 1

A.2. Model comparison

To compare the various alternative model specifications we utilize *Bayes factors*, or ratios of *marginal likelihoods*. The marginal likelihood of a particular model is the normalizing constant of the posterior density and is defined as

$$m(y) = \int l|\psi)\pi(\psi)\mathrm{d}\psi \tag{8}$$

the integral of the likelihood function with respect to the prior density. If we have two models $\mathcal{M}_k$ and $\mathcal{M}_l$, then the Bayes factor is the ratio

$$B_k = \frac{m(y|\mathcal{M}_k)}{m(y|\mathcal{M}_l)} \tag{9}$$

We estimate the marginal likelihood of each model by the method of Chib (1995).

Begin by noting that $m(y)$ by virtue of being the normalizing constant of the posterior density can be expressed as

$$m(y) = \frac{l(\psi^*)\pi(\psi^*)}{\pi(\psi^*|y)}, \tag{10}$$

for any given point θ^* (generally taken to be a high density point such as the posterior mean). Thus, provided we have an estimate $\hat{h}\,(\theta^*|\mathbf{y})$ of the posterior ordinate, the marginal likelihood can be estimated on the log scale as

$$\log m(y) = \log l(\psi^*) + \log\pi(\psi^*) - \log\hat{\pi}(\psi^*|\mathbf{y}), \tag{11}$$

The log-likelihood of the model is estimated at the single point ψ^* by a simulation-based method. The second term is the prior ordinate and is available directly. For the third term, the ordinate of the posterior at ψ^*, we employ the marginal-conditional decomposition

$$\pi(\mathbf{D}^{-1*}, \sigma^{2*}, \beta^*|y) = \pi(D^{-1*}|y)\pi(\sigma^{2*}|y, D^*)\pi(\beta^*|y, D^*, \sigma^{2*}),$$

where the first term is obtained by averaging the Wishart density over draws on $\{z_i\}$ and $\{b_i\}$ from the full run. To estimate the second ordinate, which is conditioned on D^*, we run a reduced MCMC simulation with the full conditional densities

$$\pi(\beta|y_z, D^*, \sigma^2); \pi(\{z_i\}|y, \beta, D^*, \sigma^2); \pi(\sigma^2|y_z, \beta, D^*, \{b_i\}); \pi(\{b_i\}|y_z, \beta, D^*, \sigma^2),$$

where each conditional utilizes the fixed value of $\mathbf{D}$. The second ordinate can now be estimated by averaging the inverse gamma full conditional density of σ^2 at σ^{2*} over the draws on (β, $\{z_i\}$, $\{b_i\}$) from this reduced run. Finally, to estimate the last ordinate we also fix σ^2 at σ^{2*} and continue the reduced runs with the full-conditional densities

$$\pi(\beta|y_z, D^*, \sigma^{2*}); \pi(\{z_i\}|y, \beta, D^*, \sigma^{2*}); \pi(\{b_i\}|y_z, \beta, D^*, \sigma^{2*}),$$

and average the multivariate normal density given in Step 1 of the MCMC algorithm at the point β^*.

References

Aggarwal, R.K., Samwick, A.A., 1999. The other side of the trade-off: the impact of risk on executive compensation. Journal of Political Economy 107 (1), 65–105.

Albert, J.H., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 88, 669–779.

Anderson, R.C., Bizjak, J.M., 2003. An empirical examination of the role of the CEO and the compensation committee in structuring executive pay. Journal of Banking and Finance 27:7, 1323–1348.

Bernardo, J.M., Smith, A.F.M., 1994. Bayesian Theory. John Wiley and Sons, New York.

Bertrand, M., Mullainathan, S., 2001. Are CEOs rewarded for luck? The ones without principles are. Quarterly Journal of Economics 116, 901–932.

Chen, Q., Jiang, W., 2003. Carrer Concerns and the Optimal Pay-for-Performance Sensitivity. Working Paper.

Chib, S., 1992. Bayes inference in the Tobit censored regression model. Journal of Econometrics 51, 79–99.

Chib, S., 1995. Marginal likelihood from the Gibbs output. Journal of the American Statistical Association 90, 1313–1321.

Chib, S., 2001. Markov Chain Monte Carlo methods: computation and inference. In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 5. North Holland, Amsterdam.

Chib, S., Carlin, B., 1999. On MCMC sampling in hierarchical longitudinal models. Statistics and Computing 9, 17–26.

Chib, S., Greenberg, E., 1995. Understanding the Metropolis–Hastings algorithm. American Statistician 49, 327–335.

Chib, S., Greenberg, E., 1996. Markov Chain Monte Carlo simulation methods in econometrics. Econometric Theory 12, 409–431.

Core, J.E., Guay, W.R., 1999. The use of equity grants to manage optimal equity incentive levels. Journal of Accounting and Economics 28, 151–184.

Core, J.E., Guay, W.R., 2002. Estimating the value of employee stock option portfolios and their sensitivity to stock prices and volatility. Journal of Accounting Research 40 (3), 613–630.

Core, J.E., Guay, W.R., Larcker, D.F., 2003. Executive equity compensation and incentives: a survey. Federal Reserve Bank of New York Economic Policy Review 9 (1), 27–50.

Core, J.E., Holthausen, R.W., Larcker, D.F., 1999. Corporate governance, chief executive officer compensation, and firm performance. Journal of Financial Economics 51, 371–406.

Gaver, J.J., Gaver, K.M., 1993. Additional evidence on the association between the investment opportunity set and corporate financing, dividend, and compensation policies. Journal of Accounting and Economics 16, 1125–1160.

Gompers, P.A., Ishii, J.L., Metrick, A., 2003. Corporate governance and equity prices. Quaterly Journal of Economics 118 (1), 107–155.

Hermalin, B.E., Wallace, N.E., 2001. Firm performance and executive compensation in the savings and loan industry. Journal of Financial Economics 61, 139–170.

Himmelberg, C.P., Hubbard, R.G., Palia, D.N., 1999. Understanding the determinants of managerial ownership and the link between ownership and performance. Journal of Financial Economics 53, 353–384.

Holmstrom, B., Milgrom, P., 1987. Aggregation and linearity in the provision of intertemporal incentives. Econometrica 55, 303–328.

Ju, N., Leland, H., Senbet, L.W., 2003. Options, Option Repricing and Severance Packages in Managerial Compensation: Their Effects on Corporate Risk. Working Paper.

Kaplan, S.N., Stromberg, P., 2002. Characteristics, Contracts and Actions: Evidence from Venture Capitalist Analyses. Working Paper.

Milbourn, T., 2003. CEO reputation and stock-based compensation. Journal of Financial Economics 68, 233–262.

Ofek, E., Yermack, D., 2000. Taking stock: equity-based compensation and the evolution of managerial ownership. Journal of Finance 55, 1367–1384.

Palia, D., 2001. The endogeneity of managerial compensation in firm valuation: a solution. The Review of Financial Studies 14 (3), 735–764.

Ross, S.A., 2004. Compensation, incentives, and the duality of risk aversion and riskiness. The Journal of Finance 59 (1), 207–225.

Smith Jr., C.W., Watts, R.L., 1992. The investment opportunity set and corporate financing, dividend, and compensation policies. Journal of Financial Economics 32, 262–292.

Yermack, D., 1995. Do corporations award CEO stock options effectively? Journal of Financial Economics 39, 237–269.

Yermack, D., 1996. Higher valuation of companies with small boards of directors. Journal of Financial Economics 40, 185–212.

Zellner, A., 1971. Introduction to Bayesian Inference in Econometrics. John Wiley and Sons, New York.