

Optimal CEO Compensation and Stock Options

Arantxa Jarque*
University of Rochester

Job Market Paper

Abstract

We study the incentive problem between CEO's and the owners of a firm due to the unobservability of the manager's actions. Our model departs from the rest of the literature in two ways. First, we acknowledge that, in contrast with the standard repeated moral hazard case, actions taken by CEO's have a persistent effect in time. Second, in our model the effect of effort on stock prices is derived from primitives; i.e., effort affects directly the conditional distribution of profits, and not the distribution of prices. The stock market determines the price of the stock of the firm using information about past profits.

In this setup, we study the properties of a stylized compensation package that includes stock options, trying to capture the main features of real life executive compensation packages. We find a numerical solution of the contract and parameter values under which the decentralization of the compensation through stock is a good approximation of the optimal contract with unrestricted contingent consumption.

Key Words. Moral Hazard, Optimal Contracts, Persistence, CEO Compensation, Stock Options

JEL Codes: D21, D82, G32

*Department of Economics, University of Rochester, Rochester, NY 14627. Email: arll@troi.cc.rochester.edu. Special thanks to my advisor Hugo Hopenhayn for encouragement and insightful conversations. I would also like to thank Marina Azzimonti, Juan Carlos Hatchondo, Leonardo Martínez, Josef Perktold and participants at Rochester and UCLA seminars for very useful suggestions and discussions. I gratefully acknowledge financial support from the Banco de España.

1 Introduction

Compensation of CEO's with stock options is a widely spread practice nowadays. This paper analyzes the performance of such executive compensation packages relative to an unconstrained scheme that prescribes sequences of contingent consumption. The model used for the analysis differs from the standard moral hazard models used in the literature in that the type of actions that CEO's are asked to exert have long lasting consequences: they affect the profit of the firm for many periods after they are implemented. Persistence is modeled in this paper as follows: the CEO takes a single action at the beginning of the contractual relationship, which determines the conditional distribution of the profit of the firm in each of the following periods until the end of the contract. The second feature of the model is that the effect of the action of the manager on the distribution over stock prices of the firm is not assumed directly, but it is derived from more fundamental structure: effort affects the profits of the firm, and the stock market rationally determines prices based on the history of profits.

The purpose of this paper is to analyze formally how well contracts that specify compensation contingent on the performance of the firm can be approximated through the use of stock options. Most observed executive compensation packages have four components: a base salary, an annual bonus tied to accounting performance, stock options, and long term incentive plans. As documented in [10], in the mid 90's stock options represented a 17 to 36% of total executive compensation. Bonuses represented 19 to a 26% of compensation, while the base salary proportion ranged from a 21 to 40%. Restricting the sample to companies with sales above the median, options represented a 39% of the total compensation, while salary was 24%.¹ Given their empirical relevance, this paper focuses on options as the main instrument for executive compensation. In the stylized model presented here, the manager receives a base salary every period until he is allowed to exercise the options granted to him at the beginning of the contract. The price and time of exercise are optimally chosen by the stockholders, and commitment is assumed on both parts.

In contrast with the previous literature, we provide a structural model based on primitives that illustrates explicitly the effect that the action of

¹The data cited in Murphy (1999) includes all companies in the S&P 500, based on ExecuComp data.

the manager has on the price of the stock of the firm.^{2,3} The probability distribution over the profit of the firm is affected both by the effort of the manager and by the quality of the technology of the firm, an unknown parameter that the manager cannot control. Buyers in the stock market price a share of the firm as the expected stream of future profits. In order to calculate the expectation, they incorporate all past public information. More precisely, they use the history of profit realizations to update their priors over the technology parameter. The market understands the incentives given by the firm to its manager, and buyers correctly condition their beliefs on the equilibrium action taken by the CEO. Given this, any variation in the price of the stock comes from learning about the quality of the firm's technology and not about the action taken by the manager. However, in the presence of learning, an off the equilibrium path change in the effort of the manager will affect the market's posterior about the firm's type, and thus its market value. Payment schemes that include options exploit the effect that poor outcomes have on the price of the firm: lower effort increases the probability of poor outcomes, which will be interpreted by the market as a sign of a bad quality firm and will bring down the stock price, lowering the profits to the manager from selling the stock options. In order to provide a benchmark to evaluate the performance of option contracts we solve for the best incentive compatible contract that specifies a consumption for the CEO contingent on the performance of the firm, the Second Best contract. Our goal is to find parameter values under which the indirect link between effort and prices exploited by options schemes is sufficient to implement the right action of the principal with a similar cost to that of the Second Best scheme.

We solve for the optimal option contract using numerical methods. Qualitatively, two properties of the option scheme stand out: first, punishments to the manager are bounded below, suggesting evidence of limited liability for the managers. Second, compensation is always monotonic in the profit of

²Aleff and Santos (2002) perform a similar analysis using a reduced form model: effort has an effect on the distribution of prices of the firm, as opposed to having an effect on output and then letting the market price the stock, as it is modeled in the present paper. This reduced form prevents them from pointing out the differences between a compensation scheme with options and one with contingent consumption in which this paper will focus.

³Clementi, Cooley and Wang (2003) propose a justification of the use of stock options as a commitment device that increases efficiency. In their model, the manager receives a portion of the profit of the firm as payment, so there is a direct effect of effort on the price of the stock.

the firm. A complete characterization of the Second Best contract assuming limited liability is given, and the cost of the two contracts is compared.

2 The Model

We model the moral hazard problem existing between the owners of a firm and the agent they hire to perform as the CEO of the company. The stockholders of the firm act as a unique risk neutral principal when designing the contract. The manager is assumed to be risk averse, with a strictly concave utility function $u(\cdot)$ with $u'(0) = \infty$. He has an outside opportunity with an expected utility of \underline{U} at period zero. The contract lasts for an infinite number of periods, i.e., the agent receives payments from the firm during his whole lifetime. In any period profits can take two values, $y_L = 0$, and $y_H = 1$. Both the level of effort of the manager e , and the quality of the technology of the firm θ , affect the probability of the outcome. Effort can take two values, $e_L < e_H$, which also capture the disutility that the manager experiences exerting these levels of effort is higher than that of the low effort. We assume that the parameters are such that the firm owners always want to implement e_H . The type of the firm can take two values: a “good” one G , and a “bad” one B . The stockholders and the CEO have the same prior q_0 about $\theta = G$. New observations are used to update this prior according to Bayes’ rule. The posteriors are denoted $q(y^t)$ if the action is e_H and $\hat{q}(y^t)$ if the action is e_L .

The effect of any action taken by the manager of a firm decreases as time passes. In order to capture this feature, the probability of a high level of profits at period t when the CEO chooses e_H is determined as follows:

$$\begin{aligned} \Pr_t(y_t|y^{t-1}) &= \alpha_t [q(y^{t-1}) \pi_G + (1 - q(y^{t-1})) \pi_B] \\ &\quad + (1 - \alpha_t) [q(y^{t-1}) \bar{\pi}_G + (1 - q(y^{t-1})) \bar{\pi}_B], \end{aligned}$$

where α_t is the rate at which the effect of effort decreases, π_θ is the probability of y_H induced by the high effort when the firm is of type $\theta = G, B$, and $\bar{\pi}_\theta$ represents the probability that drives the profits of the firm when the effect of effort dies out. The Bayesian updating is done in the usual way:

$$q(y^t) = \frac{q(y^{t-1}) [\alpha_t \pi_G + (1 - \alpha_t) \bar{\pi}_G]}{q(y^{t-1}) [\alpha_t \pi_G + (1 - \alpha_t) \bar{\pi}_G] + (1 - q(y^{t-1})) [\alpha_t \pi_B + (1 - \alpha_t) \bar{\pi}_B]}.$$

If the manager chooses e_L instead, the probability will be:

$$\begin{aligned}\widehat{\text{Pr}}_t(y_t|y^{t-1}) &= \alpha_t [\widehat{q}(y^t) \widehat{\pi}_G + (1 - \widehat{q}(y^t)) \widehat{\pi}_B] \\ &\quad + (1 - \alpha_t) [\widehat{q}(y^t) \overline{\pi}_G + (1 - \widehat{q}(y^t)) \overline{\pi}_B],\end{aligned}$$

with $\pi_G > \widehat{\pi}_G$ and $\pi_B > \widehat{\pi}_B$, and

$$\widehat{q}(y^t) = \frac{\widehat{q}(y^{t-1}) [\alpha_t \pi_G + (1 - \alpha_t) \overline{\pi}_G]}{\widehat{q}(y^{t-1}) [\alpha_t \pi_G + (1 - \alpha_t) \overline{\pi}_G] + (1 - \widehat{q}(y^{t-1})) [\alpha_t \pi_B + (1 - \alpha_t) \overline{\pi}_B]}.$$

Typically, $\{\alpha_t\}_{t=1}^T$ will be a decreasing sequence. Whenever $\alpha = 1$, profits are driven by:

	G	B
e_L	$\widehat{\pi}_G$	$\widehat{\pi}_B$
e_H	π_G	π_B

We assume that $0 < \pi_\theta < 1$ and $0 < \widehat{\pi}_\theta < 1$ for both $\theta = G, B$. When $\alpha = 0$, the probability of the high level of output is $\overline{\pi}_G$ if the firm is of the good type, $\overline{\pi}_B$ if it is of the bad one. We assume, for tractability, that $\alpha_t = \alpha^t$, where α is a constant number between zero and one.

Given these assumptions about the stochastic structure, we can construct the probability of a given history y^T :

$$\text{Pr}(y^T) = \prod_{t=1}^T \text{Pr}_t(y_t|y^{t-1}).$$

The probability of each outcome depends on time through the α 's, and on the history through the bayesian updating of the probability of the good technology.

3 Options scheme

We now introduce a stylized model that tries to capture the main features of typical executive compensation schemes. An Option Scheme is a vector:

$$\{T, c_0, n, p_{ex}\},$$

where T is the time at which options are exercised, n is the number of stock options granted, p_{ex} is the price at which the stock can be bought if

the options are exercised, and c_0 is the constant consumption for the CEO for every period before the time at which options are exercised. This scheme is a major simplification in several ways. First, it allows for only one exercise time, while in reality we observe that some proportions of the same stock options award become available for exercise at different times. Second, the base wage could be thought as changing over time, since stock grants are usually complemented with bonus plans based on accounting measures. Also, some part of the variable compensation is given in restricted stock (shares that the CEO has to hold until a pre specified time); these are just stock options with strike price equal to zero, which are available instruments in the model presented.

In order for options to be used as an incentive mechanism, the price sequence for the stock of the firm $\{p_t\}$ must be sensitive to profit realizations. In our model, this dependence comes from learning. Investors, the third player of the game, are trying to learn about the quality of the technology of the firm, so that they can accurately price the expected future stream of profits. Given the stochastic structure defined in the previous section, and for a given interest rate r , we can calculate the price of the stock:

$$p_t(y^t) = E_t \left[\sum_{\tau=t}^{\infty} y_{\tau} \left(\frac{1}{1+r} \right)^{\tau-t} \right]. \quad (1)$$

The timing of the game is as follows: first, the owners of the firm (O) decide on the Options Contract. Nature (N) decides whether the firm is good or bad, and without knowing the outcome of such randomization the manager (M) decides his level of effort. The level of profits of the firm is realized, conditional on the choice of effort, and the Stockholders (S) price the stock based on their bayesian updating of the quality of the firm (see Fig. 1).

A Perfect Bayesian Equilibrium of this game will be an Option Contract $\{T, c_0, n, p_{ex}\}$, a level of effort of the agent equal to e_H , a pricing function $p(y^T)$ and beliefs of the Stockholders consistent with the equilibrium choice of effort. Since the probability of observing any history is positive under the equilibrium level of effort, Bayesian updating provides consistent beliefs and no refinement is necessary.

Note that, if there was no extra noise in the outcome of the firm, a payment with options would not be feasible; the price of the stock would be constant and independent of the effort level of the CEO. In our set up with learning we explicitly make the distribution over stock prices depend on the

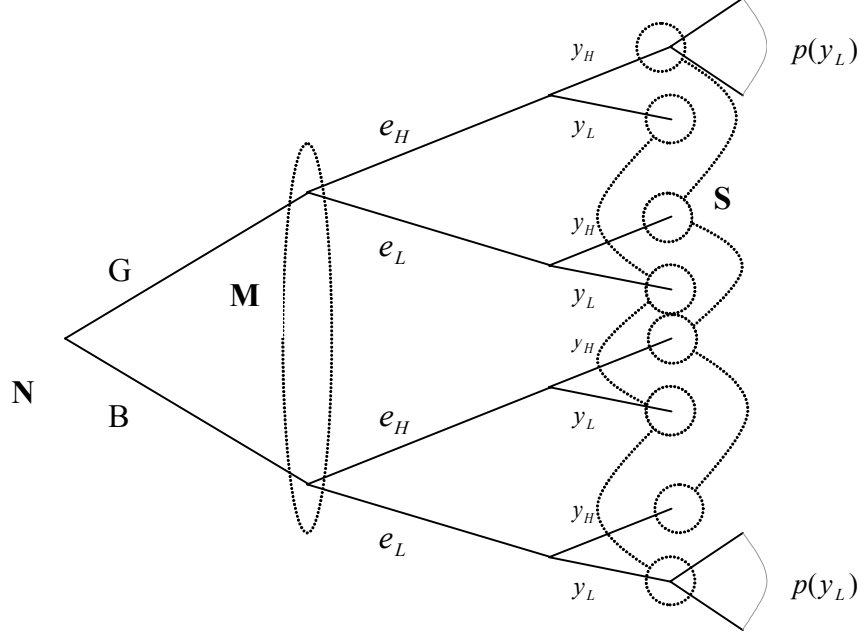


Figure 1: Game tree for a given Options Contract.

recommended level of effort: when the CEO is considering a deviation he understands that prices are determined under the equilibrium beliefs of the market, so the reduced form probability vector over prices will depend on the recommended effort:

$$\Pr(p(y^t) | e; e_H).$$

The problem of the stockholders is to choose the options scheme that minimizes the cost of implementing the high level of effort. Let

$$\Gamma = \{y^T \mid p(y^T) > p_{ex}\}$$

be the set of histories of outcomes for which the market price is higher than the exercise price; i.e., the set of histories for which the option is exercised.

Define

$$\begin{aligned}\Pr(\Gamma) &= \sum_{y^T \in \Gamma} \Pr(y^T), \\ \widehat{\Pr}(\Gamma) &= \sum_{y^T \in \Gamma} \widehat{\Pr}(y^T).\end{aligned}$$

For a given exercise time T , and for a given options scheme the cost to the principal is:

$$V(\underline{U}, T, c_0, n, p_{ex}) = \frac{1}{1-\beta} c_0 + \beta^{T-1} n \sum_{y^T \in \Gamma} (p(y^T) - p_{ex}) \Pr(y^T).$$

The problem above will have the two usual constraints. First, the Participation Constraint, which makes sure that the agent gets as much utility in the contract as in his outside option:

$$\begin{aligned}\underline{U} &= \frac{1-\beta^{T-1}}{1-\beta} u(c_0) + \beta^{T-1} (1 - \Pr(\Gamma)) \frac{u(c_0)}{1-\beta} \\ &\quad + \beta^{T-1} \sum_{y^T \in \Gamma} \frac{u(c_0 + n[p(y^T) - p_{ex}](1-\beta))}{(1-\beta)} \Pr(y^T) - e_H.\end{aligned}$$

Up to the time of exercise, the CEO will receive a payment of c_0 . At T , if the market price of the stock $p(y^T)$ is bigger than p_{ex} the manager will exercise the n options that he was granted, thus making a profit of $n[p(y^T) - p_{ex}]$. We are assuming that the agent cannot borrow against his future income, but we allow him to save; at time T he will use the savings technology to smooth his consumption in time. His per period consumption after the exercise of the options amounts to $c_0 + \max\{0, n[p(y^T) - p_{ex}](1-\beta)\}$.

Second, since effort is unobservable, the Incentive Constraint has to be satisfied; the agent should get at least as much utility from taking the recommended level of effort than from the alternative:

$$\begin{aligned}
& \frac{1 - \beta^{T-1}}{1 - \beta} u(c_0) + \beta^{T-1} (1 - \Pr(\Gamma)) \frac{u(c_0)}{1 - \beta} \\
& + \beta^{T-1} \sum_{y^T \in \Gamma} \frac{u(c_0 + n[p(y^T) - p_{ex}](1 - \beta))}{(1 - \beta)} \Pr(y^T) - e_H \\
\geq & \frac{1 - \beta^{T-1}}{1 - \beta} u(c_0) + \beta^{T-1} (1 - \widehat{\Pr}(\Gamma)) \frac{u(c_0)}{1 - \beta} \\
& + \beta^{T-1} \sum_{y^T \in \Gamma} \frac{u(n[p(y^T) - p](1 - \beta))}{(1 - \beta)} \widehat{\Pr}(y^T) - e_L
\end{aligned}$$

Under the deviation effort, the probability of falling into the set Γ will be lower, but the set will not change precisely because the market uses the equilibrium action to calculate prices. The manager can exploit the market's equilibrium beliefs.

Finding the optimal options scheme is a fairly complicated problem that cannot be solved analytically. Some properties of the optimal scheme can be derived, however.

Remark 1 *The optimal scheme will always have a positive number of stock options and the exercise price will be such that the options will be exercised after at least one of the possible histories of profits.*

This property comes straight out of the need of satisfying the Incentive Constraint.

Remark 2 *The consumption of the manager is bounded below, since he always has the right not to exercise the option if the stock price is below the exercise price.*

This is in fact an important limitation for the design of the contract. It may be understood as an assumption of limited liability of the CEO, and it implies that punishments are bounded.

Proposition 1 In an Option Scheme, the utility of the CEO is always weakly monotonic in the profit of the firm; i.e., after a given history y^t , whenever the observation at time $t + 1$ is a high outcome, the manager is given higher utility than when the realization is low.

Proof See Appendix.

4 Comparison to the Second Best

There exists a contingent wage profile that would implement the high level of effort at less or at least equal cost than the option scheme. In this section we characterize this contract, which will be referred to as the second best. A numerical simulation of the first and the second best, along with the optimal option scheme, should give us an idea of the relative performance of the stock options. Furthermore, we can compare the properties of both schemes and learn about the circumstances under which the use of options constitutes a good approximation to an optimal contract.

The optimal contract defines an unrestricted sequence of contingent consumption that minimizes the cost of implementing the high level of effort. In this benchmark model we try to incorporate as many realistic features as possible, in order to make set up the most favorable scenario for the stock option scheme. On one hand, since the scheme with options puts a lower bound on the available punishments to the agent, here we assume that there exists a minimum utility level b , that needs to be guaranteed to the agent. Moreover, we incorporate as a new parameter the number of periods for which information about the profits of the firm is available, T . This, of course, is meant to have a similar role to the exercise time in the stock option scheme. This parameter is exogenous to the problem of designing the optimal contract, but it can be easily endogenized once the optimal consumption scheme for each T is known.

Let $c_\tau(y^\tau)$ be the consumption levels delivered to the agent contingent on the realization of the profits of the firm, and $\omega(y^t)$ the level of constant consumption that he will get from time $T + 1$ on. We assume that any information revealed after T is ignored when designing the contract. The cost of the principal is:

$$V\left(\underline{U}, \{c_\tau(y^\tau)\}_{\tau=0}^T, \omega(y^T)\right) = \sum_{t=0}^T \sum_{y^t} \beta^t \{c_t(y^t)\} \Pr(y^t) + \beta^T \sum_{y^T} \frac{\omega(y^T)}{1-\beta} \Pr(y^T)$$

The CEO is assumed to get payments from the firm all his life, which is infinite. His Participation Constraint is:

$$\underline{U} \leq \sum_{t=0}^T \sum_{y^t} \beta^t u(c_t(y^t)) \Pr(y^t) + \beta^T \sum_{y^T} \frac{u(\omega(y^T))}{1-\beta} \Pr(y^T) - e_H, \quad (\text{PC})$$

where the payment $\omega(\cdot)$ received from period T onwards depends only on the history of profits up to time T . The Incentive Constraint needs to be satisfied:

$$\begin{aligned} & \sum_{t=0}^T \sum_{y^t} \beta^t u(c_t(y^t)) \Pr(y^t) + \beta^T \sum_{y^T} \frac{u(\omega(y^T))}{1-\beta} \Pr(y^T) - e_H \\ & \geq \sum_{t=0}^T \sum_{y^t} \beta^t u(c_t(y^t)) \widehat{\Pr}(y^t) + \beta^T \sum_{y^T} \frac{u(\omega(y^T))}{1-\beta} \widehat{\Pr}(y^T) - e_L. \quad (\text{IC}) \end{aligned}$$

We also impose the limited liability constraint:

$$c_t(y_t) \geq b \quad \forall y^t. \quad (\text{LL})$$

Proposition 2 The optimal sequence of contingent consumption $\{c_\tau(y^\tau)\}_{\tau=0}^T$ in the Second Best contract will be ranked according to the likelihood ratios of the histories of profit realizations up to time T .

Proof See Appendix.

The ordering of the payments is determined by the ordering of the likelihood ratios of the histories of outcomes. Consider the case in which the effect of effort does not depreciate in time (i.e., $\alpha = 1$). In our two-outcome world, the number of times that the high outcome is realized, call it x , contains all the information needed to determine consumption following a given history. At any particular t , any two histories y^t, \tilde{y}^t with the same x will be assigned the same consumption, that is $c_t(y^t) = c_t(\tilde{y}^t) = c_t(x)$. Fig. 1 illustrates the case of perfectly persistent actions: in the first period, $c(y_H) > c(y_L)$. After a high realization in the first period, the variability in consumption in the second period is smaller than the one following a low profit. The optimal contract assigns a much lower consumption to the history (y_L, y_L) , since it is relatively more likely under the low level of effort. The solution to the contract will concentrate punishments in the poor performance histories; from the first order condition of consumption we can see that for a common pair of multipliers, and provided the limited liability constraint does not bind, histories with very high likelihood ratios will have lower consumption.

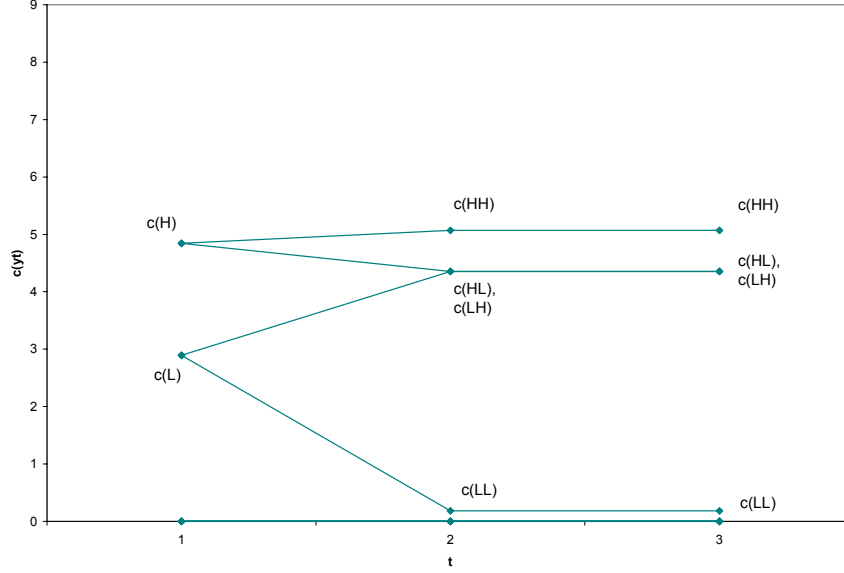


Fig. 1. Consumption under the Second Best Scheme ($\alpha = 1$).

For any $\alpha < 1$, realizations in the early periods will be more valuable information than later ones (see Fig. 2): consumption after (y_H, y_L) is higher than after (y_L, y_H) . As T increases, and provided α_t does not become very small, histories of outcomes become more informative; this implies that incentives can be better tailored, so longer contracts will rely more on later periods for incentive purposes and their cost will be lower than for shorter contracts. Presumably, options schemes with longer waiting periods will try to take advantage of this increase in the quality of information. In order to get some intuition about the trade off implicit in the setting of the exercise time, we look now at the implications of changes in T for the optimal consumption paths.

Conjecture 1 As T increases, dispersion in consumption decreases; variation in consumption is concentrated in later periods, following relatively unlikely histories of profits.

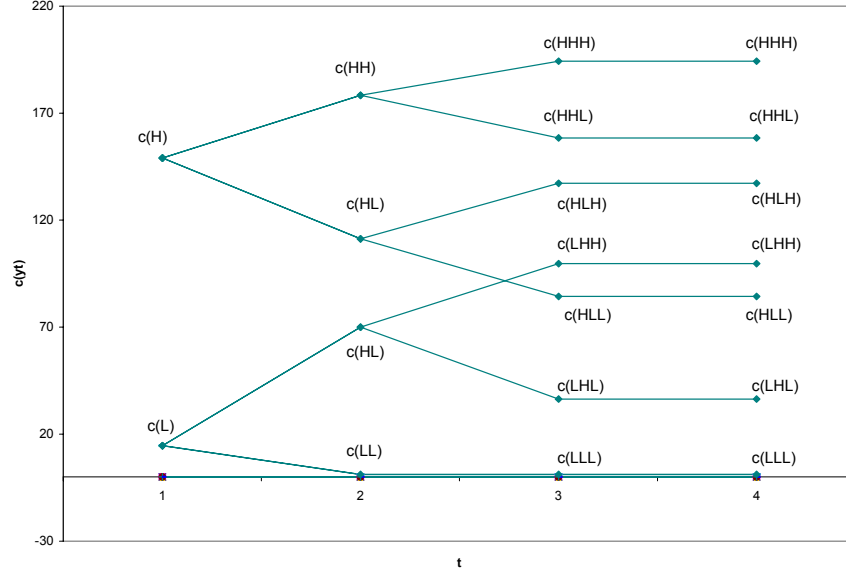


Fig. 2 Consumption in the Second Best with $\alpha < 1$.

As more periods of observations are added to the contract, the space of signals becomes richer; longer histories of profits have more extreme values for their likelihoods under the two possible levels of effort. In particular, the likelihood of bad histories under high effort will decrease faster than their likelihood under low effort. These very high likelihood ratios imply low consumption. The histories that have high likelihood ratios do not have a high probability of occurring under the equilibrium effort level; time makes available a cheap punishment scheme, so the cost of the contract decreases.

Conjecture 2 The cost of the Second Best contract weakly decreases when T increases (strictly if $\alpha = 1$, or if the rate at which the sequence $\{\alpha_t\}$ decreases is smaller than the speed of convergence by the LLN)

If α_t decreases very fast, later outcomes do not contain much information: the likelihood ratios do not diverge and consumption dispersion does not decrease. Intuitively, when the effect of effort on profits dies out rapidly, there is not as much value in using later information to decide the consumption of the CEO.

Conjecture 3 If the limited liability constraint is more binding, the cost of the contract increases and consumption is more volatile.

Forcing consumption to be above a certain threshold has a similar effect to that of having α_t decrease. Any history of outcomes that would have been assigned a consumption equal or lower to the minimum will now be pooled together and assigned the same consumption. This is a waste of information that will increase the cost of the contract. In order to satisfy the IC, rewards will be used after good histories. Using big rewards instead of harsh punishments is more costly for the principal, since rewards are tied to histories of profits that have a higher probability of occurring under the equilibrium action, while histories with punishments are less likely under a high effort than under a low one. Rewards introduce bigger distortions in the consumption smoothing of the agent, since the jumps in consumption occur with high probability.

Another qualitative property of the model is the possibility of non-monotonicities. Given our assumptions about the probability of the profits, the Monotone Likelihood Ratio Property (MLRP) holds. In standard moral hazard problems, this assumption is enough to guarantee monotonicity of consumption in the profit.

Proposition 3 In the first period of the contract, compensation is always monotonic in the profit.

Proof See Appendix.

For later periods monotonicity does not hold, even under the MLRP. This is pointing out a limitation of the option scheme over the contingent consumption schedule.

Proposition 4 Optimal consumption will not necessarily be monotonic in profit.

Proof See Appendix.

The existence of non-monotonicities comes from combining learning with the provision of incentives. Since the quality of the technology of the firm is not controllable by the CEO, ideally the contract would tend to insure him against this risk. However, under such a contract the manager will tend to

shirk and blame poor performance on a bad technology. The optimal contract demands exposing the agent to some technology related risk, and it can derive into non monotonicities of consumption. The owners of the firm evaluate the relative likelihood of effort and learn about the quality of the technology at the same time. This is what introduces the non-monotonicities in the optimal contract. A low outcome will sometimes be stronger evidence of a high level of effort than a high outcome; this will occur in cases in which there is more precise learning about the technology under the low effort assumption than under the high one.

5 Relative Performance of the Option Scheme

Having established qualitative properties of the Option Scheme and the Second Best contract, we now proceed to report the preliminary quantitative findings of the numerical solution of both contracts.

In order to have a benchmark in terms of cost, we look at the unconstrained optimal contract achievable when the principal can observe the effort level chosen by the agent. In this First Best contract, the manager is given a constant consumption each period that gives him a present discounted value of utility equal to his outside option, \underline{U} . This is the way of optimally sharing the risk in the production technology, given that we assumed the owners of the firm to be risk neutral and the CEO to be risk averse. The optimal level of consumption in this First Best contract is determined by:

$$\underline{U} + v(e_H) = \frac{1 - \beta^{T+1}}{1 - \beta} u(c).$$

We can calculate numerically the cost under observable effort, and compare the relative cost of both the Second Best and the Option Scheme.

From the preliminary computational work, we can already extract some conjectures about the properties of the optimal Option Scheme.

Conjecture 4 The cost of the optimal options scheme decreases when α increases

An Options Scheme bases all the incentives exclusively on the time of exercise. If α_{Tex} is very small, the effect of the action of the CEO in $p(y^T)$ is very small, and incentives become very expensive. In Table 1 we report the

relative cost of the different schemes with respect to the cost of the First Best contract. The Second Best scheme always performs better when information is available for three periods. Also, the cost is much closer to that of the First Best than when using options. The scheme with options does better for higher persistence of the effort.

			α			
	0.4	0.5	0.6	0.7	0.8	0.9
Second Best (1 per)	1.102	1.102	1.102	1.102	1.102	1.102
Second Best (2 per)	1.045	1.042	1.038	1.035	1.031	1.028
Second Best (3 per)	1.045	1.040	1.036	1.031	1.026	1.021
Options Contract	1.559	1.546	1.544	1.543	1.543	1.543

Table 1. Cost of Contract / Cost of First Best

Our conclusion from the numerical examples is that compensation with options performs best as an approximation of the Second Best when the discount factor of the CEO is the highest and effort is persistent.

6 Using Multiple Option Grants

The analysis in the previous sections was carried assuming a single grant of stock options was granted to the CEO. This assumption is obviously not realistic and is raising the cost of the option schemes, since there is a limit to how well they can approximate the Second Best contract. In fact, if there would be a big enough number of available grants, at different exercise prices and times, the executive compensation packages could reproduce any contract that used arbitrary consumption transfers.

Define a Multiple Stock Options Scheme as a tuple

$$\langle \{n_i, pex_i\}_{i=1}^{t+1} \rangle_{t=1}^T,$$

where $\{n_i\}_{i=1}^{t+1}$ is the number of options with exercise price pex_i that will be available to exercise at time t .

Proposition 5 Whenever $y^t \neq y^{t'}$ implies $p(y^t) \neq p(y^{t'})$ for all y^t , the second best contingent wage contract can be implemented through a multiple stock options scheme.

Proof See Appendix.

Real-life option schemes do not include as much variety of options as the above proposition suggests that is necessary to implement the second best correctly. We do observe, however, changes in constant salaries and sequential award of different options with several vesting times. The next step in the paper is to try to introduce these variations into the model. Some simulations will then be used to assess the costs of the modified options scheme.

7 Conclusion

This paper studies the form of contracts used to solve the incentive problem between CEO's and the owners of a firm due to the unobservability of the manager's actions. We provide an innovative framework to evaluate the decentralization of CEO compensation through stock; we model the long lasting effects of CEO's actions and we derive the effect of effort on stock prices from primitives. Effort affects directly the conditional distribution of profits, and not the distribution of prices.

In this setup we are able to analyze the extent to which option schemes can be a good approximation to an optimal contract with unrestricted contingent consumption. In order for options to be used as an incentive mechanism, the price sequence for the stock of the firm $\{p_t\}$ must be sensitive to profit realizations, and thus to the effort of the CEO. In our model, we assume that buyers in the stock market understand that the owners of the firm design the compensation of the CEO in order to provide him with incentives to choose the right effort; under this assumption, the dependence of prices on profits comes only from learning about the quality of the technology of the firm through time. This points out the main characteristic of compensation schemes that include stock options: the link between effort and payoff is indirect and can make the use of options a more expensive way of providing incentives.

In our preliminary numerical exercises we find that the cost of choosing option schemes over unrestricted contracts diminishes with the level of persistence of the CEO's actions. Using the comparison with the unrestricted optimal contract, future work will explore the effect of changes in limited liability levels, as well as that of differences in technologies across firms and industries on the performance of stock option schemes.

8 Appendix

Proof of Proposition 1 Given the assumptions about the conditional distribution of profit, a good outcome is always stronger evidence of a good technology, i.e., $q(y^t, y_H) > q(y^t, y_L)$. From eq. 1,

$$\begin{aligned} p(y^t, y_H) - p(y^t, y_L) &= \frac{1+r}{1+r+\alpha} [q_t(y^t, y_H) - q_t(y^t, y_L)] [(\pi_G - \bar{\pi}_G) - (\pi_B - \bar{\pi}_B)] \\ &\quad + \frac{1+r}{r} [q_t(y^t, y_H) - q_t(y^t, y_L)] (\bar{\pi}_G - \bar{\pi}_B) > 0. \end{aligned}$$

This implies that market prices will always be monotonic in profit: $p(y^t, y_H) > p(y^t, y_L)$. From the PC of the problem, the utility of the CEO is an increasing function of market prices.

Proof of Proposition 2 Using the Kuhn-Tucker theorem, the first order conditions with respect to $c_t(y^t)$ will be:

$$\frac{1}{u'(c_t(y^t))} = \lambda + \mu \left[1 - \frac{\widehat{\text{Pr}}(y^t)}{\text{Pr}(y^t)} \right] + \gamma(y^t) \frac{1}{\text{Pr}(y^t)} \quad \forall y^t.$$

Also, $\omega(y^t) = c_t(y^T)$, since the corresponding first order condition will be

$$(\omega(y^t)) : \frac{1}{u'(c_t(y^T))} = \lambda + \mu \left[1 - \frac{\widehat{\text{Pr}}(y^T)}{\text{Pr}(y^T)} \right] + \gamma(y^{T+1}) \frac{1}{\text{Pr}(y^T)}.$$

Whenever $\gamma(y^t) > 0$, we know that the limited liability constraint holds, so $c_t(y^t) = b$. For histories such that $\gamma(y^t) = 0$,

$$\frac{1}{u'(c_t(y^t))} = \lambda + \mu \left[1 - \frac{\widehat{\text{Pr}}(y^t)}{\text{Pr}(y^t)} \right]$$

holds and by strict concavity of the utility function a higher value for the likelihood ratio $\frac{\widehat{\text{Pr}}(y^t)}{\text{Pr}(y^t)}$ implies lower consumption.

Proof. The first order conditions of the problem say:

$$c(y_i) = \lambda + \mu(1 - LR(y_i)) \quad i = L, H$$

We have that

$$LR(y_L) = \frac{q_0(1 - \hat{\pi}_G) + (1 - q_0)(1 - \hat{\pi}_B)}{q_0(1 - \pi_G) + (1 - q_0)(1 - \pi_B)}$$

$$LR(y_H) = \frac{q_0\hat{\pi}_G + (1 - q_0)\hat{\pi}_B}{q_0\pi_G + (1 - q_0)\pi_B}$$

Since

$$\begin{aligned}\pi_G &> \hat{\pi}_G \\ \pi_B &> \hat{\pi}_B,\end{aligned}$$

we find that, as in the standard moral hazard problem, consumption in the first period is monotonic in the outcome:

$$LR(y_L) > LR(y_H) \Rightarrow c(y_L) < c(y_H).$$

■

Proof of Proposition 4 (Generalization of Miller 1999) The difference $c(y^t, y_H) - c(y^t, y_L)$ will have the same sign as $LR(y^t, y_L) - LR(y^t, y_H)$. To simplify the notation, let π and $\hat{\pi}$ be the probability of the high outcome after history y^t under the high and low effort respectively. Then :

$$\begin{aligned}LR(y^t, y_L) - LR(y^t, y_H) &= \frac{\hat{\pi}}{\pi} - \frac{1 - \hat{\pi}}{1 - \pi} \\ &= \frac{\pi - \hat{\pi}}{\pi(1 - \pi)}.\end{aligned}$$

So,

$$\text{sign}(c(y^t, y_H) - c(y^t, y_L)) = \text{sign}(\pi - \hat{\pi}).$$

$$\begin{aligned}\pi - \hat{\pi} &= [q(y^t)(\alpha^t\pi_G + (1 - \alpha^t)\bar{\pi}_G) + (1 - q(y^t))(\alpha^t\pi_B + (1 - \alpha^t)\bar{\pi}_B)] \\ &\quad - [\hat{q}(y^t)(\alpha^t\hat{\pi}_G + (1 - \alpha^t)\bar{\pi}_G) + (1 - \hat{q}(y^t))(\alpha^t\hat{\pi}_B + (1 - \alpha^t)\bar{\pi}_B)]\end{aligned}$$

or, rearranging terms,

$$\begin{aligned}\pi - \hat{\pi} &= \alpha^t [q(y^t)\pi_G + (1 - q(y^t))\pi_B] \\ &\quad + (1 - \alpha^t) [q(y^t)\bar{\pi}_G + (1 - q(y^t))\bar{\pi}_B] \\ &\quad - \alpha^t [\hat{q}(y^t)\hat{\pi}_G + (1 - \hat{q}(y^t))\hat{\pi}_B] \\ &\quad + (1 - \alpha^t) [\hat{q}(y^t)\bar{\pi}_G + (1 - \hat{q}(y^t))\bar{\pi}_B]\end{aligned}$$

We can add and subtract the term $\alpha^t [q(y^t) \hat{\pi}_G + (1 - q(y^t)) \hat{\pi}_B]$:

$$\begin{aligned} \pi - \hat{\pi} &= \alpha^t [q(y^t) \pi_G + (1 - q(y^t)) \pi_B] \\ &\quad + (1 - \alpha^t) [q(y^t) \bar{\pi}_G + (1 - q(y^t)) \bar{\pi}_B] \\ &\quad - \alpha^t [\hat{q}(y^t) \hat{\pi}_G + (1 - \hat{q}(y^t)) \hat{\pi}_B] \\ &\quad + (1 - \alpha^t) [\hat{q}(y^t) \bar{\pi}_G + (1 - \hat{q}(y^t)) \bar{\pi}_B] \\ &\quad + \alpha^t [q(y^t) \hat{\pi}_G + (1 - q(y^t)) \hat{\pi}_B] \\ &\quad - \alpha^t [q(y^t) \hat{\pi}_G + (1 - q(y^t)) \hat{\pi}_B] \end{aligned}$$

and rearrange:

$$\begin{aligned} \pi - \hat{\pi} &= \alpha^t [q(y^t) (\pi_G - \hat{\pi}_G) + (1 - q(y^t)) (\pi_B - \hat{\pi}_B)] \\ &\quad + (\hat{\pi}_G - \hat{\pi}_B) (q(y^t) - \hat{q}(y^t)) \\ &\quad + (1 - \alpha^t) (q(y^t) - \hat{q}(y^t)) (\bar{\pi}_G - \bar{\pi}_B) \end{aligned}$$

If $q(y^t) > \hat{q}(y^t)$, compensation will be higher after a good observation.
If $q(y^t) < \hat{q}(y^t)$, then we have to check the conditions under which the effect of learning will overcome the standard MLRP effect.

Proof of Proposition 5 The Second Best contingent consumption scheme can be thought of as a set of vectors, one for each period:

$$\vec{c}_1 = \begin{bmatrix} c_1(0) \\ c_1(1) \end{bmatrix}, \quad \vec{c}_2 = \begin{bmatrix} c_2(0,0) \\ c_2(0,1) \\ c_2(1,0) \\ c_2(1,1) \end{bmatrix}, \quad \dots, \quad \vec{c}_t = \begin{bmatrix} c_t(0, \dots, 0) \\ \vdots \\ c_t(y^t) \\ \vdots \\ c_t(1, \dots, 1) \end{bmatrix}$$

The dimension of the consumption vector at each t equals 2^t ; let this dimension be denoted by $d(t)$. We need the Option Scheme to span any point in $\mathbb{R}^{d(t)}$. Denote as $w_t(\cdot)$ the consumption of the CEO under the Options Scheme at time t and following a history with a given y^t . It will be given by:

$$w_t(y^t) = \sum_{i=1}^{d(t)} \max \{n_{t,i} [p(y^t) - pex_i], 0\}.$$

The set of vectors of consumption spannable with the multiple options scheme will be:

$$\begin{bmatrix} w_{t,1} & \dots & w_{t,t+1} \end{bmatrix} = \begin{bmatrix} n_{t,1} & \dots & n_{t,t+1} \end{bmatrix} \cdot \begin{bmatrix} \max \{ (p(y_1^t) - \text{pex}_{t,1}), 0 \} & \dots & \max \{ (p(y_1^{t+1}) - \text{pex}_{t,1}), 0 \} \\ \vdots & \ddots & \vdots \\ \max \{ (p(y_{d(t)}^1) - \text{pex}_{t,t+1}), 0 \} & \dots & \max \{ (p(y_{d(t)}^{t+1}) - \text{pex}_{t,d(t)}), 0 \} \end{bmatrix}.$$

For the result to hold, we need the rank of the matrix of payoffs equal to $t + 1$. Order the market prices from lowest to biggest, where $p_j < p_{j+1}$.

Let $\text{pex}_{t,1} = 0$, and $\text{pex}_{t,i} = \frac{p(y_{i+1}^t) - p(y_i^t)}{2}$. The resulting matrix will be

$$\begin{bmatrix} p_1 & p_2 & p_3 & \dots & p_d \\ 0 & \frac{p_2 - p_1}{2} & p_3 - \frac{p_1 + p_2}{2} & \dots & p_d - \frac{p_1 + p_2}{2} \\ 0 & 0 & \frac{p_3 - p_2}{2} & \dots & p_d - \frac{p_3 - p_2}{2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{p_d - p_{d-1}}{2} \end{bmatrix},$$

which is diagonal and thus of full rank. We can always find $(n_{t,1}, \dots, n_{t,t+1})_{t=1}^T$ so that $\vec{w}_t = \vec{c}_t \forall t$.

References

- [1] Aseff, Jorge and Manuel S. Santos. “Stock Options and Managerial Optimal Contracts”, Working Paper (2002)
- [2] Clementi, Gian Luca, Thomas F. Cooley and Cheng Wang. “Stock Grants as a Commitment Device”, Working Paper (2003)
- [3] Grossman, Sanford and Oliver D. Hart. “An Analysis of the Principal–Agent Problem.” *Econometrica* 51, Issue 1 (Jan.,1983), 7-46.
- [4] Holmström, B. and P. Milgrom. “Aggregation and Linearity in the Provision of Intertemporal Incentives”, *Econometrica*, 55, 303-328. (1987)
- [5] Holmström, B. “Managerial Incentive Problems: a Dynamic Perspective.” *Review of Economic Studies*, 66, 169-182. (1999)

- [6] Jensen, M. C. and K.J. Murphy. "Performance-Pay and Top-Management Incentives", *Journal of Political Economy*, 98, 225-264. (1990)
- [7] Kole, S.R. "The Complexity of Compensation Contracts." *Journal of Financial Economics*, 43, 79-104. (1997)
- [8] Miller, Nolan. "Moral Hazard with Persistence and Learning", Manuscript. (1999)
- [9] Mirrlees, James. "Notes on Welfare Economics, Information and Uncertainty", in M. Balch, D. McFadden, and S. Wu (Eds.), *Essays In Economic Behaviour under Uncertainty*, pgs.. 243-258. (1974)
- [10] Murphy, Kevin J. "Executive Compensation", in *Handbook of Labor Economics*, Volume 3, Edited by O. Ashenfelter and D. Card, 1999
- [11] Wang, C. "Incentives, CEO compensation, and Shareholder Wealth in a Dynamic Agency Model," *Journal of Economic Theory*, 76, 72-105 (1997)
- [12] Yermak, D. "Do Corporations Award CEO Stock Options Effectively?" *Journal of Financial Economics*, 39, 237-269. (1995)