DYNAMIC MORAL HAZARD AND EXECUTIVE STOCK OPTIONS

BAOMIN DONG  Henan University
GUIXIA GUO  University of International Business and Economics
FRANK YONG WANG*  University of International Business and Economics

Abstract. This paper shows that the optimal executive compensation scheme in a dynamic moral hazard environment is convex in the firm value. This implies that the optimal contract should include stock options. This is because the private benefit of shirking is increasing in firm value and the manager’s utility is concave. Therefore, in contrast to the previous literature that takes stock options in the incentive contract exogenously, we rationalize the optimality of their use endogenously. Moreover, we show that the optimal amount of stock options (restricted stocks) increases with agency cost and the executive’s reservation utility, decreases with the degree of risk aversion of the manager, and increases (decreases) with the firm size.

JEL Classification: C73, D86, G34

1. INTRODUCTION

Executive stock options, as a long-term incentive mechanism, give executives the right to buy their firm’s shares at a prespecified strike price. Over the past two decades, the use of stock options has represented a major change in the corporate governance structure (Holmstrom and Kaplan, 2001). Stock options are also one of the most frequently used tools in executive compensation schemes (Hall and Liebman, 1998; Murphy, 1999; Hall and Murphy, 2000, 2002, 2003).

However, with the increasingly extensive use of stock options and with the rising proportion of option value in executives’ compensation, heated debate has emerged over the ‘generosity’ of executive stock options, both in academia and industry. After the global financial crisis in 2008, several financial institutions, including those who received government bailouts, continued to provide fat bonuses for their executives. Whether such huge bonuses are worthwhile has been discussed extensively in the literature. For instance, Meulbroek (2001) and

*Address for Correspondence: School of International Trade and Economics, University of International Business and Economics, Beijing 100029, China. E-mail: wangyong9110@gmail.com. The authors are grateful to an anonymous referee for his insightful comments on an earlier version of the paper. Guixia Guo is grateful for financial support from the 2011 National Social Science Foundation Major Project (Grant no. 11&ZD007), the 2012 National Natural Science Funds of China for Young Scholars (Grant no. 71203026), the Humanities and Social Science Research Projects of the MOE in China (Grant no. 12YJC790047), Program for Innovative Research Team and the ‘211 Program’ in UIBE. Frank Yong Wang is grateful for financial support from the Program for Innovative Research Team and from the ‘211 Program’ at the University of International Business and Economics.

© 2013 Wiley Publishing Asia Pty Ltd
Hall and Murphy (2000, 2002) claim that firms have to compensate their risk-averse executives with more stock options, because the executives cannot choose efficient investment portfolios and, thus, significantly undervalue the stock options. However, Lambert and Larcker (2004) point out that such analysis is incomplete, because the incentive role of stock options is neglected.

The essence of the debate is whether executives should be granted stock options and, if so, what is the optimal amount? No consensus has been reached. Haugen and Senbet (1981) are among the first to propose stock options as an instrument to alleviate the conflict between shareholders and managers, but their focus is on executives’ excessive incentive of entrenchment due to the issuance of new shares. Therefore, their discussion is more about the issue of capital structure. In a setting with hyperbolic absolute risk aversion utility and gamma distribution functions, Hemmer et al. (2000) show that stock options are a component of the optimal compensation contract only if the executive’s absolute coefficient of risk aversion is decreasing and his or her relative coefficient of risk aversion is moderate. However, using constant relative risk aversion (CRRA) utility and lognormal distribution functions, the simulation results of Dittmann and Maug (2007) predict that the optimal compensation scheme that does not include stock options can realize the same level of incentive and executives’ utility at 20% less cost.

In our opinion, the long-run or dynamic nature of moral hazard of the executives is key for executive stock options plans. Thus, in the current paper, we use a dynamic or multi-period moral hazard framework to study the optimal compensation contract. This is in contrast to the aforementioned papers, which use a static setup in which before the output realization the manager can make his or her effort decision only once. The dynamic model is better suited for reality and tailored to address the incentive role of stock options in the long run.

Using a simple binomial output process, we show that the optimal compensation scheme is convex in the firm value. This naturally implies that stock options should be included in the optimal contract. There are two factors contributing to the convexity result. The first is that the private benefit of shirking is increasing in the firm value. Hence, the optimal contract should pay higher at the margin for a given increase in output. In addition, because the manager is risk averse, his or her marginal utility for a given increase in payment is smaller, and, hence, the firm needs to pay more to offset a given amount of private benefit derived from shirking. Note that the fixed payment plus a linear piece rate contract prescribed by the traditional principal–agent theory only yields a linear contract. Given that stock options are financial assets with a convex payoff structure in reality, they should be included in the optimal compensation contract.

In contrast to the previous literature that exogenously assumes that the compensation contract consists of a base wage, restricted stocks and stock options (e.g. Lambert and Larcker, 2004; Dittmann and Maug, 2007; Armstrong et al.,

1 When granting stock options and restricted stock to executives, there is generally a vesting period during which the options and stocks cannot be sold out.
2007), our model justifies the optimality of the widespread use of executive stock options endogenously. This result is robust even if we extend our model into infinite periods.

The rest of our paper is organized as follows. Section 2 presents the basic model. Section 3 studies the optimal compensation contract in a two-period moral hazard setting and Section 4 extends the basic model into more periods to investigate the convexity of the optimal contract in a general setting. Section 5 presents the findings of comparative static analysis and Section 6 concludes.

2. THE BASIC MODEL

A risk-neutral shareholder (the principal) employs a risk-averse manager (the agent) to run a project, whose initial value is $S > 0$. The project’s future value evolves following a binomial tree: if the project’s present value is $\tilde{S}$, its value in the next period can be either $u\tilde{S}$ (success) with probability $p \in (0, 1)$ or $d\tilde{S}$ (failure) with probability $1 - p$, where $0 < d < 1 < u$. The risk-free interest rate is normalized to 0. The project is subject to a moral hazard problem, because the manager’s effort is unobservable and it can affect the probability of success. For the validity of the first-order approach (FOA), we assume there are only two effort levels of the manager: $e = 1$ (work) or $e = 0$ (shirk). Suppose the manager’s utility function is additively separable: $v(x) + B(e, \tilde{S})$, where $v(\cdot)$ is the utility of wealth, with $v’ > 0$, $v'' < 0$, and $B(e, \tilde{S})$ is the manager’s private benefit under different effort levels $e$ and the firm value $\tilde{S}$.

Suppose the manager’s effort is independent of the firm value in the following sense:

**ASSUMPTION 1.** The probability of success is independent of the firm size, only determined by the manager’s effort; the manager’s private benefit is proportional to the firm value.

Specifically, when the firm value is $\tilde{S} > 0$, the probability of success (failure, respectively) is $p_H (1 - p_H)$ if the manager chooses to work, and his or her private benefit is normalized to 0; the probability of success (failure, respectively) is $p_L (1 - p_L)$ if the manager chooses to shirk, and the private benefit is $b\tilde{S}$, where $p_L < p_H$ and $b > 0$ is a constant. For simplicity, suppose that there is no discount across different periods, and the firm has positive net present value (NPV) if the manager works; that is, $p_H u + (1 - p_H) d > 1$. We also suppose that it is socially desirable for the manager to work rather than to shirk; that is, $p_H u + (1 - p_H) d \geq p_L u + (1 - p_L) d + b$.

---

2 In the principal–agent literature, there is always the validity problem of FOA, from Holmstrom (1979) to Grossman and Hart (1983), Rogerson (1985), Jewitt (1988), and, more recently, Ke (2012). Generally, we can just provide some sufficient conditions for the validity of FOA, such as the convex distribution function condition and monotone likelihood ratio property, but those sufficient conditions are often too stringent to be satisfied. The FOA problem is even more severe when stock options (which are convex in the firm value) are included in the compensation package, because, in this case, the objective function of the agent becomes a non-concave function of his or her action. For detailed discussion, interested readers may refer to Hemmer et al. (2000), Lambert and Larcker (2004) and Armstrong et al. (2007).
In the basic model, we consider a two-period moral hazard problem: the manager works for the shareholder for two periods before the project value is realized. Similar to Holmstrom and Milgrom (1987), we assume the manager gets paid according to the final project value, and the manager can observe the project value in each period. More specifically, at $t=1$, the shareholder proposes a compensation scheme $\{R_0, R_1, R_2\}$ to the manager, and advises the desirable effort level. $R_i (i=0, 1, 2)$ is the compensation for the manager when the project succeeds $i$ times. If the manager rejects the offer, the game ends; if he or she accepts the offer, he or she chooses the effort level at $t=1$ and $t=2$. At $t=3$, the final project value is realized and the manager is paid accordingly. Denote $U_i = v(R_i)$, and $h(x) = v^{-1}(x)$. Figure 1 depicts the timeline of the game.

3. THE OPTIMAL COMPENSATION SCHEME

Because the project has positive NPV only when the manager works, without loss of generality, we assume the principal always prefers the manager to work than to shirk. Therefore, the shareholder’s problem is to design the compensation scheme $\{R_0, R_1, R_2\}$ to maximize his or her expected profit subject to a series of constraints:\(^3\)

In the Appendix, we show that all the incentive constraints can be reduced to inequalities 3 and 4. Intuitively, if the manager chooses to work, the expected project value is higher in the next period given the assumption $p_{H} + (1 - p_H)d > 1$, and, thus, the private benefit for the manager to shirk currently is smaller than to shirk in the next period. Therefore, as long as the manager has no incentive to shirk in the last period, he or she will not do so before that. This also holds true in the extended model in the next section.

\(^3\) In the Appendix, we show that all the incentive constraints can be reduced to inequalities 3 and 4.
\[
\max_{R_l, R_H, R_0} p_H^2 (u^2 S - R_2) + 2 p_H (1 - p_H) (udS - R_1) + (1 - p_H)^2 (d^2 S - R_0),
\]

subject to
\[
\begin{align*}
\bar{U} &\equiv p_H U_2 + (1 - p_H) U_1 \geq p_H U_2 + (1 - p_H) U_1 + b u S, & (IC_{21}) \\
U &\equiv p_H U_1 + (1 - p_H) U_0 \geq p_H U_1 + (1 - p_H) U_0 + b d S, & (IC_{22}) \\
p_H \bar{U} + (1 - p_H) U \geq p_H^2 U_2 + 2 p_H (1 - p_H) U_1 + (1 - p_H)^2 U_0 \geq U_R. & (IR)
\end{align*}
\]

where \( \bar{U} \) (\( U \), respectively) is the expected utility level when the manager works and the project value is \( uS \) (\( dS \), respectively) at \( t = 2 \), and \( U_R \) is his or her reservation utility. \( IC_{21} \) and \( IC_{22} \) are the incentive compatibility constraints at \( t = 2 \) when the project value is \( uS \) and \( dS \), respectively, and \( IR \) is the participation constraint for the manager at \( t = 1 \).

Denote \( \Delta p \equiv p_H - p_L \). Because firm’s expected value at \( t = 2 \) is a constant when the manager works in both periods, the principal’s problem is equivalent to minimizing his or her (expected) incentive cost \( p_H^3 R_0 + 2 p_H (1 - p_H) R_1 + (1 - p_H)^2 R_0 \). By replacing \( U_i (i = 0, 1, 2) \) with \( h(R) \), the original programming (eqn 1) can be simplified as follows:

\[
\begin{align*}
\min_{U_0 \geq U_1 \geq U_2} &\quad p_H^2 h(U_2) + 2 p_H (1 - p_H) h(U_1) + (1 - p_H)^2 h(U_0), \\
\text{s.t.} &\quad (\Delta p) (U_2 - U_1) \geq b u S, & (IC_{21}) \\
&\quad (\Delta p) (U_1 - U_0) \geq b d S, & (IC_{22}) \\
&\quad p_H^2 U_2 + 2 p_H (1 - p_H) U_1 + (1 - p_H)^2 U_0 \geq U_R. & (IR)
\end{align*}
\]

It is not difficult to show that the participation constraint (5) must be binding at the optimum. Otherwise, the principal can always reduce \( U_2, U_1, U_0 \) by a sufficiently small amount \( \varepsilon > 0 \), while keeping both incentive compatibility constraints unaltered, and, in this way, the principal is strictly better off.

In Lemma 1, we show that the incentive compatibility constraints must also be binding at the optimum, with the proof relegated to Appendix B.

**Lemma 1.** Both incentive compatibility constraints (3) and (4) are binding at the optimum.

Intuitively, for the optimal contract \( \{U_2^*, U_1^*, U_0^*\} \), because \( U_1^* - U_0^* = \frac{b d S}{\Delta p} \) and \( U_2^* - U_1^* = \frac{b u S}{\Delta p} \) are sufficient to motivate the manager to work, there is no need for the principal to pay the manager any more bonus. If otherwise, consider another contract \( \{\hat{U}_2, \hat{U}_1, \hat{U}_0\} \) where \( \hat{U}_1 - \hat{U}_0 \geq \frac{b d S}{\Delta p} \), \( \hat{U}_2 - \hat{U}_1 \geq \frac{b u S}{\Delta p} \), and one of the two inequalities strictly holds. Because the participation constraint is always binding at the optimum, contract \( \{\hat{U}_2, \hat{U}_1, \hat{U}_0\} \) has the same mean (under the binomial probability density \( \{p_H^2, 2 p_H (1 - p_H), (1 - p_H)^2\} \) as the optimal contract.
contract. One can readily show that \( \hat{U}_2 > U^*_2 \), \( \hat{U}_0 < U^*_0 \), implying that contract \( \{\hat{U}_2, \hat{U}_1, \hat{U}_0\} \) can be considered as the mean-preserving spread of the original contract \( \{U^*_2, U^*_1, U^*_0\} \), and is, thus, riskier than the optimal contract.

Recall that the objective function is \( E[h(x)] \), where \( E[\cdot] \) is the expectation operator with the probability density \( \{p_h, 2p_h(1-p_h), (1-p_h)^2\} \) and \( h(x) \) is the inverse of the utility function and is, thus, convex. Therefore, it has a larger mean under a riskier distribution. Remember that the principal aim is to minimize the incentive cost, so contract \( \{\hat{U}_2, \hat{U}_1, \hat{U}_0\} \) is dominated by the optimal contract \( \{U^*_2, U^*_1, U^*_0\} \). Therefore, the incentive compatibility constraints are binding at the optimum.

With the help of Lemma 1, there are three unknowns \( \{U_2, U_1, U_0\} \) and three linear equations, and we are able to derive the optimal contract, as summarized in Proposition 1.

**Proposition 1.** The optimal compensation scheme consists of a base wage, restricted stocks and stock options.

**Proof.** From Lemma 1, we have

\[
R_0^* = h(U_0^*) = h \left( U_R - \frac{bS}{\Delta p} A \right),
\]

\[
R_1^* = h(U_1^*) = h \left( U_R - \frac{bS}{\Delta p} A + \frac{bdS}{\Delta p} \right),
\]

\[
R_2^* = h(U_2^*) = h \left( U_R - \frac{bS}{\Delta p} A + \frac{(u+d)S}{\Delta p} \right),
\]

where \( A \equiv p_h(u+d)+2p_h(1-p_h)d \) is a constant.

By definition, \( R_2^* \), \( R_1^* \) and \( R_0^* \) are the payoffs in an optimal compensation when the project value is \( u^2S \), \( udS \) and \( d^2S \), respectively. Therefore, \( R_2^* - R_2 = h(U_2^*) - h(U_1^*) = h'(\xi_2)(U_2^*-U_1^*) = h'(\xi_2)\frac{bS}{\Delta p} = h(U_1^*)-h(U_0^*) = h'(\xi_0)\frac{bdS}{\Delta p} \), where \( U_0^* < \xi_0 < U_1^* < \xi_2 < U_2^* \). By the monotonicity and convexity of function \( h(\cdot) \), we have \( h'(\xi_2) > h'(\xi_0) > 0 \), and, thus,

\[
\frac{R_2^* - R_1^*}{u(u-d)S} = h'(\xi_2) \frac{b}{\Delta p(u-d)} > \frac{R_1^* - R_0^*}{d(u-d)S} = h'(\xi_0) \frac{b}{\Delta p(u-d)},
\]

from which we obtain

\[
\frac{R_2^* - R_1^*}{u^2S-udS} > \frac{R_1^* - R_0^*}{udS-d^2S} > 0.
\]
Because \( \frac{R_2^* - R_0^*}{u^2 S - u d S} \) and \( \frac{R_0^* - R_0^*}{u d S - d^2 S} \) are the piece rates when the firm value increases from \( d^2 S \) to \( u d S \), and from \( u d S \) to \( u^2 S \), respectively, equation 7 implies that the piece rate increases in the firm value, and, therefore, the optimal compensation scheme is convex. To implement this convex contract, it is natural to include stock options in the optimal contract.

Proposition 1 delivers an important message; that is, the optimal compensation scheme is a convex function of the firm value (see Fig. 2). Therefore, a compensation scheme consisting solely of a base wage and restricted stocks is suboptimal, unless restricted stock is considered as a special sort of stock option with zero exercise price.

To characterize the optimal amount of stock options, rewrite the optimal compensation contract in the following way:

\[
w(x) = \beta_0 + \beta_1 x + \beta_2 (x - k)^+, \quad x \in \{u^2 S, u d S, d^2 S\},
\]

where \( \beta_0 \) is the base salary; \( \beta_1 x \) is the value of restricted stocks, with \( \beta_1 \) as its amount; and \( \beta_2 (x - k)^+ \) is the value of stock options, where \( k \) is the exercise price and \( \beta_2 \) is the amount of stock options. However, the expression of the optimal convex compensation scheme becomes much steeper.

\( R_2^* \)

\( R_1^* \)

\( R_0^* \)

Figure 2. The optimal compensation scheme in the two-period model

It has to be pointed out that convexity of the optimal compensation scheme is not equivalent to \( U_2 - U_1 \geq U_1 - U_0 \); it is equivalent to the monotonicity of the piece rate; that is, the agent is paid more marginally when the firm value grows, and in Figure 2, this implies that the compensation scheme becomes much steeper.
contract is not unique. Because stock options are generally granted at-the-money in practice, it is natural to set the exercise price at $k = S$. Then, there are two cases to discuss:

**Case 1.** $ud \leq 1$. In this case, the parameters in equation 8 should satisfy

$$
\begin{align*}
\beta_0 + \beta_1 \times d^2 S + \beta_2 \times 0 &= R_0^*, \\
\beta_0 + \beta_1 \times uS + \beta_2 \times 0 &= R_1^*, \\
\beta_0 + \beta_1 \times u^2 S + \beta_2 \times (u^2 S - S) &= R_2^*,
\end{align*}
$$

from which

$$
\beta_0 = \frac{uR_0^* - dR_1^*}{u - d}, \quad \beta_1 = \frac{R_0^* - R_1^*}{udS - d^2 S} > 0, \quad \beta_2 = \frac{u}{u^2 - 1} \left( \frac{R_2^* - R_1^* - R_0^*}{us} - \frac{R_1^* - R_0^*}{ds} \right) > 0,
$$

and the convexity in equation 7 is applied to establish the inequalities.

**Case 2.** $ud > 1$. Similarly, in this case, the parameters in equation 8 should satisfy

$$
\begin{align*}
\hat{\beta}_0 + \hat{\beta}_1 \times d^2 S + \hat{\beta}_2 \times 0 &= R_0^*, \\
\hat{\beta}_0 + \hat{\beta}_1 \times uS + \hat{\beta}_2 \times (uS - S) &= R_1^*, \\
\hat{\beta}_0 + \hat{\beta}_1 \times u^2 S + \hat{\beta}_2 \times (u^2 S - S) &= R_2^*,
\end{align*}
$$

Again, we have $\hat{\beta}_2 = \frac{d}{1 - d^2} \left( \frac{R_2^* - R_1^*}{us} - \frac{R_1^* - R_0^*}{ds} \right) > 0$.

Therefore, in both cases, a strictly positive amount of stock options should be included in the optimal contract. This finding is in contrast to that of Dittmann and Maug (2007), among others. Intuitively interpreted, this is due to the fact that shareholders generally prefer a ‘carrot’ to ‘stick’; that is, the corporations offer their managers fat bonuses when they make a profits but no punishment when they make a loss. Theoretically speaking, the convexity of optimal contracts is ascribed to the fact that the manager’s private benefit from shirking increases linearly with the firm value. To meet the incentive compatibility constraints, the manager’s ex post utility level should also increase at least linearly. Due to the risk aversion (or, equivalently, decreasing marginal utility) of the manager, the compensation then must increase faster than the firm value at the margin.

A caveat one should bear in mind is that the convexity of the optimal contract does not arise in the one-period static models even if there are three or more output states. This is because without firm value accumulation in the interim
stage, a linear contract can always replicate the optimum. This holds even in
the repeated moral hazard settings as long as outputs are independent across
periods. For example, in such a repeated moral hazard principal–agent model,
Holmstrom and Milgrom (1987) show that the optimal contract is always linear.

The key difference that drives the results is that in Holmstrom and Milgrom
(1987), the outputs of the agent in each period are independent, while in our
model the final output follows a binomial process. Thus, their model is a
repeated game and is basically static, whereas ours is dynamic. In particular,
because our assumption resembles the long-run nature of the principal–agent
relationship, our prediction is consistent with the fact that executive stock
options are generally used to motivate the executives in the long run.5

4. AN EXTENSION TO MULTIPLE PERIODS

In this section, we extend the basic model into three periods, so that the game
in Section 2 is extended for \( t = 3 \) (see Figure 3 for the timeline). We keep
Assumption 1 in the extended model. The principal offers a compensation
scheme \( \{R_0, R_1, R_2, R_3\} \), where \( R_i(i = 0, 1, 2, 3) \) is the manager’s repayment when
the project succeeds \( i \) times at the end of the game.

5 It is common practice that executive stock options cannot be exercised in 10 to 20 years.
The shareholder maximizes his or her expected profit subject to a series of constraints:\footnote{Similarly, all the incentive constraints can be reduced into IC$_{31}$, IC$_{32}$ and IC$_{33}$, or equations 11, 12 and 13.}

$$\max_{R_1, R_2, R_3, R_0} p_1^3(u^3 S - R_1) + 3 p_1^3(1 - p_H) (u^3 dS - R_2)$$
$$+ 3 p_H (1 - p_H)^2 (u d^2 S - R_1) + (1 - p_H)^3 (d^3 S - R_0),$$

s.t. $U_{3H} \equiv p_H U_3 + (1 - p_H) U_2 \geq p_L U_3 + (1 - p_L) U_2 + b u^2 S$ \hspace{1cm} (IC$_{31}$),
$$U_{3M} \equiv p_H U_2 + (1 - p_H) U_1 \geq p_L U_2 + (1 - p_L) U_1 + b u^2 S$ \hspace{1cm} (IC$_{32}$),
$$U_{3L} \equiv p_H U_1 + (1 - p_H) U_0 \geq p_L U_1 + (1 - p_L) U_0 + b d^2 S \hspace{1cm} (IC_{33})$$
$$p_H U_{3H} + 2 p_H (1 - p_H) U_{3M} + (1 - p_H)^2 U_{3L} \geq U_R \hspace{1cm} (IR').$$

where IC$_{31}$, IC$_{32}$ and IC$_{33}$ are the incentive compatibility constraints at $t = 3$ when the project value is $u^3 S$, $u d S$ and $d^3 S$, respectively, and IR is the participation constraint for the manager at $t = 1$.

Applying the $\Delta p \equiv p_H - p_L$ and substituting $\{U_0, U_1, U_2, U_3\}$ in the objective function and constraints, the optimization, equation 9, can be reduced as follows:

$$\min_{U_0, U_1, U_2, U_3} p_1^3 h(U_3) + 3 p_1^3(1 - p_H) h(U_2) + 3 p_H (1 - p_H)^2 h(U_1) + (1 - p_H)^3 h(U_0),$$

s.t. $(\Delta p) (U_3 - U_2) \geq b u^2 S$ \hspace{1cm} (IC$_{31}$)
$$(\Delta p) (U_2 - U_1) \geq b u d S \hspace{1cm} (IC_{32})$$
$$(\Delta p) (U_1 - U_0) \geq b d^2 S \hspace{1cm} (IC_{33})$$
$$p_H U_{3H} + 2 p_H (1 - p_H) U_{3M} + (1 - p_H)^2 U_{3L} \geq U_R \hspace{1cm} (IR').$$

Applying similar logic to the basic model, the participation constraint 14 must be binding. Similarly, the incentive compatibility constraints 11, 12 and 13 are all binding, the proof of which is similar to that shown in Appendix B, and is omitted here. Therefore, the optimal compensation scheme can be solved as:

$$R_0^* = h(U_0^*) = h\left(U_R - \frac{b S}{\Delta p} A\right),$$
$$R_1^* = h(U_1^*) = h\left(U_R - \frac{b S}{\Delta p} A + \frac{b S}{\Delta p} d^2\right).$$

© 2013 Wiley Publishing Asia Pty Ltd
\[ R_i^* = h(U_i^*) = h \left( U_R - \frac{bS}{\Delta p} \hat{A} + \frac{bS}{\Delta p} d(u + d) \right), \]

\[ R_3^* = h(U_3^*) = h \left( U_R - \frac{bS}{\Delta p} \hat{A} + \frac{bS}{\Delta p} (u^2 + ud + d^2) \right), \]

where \( \hat{A} = p_{it}(u^2 + ud + d^2) + 3p_{it}^2(1-p_H) d(u+d) + 3p_H(1-p_H)^2 d^2 \) is another constant.

Similar to the two-period model, we have

\[ \frac{R_3^* - R_0^*}{u^3(u-d)S} = h'(\vartheta_3) \frac{b}{\Delta p(u-d)} > h'(\vartheta_2) \frac{b}{\Delta p(u-d)} = \frac{R_2^* - R_0^*}{ud(u-d)S}, \]

such that

\[ \frac{R_3^* - R_2^*}{u^3S - u^2dS} > \frac{R_2^* - R_1^*}{u^2dS - ud^2S} > \frac{R_1^* - R_0^*}{ud^2S - d^3S} > 0, \tag{15} \]

where \( R_0^* < \vartheta_1 < R_1^* < \vartheta_2 < R_2^* < \vartheta_3 < R_3^*. \)

Intuitively, \( \frac{R_i^* - R_{i-1}^*}{u^i d^{i-1}S - u^{i-1} d^{i-1}S} \) is the piece rate when the firm value increases from \( u^i d^{i-1}S \) to \( u^{i-1} d^{i-1}S(i = 1, 2, 3) \), and equation 15 implies that at the optimum, the piece rate should increase in the firm value, which, again, implies the convexity of the optimal compensation scheme. The convexity is depicted in Figure 4. Thus, a simple linear contract consisting of only a base wage and restricted stocks fails to replicate the optimum.

A restricted stock and two stock option packs with different exercise prices can be used to implement the optimal contract here. Similar to the basic model, the optimal compensation scheme as a function of firm value can be expressed as:

\[ w(x) = \beta_0 + \beta_1 x + \beta_2 (x - k_1)^* + \beta_3 (x - k_2)^*, \quad x \in \{u^3S, u^2dS, ud^2S, d^3S\}, \tag{16} \]

where \( k_1 = ud^2S \) and \( k_2 = u^2dS \) are the exercise prices of the two stock option packs. The four unknowns, or the optimal amount of stocks and stock options, can be obtained by solving the following linear equation system:

\[
\begin{align*}
\beta_0^* + \beta_1^* u^3S + \beta_2^* x^0 + \beta_3^* 0, \\
\beta_0^* + \beta_1^* u^2S + \beta_2^* x^0 + \beta_3^* 0, \\
\beta_0^* + \beta_1^* u dS + \beta_2^* x(u^2 dS - ud^2S) + \beta_3^* 0, \\
\beta_0^* + \beta_1^* u^3S + \beta_2^* x(u^3S - ud^2S) + \beta_3^* x(u^3S - u^2 dS) = R_i^*.
\end{align*}
\]
Of course, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_3 > 0$ hold by the convexity of optimal contract.

Next, we show that the convexity result is carried over to models with $N$ periods, even when $N \to +\infty$.

In the $N$-period model, the complex incentive compatibility constraints can be reduced into the effort choice of the manager in the last period. More specifically, if the project value is $u^n d^{N-n-1} S (n = 0, \ldots, N-1)$, its value in the next period could be $u^{n+1} d^{N-n-1} S$ or $u^n d^{N-n} S$, and the corresponding compensation and utility levels of the manager are $(R_{n+1}, U_{n+1})$ and $(R_n, U_n)$, respectively. Accordingly, the incentive compatibility constraint is

$$p_H U_{n+1} + (1-p_H) U_n \geq p_L U_{n+1} + (1-p_L) U_n + bu^n d^{N-n-1} S,$$

which can be equivalently expressed as

$$U_{n+1} - U_n \geq \frac{bS}{\Delta p} u^n d^{N-n-1},$$

and the manager’s participation constraint is

$$\sum_{n=0}^{N} C^*_n p_H^*(1-p_H)^{N-n} U_n \geq U_R.$$

---

7 Refer to the proof in Appendix A.
Similarly, all the incentive compatibility and participation constraints must be binding at the optimum. Therefore,

\[ U_n = U_0 + \frac{bS}{\Delta p} \sum_{i=0}^{n-1} d^{N-i} u' = U_0 + \frac{bS}{\Delta p(u-d)} d^{N-n} (u^n - d^n), \quad n \in \{1, \ldots, N\}. \quad (17) \]

Substituting equation 17 into the participation constraint, we have

\[
\sum_{n=0}^{N} C_n^n (1 - p_H)^{N-n} U_n
\]

\[ = U_0 + \frac{bS}{\Delta p(u-d)} \left\{ \sum_{n=0}^{N} C_n^n (1 - p_H)^{N-n} d^{N-n} (u^n - d^n) \right\}
\]

\[ = U_0 + \frac{bS}{\Delta p(u-d)} \left\{ \sum_{n=0}^{N} C_n^n p_H^n (1 - p_H)^{N-n} - \sum_{n=0}^{N} C_n^n p_H^n (1 - p_H)^{N-n} \right\}
\]

\[ = U_0 + \frac{bS}{\Delta p(u-d)} \left\{ (p_H u + (1 - p_H) d)^N - d^N \right\} = U_R. \]

Therefore,

\[ U_0 = U_R - \frac{bS}{\Delta p(u-d)} \overline{A}, \]

where \( \overline{A} \equiv [(p_H u + (1 - p_H) d)^N - d^N] \).

Notice that \( \overline{A} \) is independent of the project value \( S \), and, thus, \( U_0 \) is a linear function of \( S \). Then, from equation 17, all the possible utility levels, \( U_n \), are linear functions of \( S \). We can also reach this point by rewriting equation 17 as follows:

\[
\frac{R_{n+1} - R_n}{u^{n+1} d^{N-n} S - u^n d^{N-n} S} = \frac{b}{\Delta p(u-d)},
\]

where \( R_{n+1} \) is the manager’s utility level when the project value is \( u^{n+1} d^{N-n} S \). Proposition 2 follows.

**Proposition 2.** In an \( N \)-period dynamic moral hazard model, the optimal compensation scheme is convex in the firm value.

**Proof.** From the above analysis, we have \( R_{n+1} - R_n = h'(\xi_n)(U_{n+1} - U_n) \), where \( \xi_n \in (U_n, U_{n+1}) \) is monotonically increasing in \( n \). By the convexity of function \( h(\cdot) \), for all \( m < n \):

\[
\frac{R_{n+1} - R_n}{u^{n+1} d^{N-n} S - u^n d^{N-n} S} > \frac{R_{m+1} - R_m}{u^{m+1} d^{N-m} S - u^m d^{N-m} S},
\]

where \( R_{n+1} \) is the manager’s compensation when the project value is \( u^{n+1} d^{N-n} S \); thus, the piece rate again increases in the firm value. Therefore, the optimal contract is convex in the firm value in the finite \( N \)-period model.
Similarly, a restricted stock and \( N - 2 \) stock option packs with different exercise prices are required to implement the optimal contract in the general model with \( N \) periods.

In the limit (i.e. when \( N \to +\infty \)), the binomial process degenerates to the geometric Brown motion (GBM), and the manager’s ex post equilibrium utility level is continuous and linear in the firm value: 

\[
U(\tilde{S}) = \frac{b}{\Delta p(u - d)} \tilde{S} + C,
\]

where \( C \) is an integration constant, implicitly determined by the following equation:

\[
\int_0^{+\infty} U(x) dF(x) = \frac{b}{\Delta p(u - d)} \mathbb{E}(\tilde{S}) + C = U_R,
\]

in which \( F(x) \) is the distribution function of the GBM, and \( \mathbb{E}(\tilde{S}) \) is the expected firm value at the final date. Readily, the manager’s compensation is a convex function of the firm value:

\[
R(\tilde{S}) = h\left( \frac{b}{\Delta p(u - d)} \tilde{S} \right),
\]

and, thus, the \( N - 2 \) stock option packs at properly prespecified prices, together with the base wage and a linear piece rate (restricted stock), constitute the optimal compensation contract.

5. SOME COMPARATIVE STATICS

To complete our analysis, we further consider how the optimal amount of restricted stocks and stock options changes with the manager’s risk attitude (\( r \)), his or her reservation utility (\( U_R \)), the agency cost (\( p_{HD}b/\Delta p \)), as well as the firm value (\( S \)). Throughout this section, we assume that the manager has a CRRA utility function:

\[
v(x) = -\exp(-rx),
\]

so that 

\[
h(x) = -\frac{1}{r} \ln(-x),
\]

where \( r > 0 \) is the manager’s constant coefficient of absolute risk aversion. For simplicity and without loss of generality, in this section, we let \( ud = 1 \) (the case of \( un \neq 1 \) is quite similar but more complex) and concentrate on the two-period model.

**Proposition 3.** When the manager has a CRRA utility function, the optimal amounts of stock options and restricted stocks decrease with manager’s coefficient of absolute risk aversion, but increase with the reservation utility and the agency cost.

The proof of Proposition 3 is relegated to Appendix C. The intuition is as follows:

---

8 Refer to Cox et al. (1979) for details.

9 We, indeed, acknowledge the effect of the manager’s initial wealth on the optimal compensation scheme, but we neglect this effect here because wealth is generally unobservable in reality.
The first part of Proposition 3 is due to the traditional tradeoff between risk sharing and incentive provision: granting a certain amount of stock options can motivate the manager to work harder but he or she (as a risk averter) has to assume some risk. The equilibrium amount is determined by the optimal tradeoff. The optimal amount of stock options will decrease when the manager becomes more risk averse.

The second part of Proposition 3 holds because when the agent’s reservation utility increases, his or her bargaining power rises.

Finally, for the last part of Proposition 3, notice that we use $p_{hl} \Delta p$ to measure the agency cost of moral hazard, which is the likelihood ratio $p_{hl} \Delta p$ multiplied by the agent’s private benefit, $b$. The higher $p_{hl} \Delta p$ is, or, given $p_{hl}$, the higher $p_L$ is, the harder it is for the principal to infer the agent’s effort level from his or her observable information, so that the high powered incentive must be provided to avoid shirking. Similarly, a higher private benefit $b$ will entice the manager to shirk more often. Therefore, a larger value of $p_{hl} \Delta p$ represents a more severe degree of moral hazard.

Next, we turn to the analysis of firm value. The influence of firm value on the optimal amount of stock options and restricted stocks is a little more complicated: on the one hand, a higher firm value makes the private benefit of shirking more tempting, so that the amount of stock options and restricted stocks should be raised. On the other hand, when the firm value rises, the financial value of restricted stocks and options also increases, so that the amount should be reduced. It can be calculated that for the restricted stocks,

$$\frac{d\beta_1}{dS} = \frac{dR_s^*/dS - dR_0^*/dS}{d(u-d)S} - \frac{R_s^* - R_0^*}{d(u-d)S^2}, \quad (18)$$

where

$$\frac{dR_s^*/dS - dR_0^*/dS}{d(u-d)S} = -\frac{1}{r(u-d)S} \frac{b}{\Delta p} \frac{U_R^*}{U_0^*} > 0,$$

but $\frac{R_s^* - R_0^*}{d(u-d)S^2} < 0$. Similarly, for the stock options,

$$\frac{d\beta_2}{dS} = -\frac{1}{r(u-d)S} \frac{b}{\Delta p} \frac{U_R}{U_1^*} \left( \frac{1}{U_2^*} - \frac{1}{U_0^*} \right) \left( \frac{R_s^* - R_0^*}{d(u-d)S^2} - \frac{R_s^* - R_0^*}{d(u-d)S^2} \right). \quad (19)$$

Because the sign of equations 18 and 19 is ambiguous, we numerically simulate the relationship between the optimal amount of stock options and restricted stocks as in Figure 5. The results are summarized in Proposition 4.

**Proposition 4.** When the manager has a CRRA utility function, the optimal amount of stock options increases with the firm value, but that of restricted stocks decreases.
\[ \text{Beta} = 0.6, p_l = 0.3, b = 0.2, \text{ur} = -150, u = 2, d = 0.5 \]

\[ \text{Beta} = 0.8, p_l = 0.5, b = 0.3, \text{ur} = -150, u = 2, d = 0.5 \]

Figure 5. The effect of firm value on the optimal amount of restricted stocks and options

© 2013 Wiley Publishing Asia Pty Ltd
6. CONCLUSION

In a dynamic moral hazard framework, we have found that the optimal compensation scheme is convex. This implies that besides the base wage and restricted stocks, the optimal compensation scheme should include stock options. Therefore, we endogenously justify the widespread use of stock options as an incentive instrument. This result is carried over to the multiple period cases. In addition, we also provided comparative statics on the optimal amount of stock options and restricted stocks with respect to the manager’s risk attitude, reservation utility level and agency cost, as well as firm value with the help of numerical simulations.

The convexity of stock options in the optimal contract is independent of the specific functional form of the manager’s utility. It only requires the manager to be risk averse. The main driving force of the convexity result, in contrast to the linearity result obtained by Holmstrom and Milgrom (1987), is that we require interim firm value accumulation and Markov process of outputs across periods.

For further research, it would be helpful if multidimensionality of the manager’s action were introduced, as in Holmstrom and Milgrom (1991). For instance, the manager may be able to choose project quality, besides deciding his or her effort level. Moreover, tax and accounting considerations may also offer new insights into the use of executive stock options.

REFERENCES


© 2013 Wiley Publishing Asia Pty Ltd

APPENDICES

A. Simplifying the incentive compatibility constraints

In this appendix, we will show that all the incentive compatibility constraints in the basic model can be simplified into equations 3 and 4.

The expected utility of the manager is

\[ U = p_H U + (1 - p_H) U = p_H [p_H U_2 + (1 - p_H) U_1] + (1 - p_H) [p_H U_1 + (1 - p_H) U_0] \]

which are several ways of deviating away from the optimum. For instance, he or she may choose to work at \( t = 1 \), and continue to work at \( t = 2 \) if the project fails; otherwise, he or she chooses to shirk. In this case, the incentive compatibility constraint is:

\[ U \geq p_H [p_L U_2 + (1 - p_L) U_1 + bU] + (1 - p_H) [p_H U_1 + (1 - p_H) U_0], \]

which is equivalent to

\[ p_H U_2 + (1 - p_H) U_1 \geq p_L U_2 + (1 - p_L) U_1 + bU, \tag{A1} \]

and this is equation 3.

Similarly, the manager may choose to work at \( t = 1 \), and continue to work at \( t = 2 \) if the project succeeds; otherwise, he or she chooses to shirk. In this case, the incentive compatibility constraint is:

\[ U \geq p_H [p_H U_2 + (1 - p_H) U_1] + (1 - p_H) [p_L U_1 + (1 - p_L) U_0 + bU], \]

which is equivalent to

\[ p_H U_1 + (1 - p_H) U_0 \geq p_L U_1 + (1 - p_L) U_0 + bU, \tag{A2} \]

and this is equation 4.

If, instead, the manager chooses to work at \( t = 1 \), but to shirk at \( t = 2 \) regardless of whether the project succeeds or fails, then his or her expected utility is

\[ p_H [p_L U_2 + (1 - p_L) U_1 + bU] + (1 - p_H) [p_L U_1 + (1 - p_L) U_0 + bU], \]

which is always lower than \( U \), from equations A1 and A2. Therefore, under constraints A1 and A2, the manager will never choose this way of deviation.

Surely, the manager might well choose to shirk at \( t = 1 \), and to work at \( t = 2 \) regardless of whether the project succeeds or fails. Then, the corresponding incentive compatibility constraint is:
\[ U \geq p_L[p_H U_2 + (1 - p_H) U_1] + (1 - p_L)[p_H U_1 + (1 - p_H) U_0] + b S, \]

which can be simplified as
\[ \Delta p \cdot [p_H (U_2 - U_1) + (1 - p_H) (U_1 - U_0)] \geq b S. \]  

(A3)

It is not difficult to show that equation A3 always holds under constraints A1 and A2 and the assumption \( p_H U + (1 - p_H) d > 1 \).

With conditions A1, A2 and A3, it can be shown that
\[ U \geq p_L[p_H U_2 + (1 - p_H) U_1] + (1 - p_L)[p_H U_1 + (1 - p_H) U_0] + b S \]
\[ \geq p_L[p_L U_2 + (1 - p_L) U_1 + b u S] + (1 - p_L)[p_H U_1 + (1 - p_H) U_0] + b S, \]

\[ U \geq p_L[p_H U_2 + (1 - p_H) U_1] + (1 - p_L)[p_H U_1 + (1 - p_H) U_0] + b S \]
\[ \geq p_L[p_L U_2 + (1 - p_L) U_1 + bu S] + (1 - p_L)[p_L U_1 + (1 - p_L) U_0 + bd S] + b S, \]

Therefore, there are no other forms of deviation for the manager.

To summarize, all the possible incentive compatibility constraints can be simplified as constraints 3 and 4.

B. Proof of Lemma 1

The optimization problem of the principal is:
\[ \min_{U_2,U_1,U_0} p_H U_2 + 2 p_H (1 - p_H) U_1 + (1 - p_H)^2 U_0 \]
\[ \text{s.t. } U_2 - U_1 \geq bu S / \Delta p, \]
\[ U_1 - U_0 \geq bd S / \Delta p, \]
\[ p_H^3 U_2 + 2 p_H (1 - p_H) U_1 + (1 - p_H)^2 U_0 = U'_R. \]

The corresponding Lagrangian function is
\[ \mathcal{L} = p_H U_2 + 2 p_H (1 - p_H) U_1 + (1 - p_H)^2 U_0 - \lambda_2 (U_2 - U_1 - bu S / \Delta p) \]
\[ -\lambda_2 (U_1 - U_0 - bd S / \Delta p) - \mu (p_H^3 U_2 + 2 p_H (1 - p_H) U_1 + (1 - p_H)^2 U_0 - U'_R), \]

the first order conditions of which are
\[ \frac{d \mathcal{L}}{d U_2} = p_H^3 h'(U_2) - \lambda_1 - \mu p_H^3 = 0, \]
\[
\frac{d\mathcal{L}}{dU_1} = 2p_H(1-p_H)h'(U_1) + \lambda_1 - \lambda_2 - \mu(2p_H(1-p_H)) = 0,
\]

\[
\frac{d\mathcal{L}}{dU_0} = (1-p_H)^2h'(U_0) + \lambda_2 - \mu(1-p_H)^2 = 0.
\]

Therefore,

\[
\mu = p_H^2h'(U_2) + 2p_H(1-p_H)h'(U_1) + (1-p_H)^2h'(U_0).
\]

Because \(U_2 > U_1 > U_0, h'' > 0\), we have \(\mu \in (h'(U_0), h'(U_2))\), and then

\[
\lambda_1 = p_H^2[\mu - h'(U_2)] < 0,
\]

\[
\lambda_2 = (1-p_H)^2[h'(U_0) - \mu] < 0.
\]

From the complementary slackness conditions, both of the incentive compatibility conditions \(IC_{21}\) and \(IC_{22}\) are binding.

C. Proof of Proposition 3

First, consider the relationship between \(\beta_1\) (and \(\beta_2\)) and the manager’s risk attitude, \(r\). Because \(\frac{dR^*}{dr} = \frac{1}{r^2} \ln(-U^*) = -\frac{1}{r} R^*\), in the two-period model, applying the convexity of the optimal compensation contract, we have

\[
\frac{d\beta_1}{dr} = \frac{dR^*_i}{dr} - \frac{dR^*_i}{dr} \frac{dR^*_i}{dr}/dS = -\frac{1}{r} \frac{R^*_i - R^*_0}{d(u-d)S} < 0;
\]

\[
\frac{d\beta_2}{dr} = \frac{dR^*_i}{dr} - \frac{dR^*_i}{dr} \frac{dR^*_i}{dr}/dS - \frac{dR^*_i}{dr} \frac{dR^*_i}{dr} \frac{dR^*_i}{dr}/dS
\]

\[
= -\frac{1}{r} \left\{ \frac{R^*_2 - R^*_1}{u(u-d)S} - \frac{R^*_1 - R^*_0}{d(u-d)S} \right\} < 0.
\]

Second, consider the relationship between \(\beta_1\) (and \(\beta_2\)) and the manager’s reservation utility. Because \(\frac{dR^*}{dU_R} = h''_i\), in the two-period model, applying \(U^*_2 - U^*_1 = \frac{h_2S}{Ap}, U^*_1 - U^*_0 = \frac{h_1S}{Ap}\) and \(U^*_0 < U^*_1 < U^*_2 < 0\), we have

\[
\frac{d\beta_1}{dU_R} = \frac{1}{d(u-d)S} \left\{ \frac{dR^*_i}{dU_R} - \frac{dR^*_i}{dU_R} \right\} = -\frac{1}{rd(u-d)S} \left\{ \frac{1}{U^*_1} - \frac{1}{U^*_0} \right\} > 0,
\]

© 2013 Wiley Publishing Asia Pty Ltd
Finally, consider the relationship between $\beta_1$ (and $\beta_2$) and the agency cost $b/\Delta p$. In the two-period model, we can derive that \( \frac{dR_0^*}{d(b/\Delta p)} = -h_0^* AS \), \( \frac{dR_1^*}{d(b/\Delta p)} (d - A)S \), and \( \frac{dR_2^*}{d(b/\Delta p)} = h_2^* (u + d - A)S \). Therefore,

\[
\frac{d\beta_1}{d(b/\Delta p)} = \frac{\frac{dR_1^*}{d(b/\Delta p)} - \frac{dR_0^*}{d(b/\Delta p)}}{d(u - d)S} = \frac{1}{d(u - d)} \left\{ h_1^* (d - A) + h_0^* \right\} A
\]

\[
= \frac{1}{ru(u - d)} \left\{ \frac{d - A}{U_1^*} - \frac{A}{U_0^*} \right\} = -\frac{1}{r(u - d)} \frac{U_R}{U_1^* U_0^*} > 0
\]

and

\[
\frac{dR_2^*}{d(b/\Delta p)} - \frac{dR_1^*}{d(b/\Delta p)} = \frac{1}{u(u - d)S} \left\{ h_2^* (u + d - A) - h_1^* (d - A) \right\}
\]

\[
= \frac{1}{ru(u - d)} \left\{ \frac{u + d - A}{U_2^*} - \frac{d - A}{U_1^*} \right\} = -\frac{1}{r(u - d)} \frac{U_R}{U_2^* U_1^*},
\]

so that

\[
\frac{d\beta_2}{d(b/\Delta p)} = \frac{\frac{dR_2^*}{d(b/\Delta p)} - \frac{dR_1^*}{d(b/\Delta p)}}{u(u - d)S} - \frac{\frac{dR_1^*}{d(b/\Delta p)} - \frac{dR_0^*}{d(b/\Delta p)}}{d(u - d)S}
\]

\[
= -\frac{1}{r(u - d)} \frac{U_R}{U_1^*} \left( \frac{1}{U_2^*} - \frac{1}{U_0^*} \right) > 0.
\]