A Graphical Approach to the Standard Principal-Agent Model

Xianming Zhou

Abstract: An obstacle to the teaching of principal-agent theory is the technical complexity and intractability of the general model. Even in academic studies strong assumptions are often imposed so as to derive an analytical solution. The author describes a graphical approach to the standard principal-agent model. Characterizing equilibrium in the contract space defined by the incentive parameter and insurance component of pay under a linear contract, this approach provides a simple and intuitive method for analyzing the principal-agent problem, which can be easily understood by students of economics with basic knowledge of algebra and differentiation. The approach has shown to be convenient and rich for comparative statics analyses.

Key words: agency theory, contracts, managerial incentives
JEL codes: C61, J41, A22

The classic principal-agent model provides an elegant theory of the moral hazard problem, but it yields few general results.1 Perhaps because of the technical complexity and intractability in its general framework, principal-agent theory is often not taught except to graduate students concentrating in economic theory or industrial organization. Studies of applied principal-agent models, however, have grown tremendously in recent years in various fields of economics. Hence, it is both desirable and important to extend the teaching of agency theory to a broad range of students and in a more intuitive approach. In this article, I describe a graphical approach to the standard principal-agent model. In the contract space defined by the incentive parameter and insurance component of pay under a linear contract, the approach characterizes equilibrium using a “feasible contract curve/indifference curves” framework. Because the approach is simple and intuitive in showing fundamental results of principal-agent theory, it can be understood easily by economics students with basic knowledge of algebra and differentiation.

A rapidly growing literature of applied principal-agent theory examines managerial incentive contracts and executive compensation. In this literature, a simplified model is often developed to obtain theoretical guidance to certain

Xianming Zhou is an assistant professor in the Department of Economics and Finance at the City University of Hong Kong and at the School of Business at the University of Sydney, Australia (e-mail: efxzhou@cityu.edu.hk). The author is grateful to three anonymous referees for helpful comments and suggestions.
specific incentive issues. Strong assumptions are often imposed on the general principal-agent model so that an analytical solution becomes attainable. These assumptions usually include exponential utility, quadratic cost of effort, a linear production function with normally distributed noise, and a linear contract (e.g., Gibbons and Murphy 1992; Garen 1994; Murphy 1999; Baker and Hall 1998). The graphical approach I describe in this article relaxes most of these assumptions. The approach is convenient and intuitive in analyzing equilibrium and comparative statics and provides further insights into the optimal contract.

After I describe the graphical approach and characterize the optimal contract for the standard principal-agent model, I discuss applications of the approach, focusing on the effect of important contractual factors such as risk, productivity, and firm size.

**THE GRAPHICAL APPROACH**

Consider a single agent and the principal in a one-period contract. The production function of the agent is

\[ Y = F(e) + \varepsilon, \]

where \( e \) is the agent’s investment of effort, which is unobservable to the principal, and \( F(e) \) defines expected output, which is strictly concave in effort, that is, \( F'(e) > 0 \) and \( F''(e) < 0 \), and where \( F(0) = 0 \). The second term, \( \varepsilon \), is a random variable with zero mean. The variance of \( \varepsilon \) measures the risk of output. To derive a graphical framework, the distribution of output needs to be well behaved in the sense that the first-order approach applies. The main audience of the graphical approach may be uninterested in model details and insensitive to conditions for the first-order approach. Yet it is important to note that the graphical framework is subject to the qualification that the first-order approach is valid.

The agent’s utility, denoted as \( U \), is a general function of his or her net payoff, \( W - C \), where \( W \) is the realized wage and \( C \) the cost of effort. The utility function is strictly concave in net payoff, that is, \( U'(W - C) > 0 \) and \( U''(W - C) < 0 \), and where \( U'(0) \to +\infty \). The cost function is strictly convex in effort, that is, \( C'(e) > 0 \) and \( C''(e) > 0 \), except at \( e = 0 \) where \( C'(0) = 0 \). Assume a linear pay contract for the agent,

\[ W = \alpha + \beta Y, \]

where \( \alpha \) is the fixed component of pay or insurance pay, and \( \beta \) the incentive parameter or pay-performance sensitivity. In the literature of managerial incentives, \( \beta \) is interpreted as incentive strength.

The agent chooses effort to maximize his or her expected utility, \( E[U(W - C)] \), where \( E \) is the expectation operator, yielding the following incentive compatibility constraint:

\[ \beta F'(e) - C'(e) = 0. \]
The constraint dictates that effort increase with the incentive parameter. The principal is risk-neutral and, in a competitive labor market, earns zero expected profit. Then, \( E(Y - W) = 0 \), or

\[
\alpha = (1 - \beta)F(e). \tag{2}
\]

The optimization problem of the model is to maximize the agent’s expected utility subject to the incentive compatibility constraint, equation (1), and the zero expected profit condition, equation (2). Essentially, the model needs two simplifying assumptions: the distribution of production satisfies the conditions for the first-order approach and the contract is linear.

To characterize the optimal contract in a diagram, I consider the space between the incentive parameter, \( \beta \), and the insurance component of pay, \( \alpha \). Let the horizontal axis represent \( \beta \) and the vertical axis represent \( \alpha \), as shown in Figure 1. A point in this space corresponds to a contract, and hence, all possible combinations of \( \alpha \) and \( \beta \) define the contract space. The optimal contract is the point in the space at which the agent’s expected utility is maximized.

For (possible) contracts to be feasible, they must satisfy both the incentive compatibility constraint, equation (1), and the zero expected profit condition, equation (2). All feasible contracts in the contract space form a locus, which I call the feasible contract curve (FCC). The shape of FCC is easy to determine. From equation (2), \( \alpha \) is positive for \( 0 < \beta < 1 \), and it equals 0 at \( \beta = 1 \). Furthermore, if \( \beta = 0 \), equation (1) determines \( e = 0 \) because \( C'(e) = 0 \), and then \( \alpha = F(0) = 0 \) from equation (2). Thus, FCC is increasing in \( \beta \) at the origin, and after a certain point it becomes decreasing in \( \beta \). It goes back to the horizontal axis at \( \beta = 1 \).

The shape of FCC can be formally obtained by analyzing its slope. Noticing that effort is an implicit function of \( \beta \), differentiating equations (1) and (2) with respect to \( \beta \) gives, respectively,
Given the functional properties of \( F(e) \) and \( C(e) \), the second term in equation (3) is positive for \( 0 < \beta < 1 \). As \( \beta \) changes from 0 to 1, \( e \) increases monotonically, following the incentive compatibility constraint. The slope of \( FCC \) accordingly changes from \( +\infty \) to 0 and then turns negative and becomes \( -F(e) \) at \( \beta = 1 \).

\( FCC \) has the shape shown in Figures 1 through 6: It starts at the origin and goes up and outward, then bends down and further outward, and back to the horizontal axis at \( \beta = 1 \). Moving outward along \( FCC \), \( \beta \) increases from 0 to 1. The intuition of the hill-shaped \( FCC \) curve is the following. There are two opposite effects from increasing the incentive parameter. On the one hand, a larger \( \beta \) means more effort and thus higher output, which tends to raise both the incentive and insurance components of pay, given the zero expected profit constraint. On the other hand, a larger \( \beta \) means a smaller portion of total pay in the insurance component. For low values of \( \beta \), the marginal product of effort is high, and the output effect of \( \beta \) on \( \alpha \) is dominating, and hence, the curve slopes up. Past its peak, the output effect drops below the switching effect from insurance pay to incentive pay, which dictates that the curve be downward sloped.

I now turn to indifference curves (IC). By definition, an indifference curve is a curve in the contract space along which the expected utility remains constant. To derive the slope of \( IC \), I set \( E[U(\alpha + \beta (F(e) + \varepsilon) - C(e))] \) equal to \( R \), where \( R \) is a constant, then I differentiate both sides of this expression by \( \beta \), which yields

\[
E[U'(W - C)] \left[ \frac{\partial \alpha}{\partial \beta} + F(e) + \beta F'(e) \frac{\partial e}{\partial \beta} - C'(e) \frac{\partial e}{\partial \beta} \right] + E[U'(W - C)\varepsilon] = 0. \tag{4}
\]

Solving equation (4) and using the incentive compatibility constraint gives

\[
\frac{\partial \alpha}{\partial \beta} \bigg|_{IC} = -F(e) - \frac{E[U'(W - C)\varepsilon]}{E[U'(W - C)]}. \tag{5}
\]
Note that $U'(W - C)$ is decreasing in $W - C$ and hence in $\varepsilon$, and that $E(\varepsilon) = 0$, so the term $E[U'(W - C)\varepsilon]$ is negative.\(^6\)

Because the second term in equation (5) can be any positive value,

$$\frac{\partial \alpha}{\partial \beta}|_{IC},$$

is either negative or positive, depending on the functional form of $U$ and the distribution of $\varepsilon$. In other words, $IC$ can slope either downward or upward in the contract space. This is not surprising because, to the agent, the insurance component of pay is always a good whereas incentive strength ($\beta$) can be either a good or a bad. Because a higher $\beta$ raises compensation to the agent but increases risk, it is a good to the agent when risk aversion is low and/or when production has little risk such that the former effect dominates (when the first term in equation (5) is dominating). When the reverse holds, $\beta$ becomes a bad to the agent. With $IC$ being upward sloped, $\alpha$ and $\beta$ become complementary. This second case, seemingly counterintuitive, is particularly interesting. It is barely noted in previous studies although, as will be shown, it contains important implications.

The optimal contract is determined jointly by $FCC$ and $IC$. In Figure 1, I present the equilibrium for the special case with a risk-neutral agent [having constant $U'(W - C)$] or riskless production ($\varepsilon = 0$), where $E[U'(W - C)\varepsilon] = 0$. With the slope of $IC$ being $-F(e)$, it is steeper than $FCC$ everywhere except at $\beta = 1$, where the two curves are tangent. Equilibrium is then attained at $\beta = 1$ and $\alpha = 0$, where the first-best solution, denoted as $E^*$, is achieved. It is worth noting that, in this case, $IC$ is concave in the contract space. This is because the slope of $IC$, $-F(e)$, increases in magnitude with $\beta$ due to an incentive effect on effort.

In general, the agent is risk averse and production is risky, so the second term in equation (5) is positive. With $IC$ being flatter than $FCC$ at $\beta = 1$, equilibrium occurs at some point with $\beta < 1$ (and thus $\alpha > 0$), where the second-best solution is obtained. As the degree of risk aversion increases, the second term in equation (5) increases, and accordingly, $IC$ rotates counterclockwise and changes from concave to convex. This occurs because the effect of the random term in the utility function, $\beta \varepsilon$, increases with risk aversion, reducing the marginal rate of substitution between $\beta$ and $\alpha$ (with the expected marginal utility of $\beta$ declining relative to the expected marginal utility of $\alpha$).\(^7\) Two cases of the equilibrium for a risk-averse agent are shown in Figure 2. $IC_1$ and $IC_2$ present the typical case where both $\alpha$ and $\beta$ are a good to the agent. Equilibrium occurs at a point in the decreasing part of $FCC$, denoted as $E$. In this case, any small change in the agent’s taste will cause a change in equilibrium along $FCC$ characterized by the substitution between $\alpha$ and $\beta$.

As the agent becomes more risk averse, indifference curves further rotate counterclockwise and result in smaller pay-performance sensitivities. When the agent is sufficiently risk-averse, the second term in equation (5) becomes dominating, and consequently, indifference curves become upward sloped as shown by $IC_1'$ and $IC_2'$ in Figure 2. The equilibrium in this case occurs at a point (e.g.,

Summer 2002
$E'$ in the increasing part of $FCC$. The implication here is that comparative statics would predict complementary changes in the incentive parameter and insurance component of pay.

With the graphical framework, therefore, it is easy to show that the co-existence of production risk and risk-averse behavior result in agency costs and that higher agency costs dictate smaller incentive parameters. This fundamental result of the principal-agent theory (Holmstrom 1979) becomes very intuitive in such a diagram, even though in this model utility, production, and the cost of effort are all a general function.

**APPLICATION**

**The Effect of Risk**

The relationship between managers and shareholders presents the classic example of the principal-agent problem. Agency theory posits an economic tradeoff between inducing managerial effort and minimizing the cost born by risk-averse managers. Confirming this tradeoff, many applied principal-agent models predict a negative effect of risk on pay-performance sensitivities, which finds empirical support in studies of executive compensation (e.g., Garen 1994; Aggarwal and Samwick 1999).

The graphical approach described above provides a convenient framework to address this issue. Because $FCC$ does not change with the distribution of the random variable $\epsilon$, the effect of risk is simply a shift in equilibrium along $FCC$. As risk in output increases variations in managerial pay, it reduces the marginal benefit of increasing $\beta$ and thus causes $IC$ to rotate counterclockwise. As shown in Figures 3 and 4, the optimal contract shifts inward along $FCC$ from $E$ to $E'$ as a result of increased risk. In either case the sensitivity becomes smaller. The effect is qualitatively the same as that of increasing the degree of risk aversion.
The risk effect on the insurance component of pay is ambiguous, depending on the position of the initial equilibrium. The case shown in Figure 4 is particularly interesting: $\alpha$ and $\beta$ change together, both decreasing in output risk. This result appears to be surprising and inconsistent with the conventional impression that the tradeoff between incentives and insurance requires opposite changes in $\alpha$ and $\beta$. This can be explained as follows: when output is sufficiently risky, incentive strength is a bad to the risk-averse agent. This requires that equilibrium be attained at the uphill side of $FCC$, where the output effect of $\beta$ is dominating. Any small change in model parameters including riskiness will lead to a sufficient change in $\beta$, which requires the fixed component of pay to adjust in the
The Effect of Productivity

Higher expected output makes incentive pay more attractive. Hence, intuition suggests a positive relationship between managerial productivity and incentive strength. Taking the marginal product of effort to be a measure of productivity, simplified principal-agent models obtain analytical solutions confirming this relationship (Milgrom and Roberts 1992; Schaefer 1998; Baker and Hall 1998).

Using the graphical approach, one can easily analyze this relationship with the more general model. Productivity is captured by the expected component of output, $F(e)$. Higher productivity means higher $F'(e)$, given the cost function of effort. Because both $FCC$ and $IC$ will change with $F(e)$, the effect of productivity becomes slightly more complicated. On one hand, $FCC$ shifts up because $\alpha$ increases with expected output given $\beta$. On the other hand, given the distribution of the random term, the expected marginal utility of $\beta$ increases with productivity because of the increase in the component $\beta F(e) - C(e)$ in the utility function. As the expected marginal utility of $\beta$ increases relative to the expected marginal utility of $\alpha$, $IC$ rotates clockwise.

The total effect can hypothetically be decomposed into two changes as shown in Figure 5. When $\beta$ is maintained unchanged, there is an output effect, which is the upward vertical shift from $E$ to $E'$. This effect increases the agent’s utility without distorting risk allocation between the agent and the principal. Along the
new feasible contract curve, $FCC'$, there is a risk effect, which is the change from $E'$ to $E''$. This effect is qualitatively similar to that shown in Figures 3 and 4 but in the opposite direction (equivalent to a reduction in output risk). The total effect (from $E$ to $E''$) is that $\beta$ becomes larger. The change in $\alpha$ is ambiguous, depending on the position of the initial equilibrium and the strength of the productivity-induced risk effect.

**Firm Size and Incentive Strength**

On the basis of the discussions in the two previous subsections, I use the graphical approach to examine the relationship between firm size and incentive strength—an interesting and important issue in recent studies of executive compensation. For clarity in discussion, the production function is redefined as follows:

$$ Y = F(K,e) + H(K)e, $$

where $K$ denotes capital, which is a proxy for firm size. The effect of capital on productivity and output variations is captured by $F(K,e)$ and $H(K)$, respectively. The random variable, $\varepsilon$, characterizes risk heterogeneity holding company size constant. The slopes of $FCC$ and $IC$ become the following:

$$\frac{\partial \alpha}{\partial \beta}\bigg|_{FCC} = -F(K,e) + \frac{(1-\beta)[F'_e(K,e)]^2}{C''(e) - \beta F''_e(K,e)}, $$

$$\frac{\partial \alpha}{\partial \beta}\bigg|_{IC} = -F(K,e) - \frac{E[U'(W-C)H(K)e]}{E[U'(W-C)]}. $$

In equation (6), $F'_e(K,e)$ and $F''_e(K,e)$ denote the first-order and second-order derivatives, respectively, of $F(K,e)$ with respect to $e$.

In one extreme case, capital increases output variations but does not affect productivity, that is, $H'(K) > 0$ and $F'_K(K,e) = 0$. In this case, firm size causes a pure risk effect: $FCC$ remains unchanged whereas $IC$ rotates counterclockwise because of a capital-induced increase in the random term, $H(K)e$. The effect is qualitatively the same as that described in Figures 3 or 4. In another extreme case, capital increases productivity but does not change variations in output, that is, $F'_K(K,e) > 0$ and $H(K) = 0$. In this case, both $FCC$ and $IC$ change, and firm capital causes a typical productivity effect, which can be similarly analyzed as in Figure 5. In general, capital raises both productivity and output variations, and hence, $F'_K(K,e) > 0$ and $H'(K) > 0$. The total effect is the combination of the two special cases, which is shown in Figure 6. The incentive parameter either becomes larger as the resulting change from $A$ to $D$ or becomes smaller from $A$ to $E$, depending on the position of the initial equilibrium and the relative strength of the output effect and the total risk effect. The output effect corresponds to the shift from $A$ to $B$. The total risk effect corresponds to the shift from $B$ to $C$ (from the productivity-induced change in risk) and then to either $D$ or $E$ (from increased output variations). The change in the insurance component of pay again remains indeterminate.
Examining a simple principal-agent model, Schaefer (1998) and Baker and Hall (1998) also obtain an ambiguous effect of firm size on the incentive parameter. Because it is intrinsically unclear as to which (output or risk) effect of capital is dominating, agency theory does not generally establish an unambiguous relationship between firm size and incentive strength. It is essentially an empirical matter. The evidence for this relationship is so far inconclusive. Examining CEO compensation of large U.S. firms, Jensen and Murphy (1990), Garen (1994), Schaefer (1998), and Baker and Hall (1998) all found a negative relationship between firm size and the arithmetic sensitivity of CEO pay to shareholder wealth. However, investigating the elasticity of CEO pay with respect to firm value for similar U.S. companies, Hall and Liebman (1997) argued that executive compensation is more sensitive to performance in larger firms. Without a doubt, further studies are needed before a firm conclusion can be reached on this issue.

CONCLUSION

Research interest in principal-agent theory and its application has grown tremendously in the past two decades. The teaching of the theory (to undergraduate students, in particular) has not kept pace, partly because of its mathematical intractability. I provide a graphical presentation of the standard principal-agent model, which characterizes equilibrium in the contract space defined by incentive strength and insurance pay. This approach provides a simple and intuitive method for analyzing the optimal contract, which can be easily understood by students of economics with basic knowledge of algebra and differentiation. By examining the effect on the optimal contract of important contractual factors such as risk, productivity and firm size, the approach is shown to be convenient and rich for comparative statics analyses.
The graphical approach relaxes some simplifying assumptions that are often needed to derive an analytical solution. However, it needs pointing out that the approach does not present a general method for analyzing the principal-agent problem. Indeed, as noted earlier, it is based on two important assumptions: the first-order approach and linear contracts. The graphical approach can be used as an intuitive introduction to the general theory for less advanced students, though it can also be useful to understanding some applied models that are more difficult to analyze otherwise because of technical challenges.

NOTES

1. For pioneer work on principal-agent theory, see Mirrlees (1975) and Holmstrom (1979).
2. Milgrom and Roberts (1992) made use of the concept of certainty equivalent to avoid technical complexity associated with expected utility.
3. The first-order approach relaxes the constraint of the general model that the agent choose a utility-maximizing action to require instead only that the agent choose an action at which his or her utility is at a stationary point. Being more mathematically tractable, the first-order approach has been the standard method for analyzing the principal-agent problem, although it is not generally correct. For the first-order approach to be valid, two distributional assumptions (sufficient conditions) need to be satisfied, which are known as the monotone likelihood ratio condition (MLRC) and the convexity of distribution function condition (CDFC). For discussions on these conditions, see Mirrlees (1979), Grossman and Hart (1983), and Rogerson (1985).
4. A linear contract is widely used for analytical convenience. Holmstrom and Milgrom (1987) discuss conditions under which linear contracts are optimal.
5. Note that the operators of expectation and differentiation are exchangeable. Differentiating \( E[U(W - C)] \) with respect to \( \beta \) gives

\[
\frac{\partial E[U(W - C)]}{\partial \beta} = E \left[ \frac{\partial U(W - C)}{\partial \beta} \right] = E \left[ \frac{\partial U(W - C)}{\partial (W - C)} \cdot \frac{\partial (W - C)}{\partial \beta} \right] \\
= E \left[ U'(W - C) \left( \frac{\partial x}{\partial \beta} + \beta F'(e) \frac{\partial e}{\partial \beta} - C'(e) \frac{\partial e}{\partial \beta} \right) \right] + E \left\{ U'(W - C) \right\}.
\]

6. This can be formally proved as follows. Use \( f(\varepsilon) \) to denote the probability density function, then,

\[
E[U'(W - C)\varepsilon] = \int U'(W - C)\varepsilon f(\varepsilon)\,d\varepsilon = \int_{\varepsilon > 0} U'(W - C)\varepsilon f(\varepsilon)\,d\varepsilon + \int_{\varepsilon < 0} U'(W - C)\varepsilon f(\varepsilon)\,d\varepsilon.
\]

Because \( U'(W - C) \) is positive and \( U'(W - C)_{\varepsilon < 0} > U'(W - C)_{\varepsilon = 0} > U'(W - C)_{\varepsilon > 0} \), the first term above is negative and has the following relationship:

\[
\int_{\varepsilon < 0} U'(W - C)\varepsilon f(\varepsilon)\,d\varepsilon < \int_{\varepsilon < 0} U'(W - C)_{\varepsilon = 0} f(\varepsilon)\,d\varepsilon = U'(W - C)_{\varepsilon = 0} \int_{\varepsilon < 0} f(\varepsilon)\,d\varepsilon.
\]

The second term is positive, and with the similar reasoning,

\[
\int_{\varepsilon > 0} U'(W - C)\varepsilon f(\varepsilon)\,d\varepsilon < U'(W - C)_{\varepsilon = 0} \int_{\varepsilon > 0} f(\varepsilon)\,d\varepsilon.
\]

\( E[U'(W - C)\varepsilon] < 0 \) follows immediately.
7. A similar graphical analysis can be done in the \([\beta F(e) - C(e)] \times \alpha\) space, where the horizontal axis is the component of net payoff in utility that is increasing in \(\beta\). In this space, it is more straightforward to discuss changes in the slope and curvature of \(IC\) in response to changes in the degree of risk aversion and riskiness. However, in this space the interpretation of the horizontal axis is less intuitive, and so I do not pursue this approach.

8. \(FCC\) is derived from the incentive compatibility constraint, equation (1), and the zero expected profit condition, equation (2), neither of which contain the random variable.

REFERENCES


