

# Lecture Note 1

## Agency Theory<sup>1</sup>

R. Gibbons  
MIT

This note considers the simplest possible organization: one boss (or “Principal”) and one worker (or “Agent”). One of the earliest applications of this Principal-Agent model was to sharecropping, where the landowner was the Principal and the tenant farmer the Agent, but in this course we will typically talk about more familiar organization structures. For example, we might consider a firm’s shareholders to be the Principal and the CEO to be the Agent. One can also enrich the model to analyze a chain of command (*i.e.*, a Principal, a Supervisor, and an Agent), or one Principal and many Agents, or other steps towards a full-fledged organization tree.

The central idea behind the Principal-Agent model is that the Principal is too busy to do a given job and so hires the Agent, but being too busy also means that the Principal cannot monitor the Agent perfectly. There are a number of ways that the Principal might then try to motivate the Agent: this note analyzes incentive contracts (similar to profit sharing or sharecropping); later notes discuss richer and more realistic models.

Taken literally and alone, the basic Principal-Agent model may seem too abstract to be useful. But we begin with this model because it is an essential building block for many discussions throughout the course—concerning not only managing the incentives of individuals but also managing the incentives of organizational units (such as teams or divisions) and of firms themselves (such as suppliers or partners). Furthermore, this abstract model allows us to consider the nature and use of economic models more generally, as follows.

---

<sup>1</sup> This note draws on R. Gibbons, “Incentives in Organizations,” *Journal of Economic Perspectives* 12 (1998): 115-132 and R. Gibbons, “Incentives Between Firms (and Within),” forthcoming in *Management Science* and available at <http://web.mit.edu/rgibbons/www/index.html>.

## 1. An Introduction to Economic Modeling

We will use several economic models in this course, so it may be helpful to begin by describing what an economic model is and what it can do. We will defer discussion of whether such models are useful until after we have a few under our belts!

*An economic model is a simplified description of reality, in which all assumptions are explicit and all assertions are derived.* Such a model can produce qualitative and/or quantitative predictions. A qualitative prediction is that “x goes up when y falls.” A quantitative prediction is that  $x = 1/y$ . *A model’s (qualitative or quantitative) predictions are useful when they are robust within the environment(s) of interest.*

Quantitative predictions often hinge on specific assumptions from the model. If the model will be applied in one particular environment (such as a queuing model describing the lines at the Refresher Course, or the Black-Scholes model for option pricing) then the specific assumptions need to match the environment fairly closely, otherwise the quantitative predictions will not be useful in that environment. One might call this “engineering modeling” rather than “economic modeling.”

Qualitative predictions are often more robust, in two senses. First, *qualitative predictions may continue to hold if one makes small changes in the model’s specific assumptions.* For example, a model’s quantitative predictions might depend on whether a particular probability distribution is normal, exponential, or uniform, but the model’s qualitative predictions might hold for any single-peaked (*i.e.*, hill-shaped) distribution, including the three mentioned above as well as others.

Qualitative predictions can also be robust in a second (and, for our purposes, more important) sense: *a simple model’s qualitative predictions may be preserved even if one adds much more richness to the model.* The major points we will derive from the economic models in this course are robust predictions in this latter sense. That is, adding greater richness and realism to these models will certainly change the models’ quantitative conclusions, but the major points we derive from the simple models will still be part of the package of qualitative conclusions from the richer models.

## 2. Pay for Performance: The Basic Principal-Agent Model

During this course we will frequently use the term “incentives.” In some settings we will mean a cash payment for a measured outcome, but in other settings our use of this term will be much broader. Lest anyone be misled or disaffected by the narrowness of the former meaning, we will start our discussion of the basic Principal-Agent model by attempting some broader definitions: let “rewards” mean outcomes that people care about (not just dollars), let “effort” mean actions that people won’t take without rewards (not just hours worked), and let “incentives” mean links between rewards and effort (not just compensation contracts). We will refine these definitions throughout the course. For now we simply note that, according to these definitions, there are clearly lots of incentives out there, even if there are many fewer dollar-denominated incentive contracts.

To be more precise about rewards, effort, and incentives, we turn now to the elements of the basic Principal-Agent model: (A) the technology of production, (B) the set of feasible contracts, (C) the payoffs to the parties, and (D) the timing of events.

### A. *The Technology of Production*

In this simple model, the production process is summarized by just three variables: (1) the Agent’s total contribution to firm value (or, for now, the Agent’s “output”), denoted by  $y$ ; (2) the action the Agent takes to produce output, denoted by  $a$ ; and (3) events in the production process that are beyond the Agent’s control (*i.e.*, “noise”), denoted by  $\varepsilon$ .

(1) The Agent’s contribution to firm value,  $y$ : In the sharecropping context, the Agent’s contribution is simply the harvest. In the CEO context, one definition of the Agent’s contribution is the change in the wealth of the shareholders through appreciation in the firm’s stock price. For workers buried inside an organization, it is sometimes very difficult to define and measure a contribution to firm value. Later in this note we will discuss alternative objective performance measures (which sometimes raise “get what you pay for” issues); in later notes we will discuss subjective performance measures.

(2) The action the Agent takes to produce output,  $a$ : The most straightforward interpretation is that the Agent’s action is effort. This interpretation may be reasonably accurate in the sharecropping context and for low-level workers in large organizations. For a CEO, however, one should think of “effort” not in terms of hours worked but rather

in terms of paying attention to stakeholders' interests—for example, does the CEO take actions that increase shareholder value (versus taking actions that indulge pet projects)? Later in this note we will consider “multi-task” situations, in which the Agent can take some actions that help the Principal but also others that hurt. For example, the Agent might increase current earnings in two ways: by working hard to increase sales and cut production costs, but also by cutting R&D and marketing expenses, thereby hurting future earnings.

(3) Events beyond the Agent's control,  $\varepsilon$ : In the sharecropping context, one event beyond the Agent's control is the weather. In the CEO context, “animal spirits” in the stock market are similarly beyond the Agent's control. For simplicity, we assume that the expected value of  $\varepsilon$  is zero.

To keep the exposition simple, we will make a very specific assumption: the production function is  $y = a + \varepsilon$ .<sup>2</sup>

### *B. Contracts*

We will focus on contracts in which the Agent's total compensation for the period of the contract, denoted by  $w$ , is a linear function of output:  $w = s + b \cdot y$ . In such a contract,  $s$  can be thought of as salary and  $b$  as the Agent's bonus rate (so that the Agent's bonus is  $b \cdot y$ ). We sometimes call  $w$  the Agent's “wage,” but this should be understood to mean total compensation, not an hourly wage.

Linear contracts are simple to analyze, are observed in some real-world settings, and have an appealing property: they create uniform incentives, in the following sense. Think of output,  $y$ , as aggregate output over (say) a year, but think of the Agent as taking lots of little actions over the course of the year—such as one per day. A non-linear contract may create unintended or unhelpful incentives over the course of the year, depending on how the Agent has done so far.

---

<sup>2</sup> Some readers may wonder about the units in this production function: how can hours of effort plus inches of rain equal bushels of corn? As an antidote to this concern, consider the slightly more general production function  $y = g \cdot a + h \cdot \varepsilon$ , where  $g$  is the number of bushels of corn produced per hour of effort and  $h$  similarly translates inches of rain into bushels of corn. We have simply set  $g = h = 1$  for notational convenience.

As an example of unintended or unhelpful incentives, consider a contract that pays a low wage if output is below a target level and a high wage if output meets or exceeds the target level—such as \$50,000 for the year if output is below 1000 widgets, but \$100,000 for the year if output meets or exceeds 1000 widgets. (In this note, we ignore future considerations such as the Agent’s reputation in the labor market or promotion prospects in the firm.) Given this contract, once the Agent reaches the target level, he or she will stop working; also, if the end of the year draws near and the Agent is still far from reaching the target level, then he or she will stop working. Alternatively, if the Agent were paid (say) a salary of  $s = \$50,000$  and a bonus rate of  $b = \$50/\text{widget}$  then the Agent would earn \$100,000 for producing 1000 widgets but would have constant incentives regardless of performance to date: every extra widget earns the Agent \$50.

Note that a stock option creates uniform incentives on the upside, in its linear portion, but potentially unintended or unhelpful incentives if it is underwater (or even nearly so). If the option is severely underwater then there are essentially no incentives, because the Agent’s payoff is constant (at zero). More perversely, when the option is at the money, the Agent’s payoff is convex (flat below and linear above), which creates an incentive for risk-taking behavior. Similar incentives and behavior have been documented in several settings, including high-risk portfolio choices by managers of ostensibly conservative mutual funds (see Chevalier and Ellison, 1997).

### *C. Payoffs*

The Principal receives the Agent’s total contribution to firm value,  $y$ , but has to pay the Agent’s wage,  $w$ , so the Principal’s payoff (or “profit”) is the difference between the value received and the wage paid:  $\pi = y - w$ . For simplicity, we assume Principal is risk neutral—that is, the Principal simply wants to maximize the expected payoff, namely  $E(y - w)$ , where the notation  $E(x)$  denotes the expected value of the random variable  $x$ .

The Agent receives the wage  $w$  but has to take a costly action (*e.g.*, supply effort) to produce any output. Let  $c(a)$  be the dollar amount necessary to compensate the Agent for taking a particular action,  $a$ . Think of  $a = 0$  as no action at all, so  $c(0) = 0$ . (A bit more precisely, think of the action  $a = 0$  as the action the Agent would take without any financial inducements, hence  $c(0) = 0$ .)

The Agent’s payoff (or “utility”) is the difference between the wage received and the cost of the action taken:  $U = w - c(a)$ . In this note (and throughout this course), we

will focus on the simple case in which the Agent is risk-neutral—that is, the Agent simply wants to maximize the expected payoff  $E(w) - c(a)$ . (Since  $\varepsilon$  is the only uncertainty in the model and does not affect the Agent's cost function, no expectation is necessary in the second term of the Agent's payoff.)

We will assume that the Agent's cost function has the (intuitive) shape shown in Figure 1. There are really two assumptions being made here: (1) bigger actions are more costly (*i.e.*, the cost function  $c(a)$  is increasing), and (2) a small increase in the Agent's action is more costly starting from a big action than starting from a small action (*i.e.*, the cost function  $c(a)$  is convex—or, equivalently, the marginal cost of actions is increasing). The latter assumption implies that an extra five hours of work per week is tougher for someone currently working 80 hours a week than for someone currently working 40.

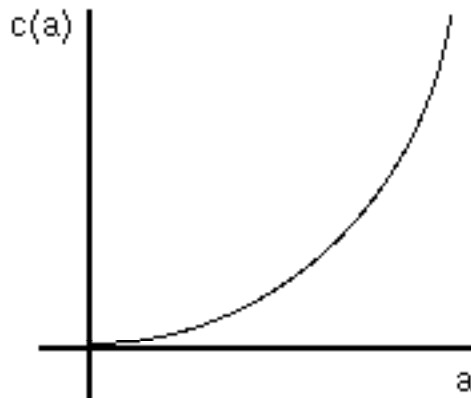


Figure 1

#### D. The Timing of Events

Putting all the model's elements together yields the following timing of events:

1. The Principal and the Agent sign a compensation contract ( $w = s + by$ ).
2. The Agent chooses an action ( $a$ ), but the Principal cannot observe this choice.
3. Events beyond the Agent's control ( $\varepsilon$ ) occur.
4. Together, the action and the noise determine the Agent's output ( $y$ ).
5. Output is observed by the Principal and the Agent (and by a Court, if necessary).

6. The Agent receives the compensation specified by the contract.

We turn next to analyzing this basic Principal-Agent model.

### 3. A Risk-Neutral Agent's Response to a Linear Contract

A risk-neutral Agent wants to choose the action that maximizes the expected value of the payoff  $w - c(a)$ . Since  $w = s + b \cdot y$  and  $y = a + \varepsilon$ , the Agent wants to maximize the expected value of  $s + b(a + \varepsilon) - c(a)$ . Since the only uncertainty in the Agent's payoff arises from the productivity shock  $\varepsilon$ , and since  $E(\varepsilon) = 0$ , the Agent's problem is

$$\max_a s + b \cdot a - c(a).$$

Starting from an arbitrary action  $a_0$ , the marginal benefit of choosing a slightly higher action is  $b$ , and the marginal cost of choosing a slightly higher action is the slope of the cost function at  $a_0$ , denoted  $c'(a_0)$ . Thus, if  $c'(a_0)$  is less than  $b$  then it is optimal for the Agent to choose a higher action than  $a_0$ , whereas if  $c'(a_0)$  is greater than  $b$  then it is optimal for the Agent to choose a lower action than  $a_0$ . Therefore, the optimal action for a risk-neutral Agent to choose in response to a contract with slope  $b$  is the action at which the slope of the cost function equals  $b$ , as shown in Figure 2.

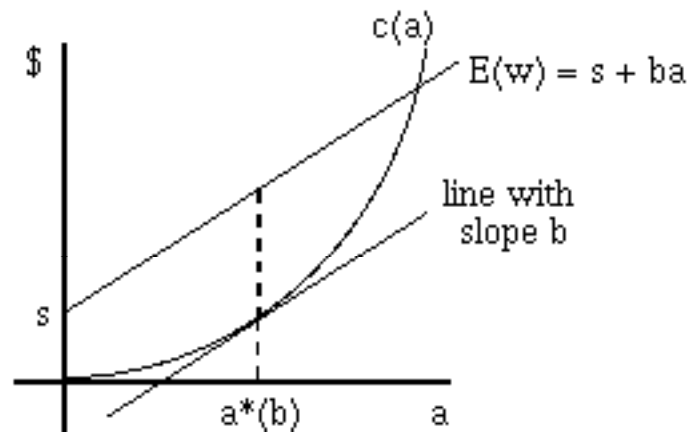


Figure 2

As an illustration, in the simple case where the cost function is the quadratic  $c(a) = \frac{1}{2} a^2$ , we have that the marginal cost function is  $c'(a) = a$  and so the Agent's optimal action in response to a contract with slope  $b$  is  $a^*(b) = b$ .<sup>3</sup>

Because the slope of  $c(a)$  increases as the Agent's action increases, a larger value of  $b$  will increase the Agent's optimal action,  $a^*(b)$ . That is, *steeper slopes create stronger incentives*. This result is intuitive: a profit-sharing plan that gives a worker 50% of the firm's profit is more likely to get the worker's attention than a plan paying 1%. But does this imply that the Principal would prefer a contract with a very large value of  $b$ ?

Definitely not. First, holding the Agent's salary constant, steeper slopes create stronger incentives but also give away more of the output to the Agent. For example if  $b = 1$  then the Principal is sure to lose money (assuming a positive salary) because the Agent gets all the output. But in some settings it may make sense to combine a steep slope with a *negative* salary: the slope does good things for the Agent's effort, the negative salary is interpreted as a fee the Agent pays the Principal for the privilege of getting the job, and the Agent expects to recover the fee by keeping most of the output. Most franchise contracts combine an up-front fee ( $s < 0$ ) and a steep slope ( $b$  near one) in this way. To pay such a fee up front, the Agent would need plenty of wealth or access to credit. There could also be disagreements about the appropriate size for such a fee: a Principal with a poor production process might try to claim otherwise to justify a high fee.

A second potential problem with a steeper slope is that it imposes more risk on the Agent. This risk doesn't matter if the Agent is risk-neutral, as assumed here, but would matter if the Agent were risk-averse. Interested readers can find more on the interplay between risk and incentives in an Appendix to this note (available upon request), but this material is *not* required elsewhere in this course.

---

<sup>3</sup> Note that in this model the salary  $s$  does not affect the Agent's optimal action: the slope of the contract completely determines the Agent's incentives; the salary is just a transfer of wealth that does not affect incentives.



#### 4. Lessons and Limits of the Basic Principal-Agent Model

The basic Principal-Agent model is extremely simple. On its own it tells us only a little about managing incentives. Among its lessons are (L1) contract shape matters (*e.g.*, linear versus kinked or jumped), (L2) steeper slopes create stronger incentives, and (L3) steeper slopes are not always better.

Beyond these simple lessons, the basic Principal-Agent model has two main values. First, the model gives us a language for expressing and analyzing abstract concepts such as reward, effort, and incentives in terms of more concrete model elements such as production ( $y = a + \varepsilon$ ), contract ( $w = s + by$ ), and payoffs ( $U = w - c(a)$  to the Agent and  $\pi = y - w$  to the Principal). Second, and probably more important, the model teaches us much by what it leaves out!

One instructive exercise is to compare this stick-figure model to the rich incentive issues in the case study on Lincoln Electric (Fast and Berg, 1975). For example, Lincoln pays not only piece rates, akin to the contracts  $w = s + by$  analyzed here, but also subjective bonuses. A second useful exercise is to compare this model to the compelling examples of failed (or at least agonizing) incentive plans recounted by Kerr (1975), in his classic article on “Rewarding A, While Hoping for B.” To begin to address both Lincoln Electric and Kerr’s examples, we turn next to the “multi-task” agency model, which emphasizes “get what you pay for” problems.

#### 5. Getting What You Pay For<sup>4</sup>

In the remainder of this note, we overturn an important but unremarked assumption in the basic Principal-Agent model: instead of assuming that the Agent’s contribution to firm value can be observed by the Principal and the Agent (and also by a court if necessary to enforce the compensation contract), we now assume that the only performance measures that can be observed by the Principal and the Agent (and a court) are distortionary, in a sense made formal below. We overturn this assumption from the basic model in order to analyze situations in the spirit of Kerr’s (1975) examples, such as the following.

---

<sup>4</sup> This section draws on Holmstrom and Milgrom (1991) and Gibbons (1998).

Business history is littered with firms that got what they paid for. At the H.J. Heinz Company, for example, division managers received bonuses only if earnings increased from the prior year. The managers delivered consistent earnings growth by manipulating the timing of shipments to customers and by prepaying for services not yet received. At Dun & Bradstreet, salespeople earned no commission unless the customer bought a larger subscription to the firm's credit-report services than in the previous year. In 1989, the company faced millions of dollars in lawsuits following charges that its salespeople deceived customers into buying larger subscriptions by fraudulently overstating their historical usage. In 1992, Sears abolished the commission plan in its auto-repair shops, which paid mechanics based on the profits from repairs authorized by customers. Mechanics misled customers into authorizing unnecessary repairs, leading California officials to prepare to close Sears' auto-repair business statewide.

In each of these cases, employees took actions to increase their compensation, but these actions were seemingly at the expense of long-run firm value. At Heinz, for example, prepaying for future services greatly reduced the firm's future flexibility, but the compensation system failed to address this issue. Similarly, at Dun & Bradstreet and Sears, although short-run profits increased with the increases in subscription sizes and auto repairs, the long-run harm done to the firms' reputations was significant (and plausibly much larger than the short-run benefit), but the compensation system again ignored the issue. Thus, in each of these cases, the cause of any dysfunctional behavior was not pay-for-performance *per se*, but rather pay-for-performance based on an inappropriate performance measure. (Baker, Gibbons, and Murphy, 1994)

To analyze these examples, we must stop calling  $y$  "output," as though it could easily be measured. This label is misleadingly simple: in the basic Principal-Agent model,  $y$  reflects everything the Principal cares about, except for wages (that is, the Principal's payoff is  $y - w$ ). Therefore, I henceforth call  $y$  the Agent's "total contribution to firm value," to emphasize that it encompasses all the Agent's actions (including mentoring, team production, and so on) and all the effects of these actions (both long- and short-run). In many settings, it is very difficult to measure synergies or sabotage across Agents and/or very difficult to predict the long-run consequences of an Agent's actions based on the observed short-run contribution. To analyze such settings, I impose the following assumption (for the rest of this course!): *no* contract based on  $y$  can be enforced in court, including but not limited to the linear contract  $w = s + by$ .

Of course, even when contracts based on  $y$  are not available, other contracts can be enforced in court. Such contracts are based on alternative performance measures—such as the number of units produced, with limited adjustment made for quality, timely delivery, and so on. Let  $p$  denote such an alternative performance measure; the wage contract might then be linear,  $w = s + bp$ . As in the basic Principal-Agent model, a large value of  $b$  will create strong incentives, but now the Agent's incentives are to produce a high value of  $p$ , not of  $y$ . But the firm does not directly benefit from increased realizations of measured performance,  $p$ ; rather, the firm benefits from increased

realizations of the Agent's total contribution to firm value,  $y$ . *The essence of the incentive problem is this divergence between the Agent's incentives to increase  $p$  and the firm's desire for increases in  $y$ .* It is no use creating strong incentives for the wrong actions. If attaching a large bonus rate  $b$  to the performance measure  $p$  would create strong but distorted incentives, then the optimal bonus rate may be quite small.

To begin to investigate these issues more formally, consider a simple extension of the basic Principal-Agent model:  $y = a + \varepsilon$  and  $p = a + \phi$ . In this case the contract  $w = s + bp$  creates incentives to increase  $p$  and the induced action also increases  $y$ . But now suppose that there are two kinds of actions (or "tasks") that the Agent can take,  $a_1$  and  $a_2$ . In this setting, the contract  $w = s + bp$  creates incentives that depend on the bonus rate  $b$  and on the way the actions  $a_1$  and  $a_2$  affect the performance measure  $p$ . For example, if  $y = a_1 + a_2$  and  $p = a_1$  then a contract based on  $p$  cannot create incentives for  $a_2$  and so misses this potential contribution to  $y$ . Alternatively, if  $y = a_1$  and  $p = a_1 + a_2$  then a contract based on  $p$  creates an incentive for the Agent to take action  $a_2$ , even though  $a_2$  is irrelevant to the Agent's total contribution to firm value. Finally, in an extreme case such as  $y = a_1 + \varepsilon$  and  $p = a_2 + \phi$ , the contract  $w = s + bp$  creates no value at all.

All of the examples above are special cases of the "multi-task" agency model described in the next section (which itself can easily be extended to include more actions and other enrichments). Compared to the basic Principal-Agent model, the chief departure in the multi-task model is the introduction of a fifth element of the model: in addition to the technology of production, the contract, the payoffs, and the timing of events, we now also require a technology of performance measurement.

## 6. The Multi-Task Agency Model<sup>5</sup>

Suppose that the technology of production is  $y = f_1a_1 + f_2a_2 + \varepsilon$ , the technology of performance measurement is  $p = g_1a_1 + g_2a_2 + \phi$ , the contract is  $w = s + bp$ , and the payoffs are  $\pi = y - w$  to the Principal and  $U = w - c(a_1, a_2)$  to the Agent. To keep things simple, assume that  $E(\varepsilon) = E(\phi) = 0$  and

$$c(a_1, a_2) = \frac{1}{2} a_1^2 + \frac{1}{2} a_2^2,$$

---

<sup>5</sup> This section draws on Feltham and Xie (1994), Datar, Kulp, and Lambert (2001), and Baker (2000).

but notice that the latter assumption rules out the potentially important case where the actions compete for the Agent's attention (*i.e.*, increasing the level of one action increases the marginal cost of the other).

The timing of events in this model is essentially the same as in the basic Principal-Agent model, except that it is modified to incorporate the new distinction here between  $y$  and  $p$ :

1. The Principal and the Agent sign a compensation contract  $w = s + bp$  (which we take to be linear for the reasons discussed above—namely, analytical simplicity and constant incentives).
2. The Agent chooses actions ( $a_1$  and  $a_2$ ) but the Principal cannot observe these choices.
3. Events beyond the Agent's control ( $\varepsilon$  and  $\phi$ ) occur.
4. The actions and the noise terms determine the Agent's total contribution to firm value ( $y$ ) and measured performance ( $p$ ).
5. Measured performance is observed by the Principal and the Agent (and by a Court, if necessary).<sup>6</sup>
6. The Agent receives the compensation specified by the contract, as a function of the realized value of  $p$ .

In this setting, the (risk-neutral) Agent chooses the actions  $a_1$  and  $a_2$  to maximize the expected payoff  $E(w) - c(a_1, a_2)$  and so must solve the following problem:

$$\max_{a_1, a_2} s + b \cdot (g_1 a_1 + g_2 a_2) - \frac{1}{2} a_1^2 - \frac{1}{2} a_2^2.$$

The Agent's optimal actions are therefore  $a_1^*(b) = g_1 b$  and  $a_2^*(b) = g_2 b$ , analogous to the special case of the basic Principal-Agent model where  $c(a) = \frac{1}{2} a^2$  and so  $a^*(b) = b$ .

To finish analyzing this model we must determine the optimal level of  $b$ . It turns out that there is a single value of  $b$  that is *efficient*. That is, if the parties were to sign a contract with a slope other than this efficient value then they could both be made better

---

<sup>6</sup> In this model, no one ever observes the Agent's total contribution, even though the Principal eventually receives the payoff  $y - w$ . See Lecture Note 3, on relational contracts and subjective performance assessment, for a more realistic approach to this issue.

off by switching to a contract with the efficient slope (and a new value of  $s$  chosen to make both parties better off after the switch).

To derive the efficient value of  $b$ , note that the Principal's expected payoff from the contract  $w = s + bp$  is

$$E(\pi) = E(y - w) = f_1 a_1^*(b) + f_2 a_2^*(b) - s - b g_1 a_1^*(b) - b g_2 a_2^*(b),$$

where the Agent's optimal actions in response to the contract have been included in the calculation of the Principal's expected payoff. Similarly, the Agent's expected payoff from the contract  $w = s + bp$  is

$$E(U) = E(w) - c(a_1, a_2) = s + b[g_1 a_1^*(b) + g_2 a_2^*(b)] - \frac{1}{2} a_1^*(b)^2 - \frac{1}{2} a_2^*(b)^2.$$

The sum of these expected payoffs is the expected total surplus,  $E(\pi + U)$ :

$$E(y) - c(a_1, a_2) = f_1 a_1^*(b) + f_2 a_2^*(b) - \frac{1}{2} a_1^*(b)^2 - \frac{1}{2} a_2^*(b)^2.$$

The efficient value of  $b$  is the value that maximizes this expected total surplus. A little math shows that this efficient value of  $b$  is

$$b^* = \frac{f_1 g_1 + f_2 g_2}{g_1^2 + g_2^2}.$$

This expression for the efficient value of  $b$  may not seem very helpful! Fortunately, it can be restated using Figure 3, which plots both the coefficients  $f_1$  and  $f_2$  from the technology of production and the coefficients  $g_1$  and  $g_2$  from the technology of performance measurement. (The figure is drawn assuming that  $f_1$ ,  $f_2$ ,  $g_1$ , and  $g_2$  are all positive, but this is not necessary.) This figure happens to represent a case in which  $g_1$  is larger than  $f_1$  but  $f_2$  is larger than  $g_2$ . In such a case, paying the Agent on  $p$  will create stronger incentives than the Principal wants for  $a_1$  but weaker incentives than the Principal wants for  $a_2$ .

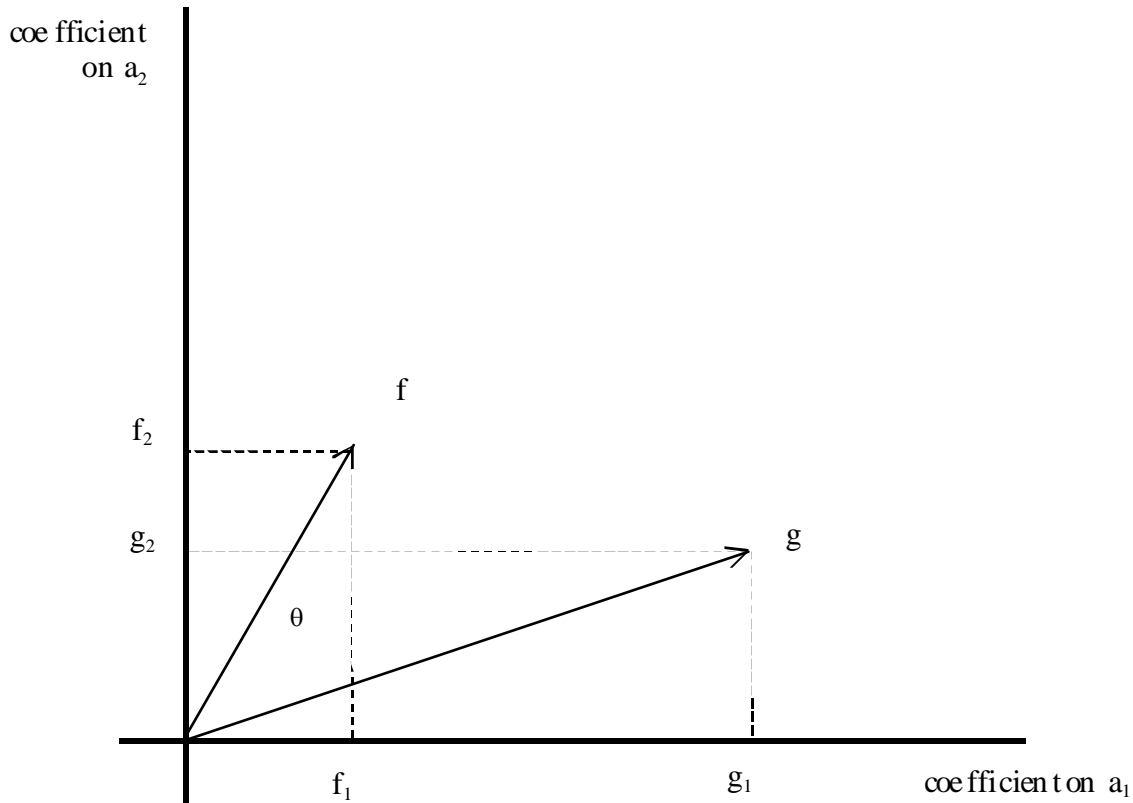


Figure 3

There are two important features in Figure 3: scale and alignment. To understand scale, imagine that  $g_1$  and  $g_2$  were both much larger than  $f_1$  and  $f_2$ . Then the Agent could greatly increase  $p$  by choosing high values of  $a_1$  and  $a_2$  but these actions would result in a much smaller value of  $y$  (ignoring the realizations of the noise terms for the moment). As a result, the efficient contract should put a small bonus rate on  $p$ , as will emerge below. To understand alignment, imagine first that the  $f$  and  $g$  vectors are closely aligned—they lie almost on top of one another (even if one is longer than the other). In this case the incentives created by paying on  $p$  are valuable for increasing  $y$ . Alternatively, imagine that the  $f$  and  $g$  vectors are badly aligned—for example, they might be orthogonal to each other (e.g.,  $f_1 = 0$  and  $g_2 = 0$ , so that  $y$  depends on only  $a_2$  and  $p$  depends on only  $a_1$ ). In this second case the incentives created by paying on  $p$  are useless for increasing  $y$ .

It turns out that scale and alignment are hiding in the expression for  $b^*$  derived above. With a little more math we can rewrite that efficient slope as

$$b^* = \frac{\sqrt{f_1^2 + f_2^2}}{\sqrt{g_1^2 + g_2^2}} \cos(\theta) ,$$

where  $\theta$  is the angle between the  $f$  and  $g$  vectors, as shown in Figure 3. Dusting off the Pythagorean Theorem reminds us that  $\sqrt{f_1^2 + f_2^2}$  is the length of the  $f$  vector, and correspondingly for  $\sqrt{g_1^2 + g_2^2}$ , so  $\sqrt{f_1^2 + f_2^2} / \sqrt{g_1^2 + g_2^2}$  reflects scaling. For example, if  $g$  is much longer than  $f$  (as considered above) then the efficient contract should put a small weight on  $p$ , as shown in this second expression for  $b^*$ . Recall also that  $\cos(0) = 1$  and  $\cos(90) = 0$ , so  $\cos(\theta)$  reflects alignment. For example, if the  $f$  and  $g$  vectors are closely aligned then  $\cos(\theta)$  is nearly 1 so  $b^*$  is large, whereas if the  $f$  and  $g$  vectors are almost orthogonal then  $\cos(\theta)$  is nearly 0 so  $b^*$  is small.

Let me note immediately that I have never seen a real incentive contract (or any other business document) that involves a cosine! In this sense, the formula for the efficient slope,  $b^*$ , illustrates the distinction between quantitative and qualitative predictions, as described in Section 1 above. That is, the two key ideas in the formula are not the Pythagorean Theorem and the cosine of an angle, but rather are scale and alignment – qualitative ideas that will persist as important factors in the determination of the efficient slope in many variations on and extensions of the simple model analyzed here.

## 7. Lessons and Limits of the Multi-Task Agency Model

The multi-task agency model delivers two important lessons, beyond the three lessons of the basic Principal-Agent model: (L4) objective performance measures typically cannot be used to create ideal incentives, and (L5) efficient bonus rates depend on scale and alignment. Furthermore, we can now dispel a persistent confusion about what makes a good performance measure, as follows.

One might be tempted to say that  $p$  is a good performance measure if it is highly correlated with  $y$ . But what determines the correlation between  $p$  and  $y$ ? That is, what would cause  $p$  and  $y$  to move together if we watched them over time? Given technologies such as  $y = f_1a_1 + f_2a_2 + \varepsilon$  and  $p = g_1a_1 + g_2a_2 + \phi$ , one important part of the answer involves the two variables we have not discussed thus far—the noise terms  $\varepsilon$  and  $\phi$ . Simply put,  $p$  and  $y$  will move together over time if the noise terms are highly correlated, regardless of the Agent's actions. For example, suppose that  $p$  is a division's accounting

earnings and  $y$  is the firm's stock price: both are hit by business-cycle variations (noise terms), but earnings reflect only the short-run effects of current actions while the stock price incorporates both short- and long-term effects of current actions. Thus, the earnings and the stock price might be highly correlated because of their noise terms, even though paying on the former creates distorted incentives for the latter (namely, incentives to ignore the long-run effects of current actions). Put more abstractly,  $y$  and  $p$  will be highly correlated if  $y = a_1 + \varepsilon$  and  $p = a_2 + \varepsilon$ , but in this case  $p$  is clearly a lousy performance measure. This argument leads to the central conclusion of the multi-task agency model:  *$p$  is a valuable performance measure if it induces valuable actions*, not if it is highly correlated with  $y$ . In short, alignment is more important than noise.

Although the multi-task agency is an important improvement on the basic Principal-Agent model, even the multi-task model clearly omits important issues. For example, in this model there would be no effort if  $b = 0$ , but in the real world we sometimes see great effort even if there is no direct link between pay and performance. In some contexts, such efforts may be inspired by intrinsic motivation, rather than the extrinsic motivation analyzed in economic models. But even economic models can begin to address incentives in settings that have no formal incentive contract such as  $w = s + bp$ , as follows.

In Lecture Note 2 (“Relational Contracts”), we move beyond the static models discussed here to a dynamic model of “relational contracts” that are enforced by the parties' concerns for their reputations rather than by the power of a court. In this dynamic model the Agent's contribution to firm value is observed by the Agent and the Principal but not by a court. In this sense, the Agent's performance can be subjectively assessed but not objectively measured. The model explores the extent to which the Principal can credibly promise to pay a bonus based on such a subjective assessment of the Agent's contribution to firm value. We have already mentioned that such subjective bonuses are used at Lincoln Electric, where they complement piece-rate compensation based on objective performance measures. In later cases we will see that such subjective bonuses are very important in many other settings, both literally (*e.g.*, in investment banking) and by analogy (*e.g.*, in promotion decisions).



## References

- Baker, George. 2002. "Distortion and Risk in Optimal Incentive Contracts." *Journal of Human Resources* 37: 728-51.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109:1125-56.
- Chevalier, Judith, and Glen Ellison. 1997. "Risk Taking by Mutual Funds as a Response to Incentives." *Journal of Political Economy* 105:1167-1200.
- Datar, Srikant, Susan Kulp, and Richard Lambert. 2001. "Balancing Performance Measures." *Journal of Accounting Research* 39: 75-92.
- Fast, Norman, and Norman Berg. 1975. "The Lincoln Electric Company." Harvard Business School Case #376-028.
- Feltham, Gerald and Jim Xie. 1994. "Performance Measure Congruity and Diverstiy in Multi-Task Principal/Agent Relations." *The Accounting Review* 69: 429-53.
- Gibbons, Robert. 1998. "Incentives in Organizations," *Journal of Economic Perspectives* 12: 115-32.
- Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7:24-52.
- Kerr, Steven. 1975. "On the Folly of Rewarding A, While Hoping for B." *Academy of Management Journal* 18:769-83.