MECS 570-1: Organizational Economics I

Michael Powell

Incentives in Organizations, Part I: Risk–Incentives, Limited Liability, Misaligned Measures

Incentives in Organizations

In order to move away from the Neoclassical view of a firm as a single individual pursuing a single objective, different strands of the literature have proposed different approaches. The first is what is now known as "team theory" (going back to the 1972 work of Marschak and Radner). Team-theoretic models focus on issues that arise when all members of an organization have the same preferences—these models typically impose constraints on information transmission between individuals and information processing by individuals and look at questions of task and attention allocation.

The alternative approach, which we will focus on in the majority of the course, asserts that different individuals within the organization have different preferences (that is, "People (i.e., individuals) have goals; collectivities of people do not." (Cyert and March, 1963: 30)) and explores the implications that these conflicts of interest have for firm behavior. In turn, this approach examines how limits to formal contracting restrict a firm's ability to resolve these conflicts of interest and how unresolved conflicts of interest determine how decisions are made. We will talk about several different sources of limits to formal contracts and the trade-offs they entail.

We will then think about how to motivate individuals in environments where formal contracts are either unavailable or they are so incomplete that they are of little use. Individuals can be motivated out of a desire to convince "the market" that they are intrinsically productive in the hopes that doing so will attract favorable outside offers in the future—that is, they are motivated by their own career concerns. Additionally, individuals may form long-term attachments with an organization. In such long-term relationships, goodwill can arise as an equilibrium phenomenon, and fear of shattering this goodwill can motivate individuals to perform well and to reward good performance.

Formal Incentive Contracts

We will look at several different sources of frictions that prevent individuals from writing contracts with each other that induce the same patterns of behavior they would choose if they were all acting as a single individual receiving all the payoffs. The first will be familiar from core microeconomics—individual actions chosen by an agent are not observed but determine the distribution of a verifiable performance measure. The agent is risk-averse, so writing a high-powered contract on that noisy performance measure subjects him to costly risk. As a result, there is a trade-off between incentive provision (and therefore the agent's effort choice) and inefficient risk allocation. This is the famous **risk–incentives trade-off**.

The second contracting friction that might arise is that an agent is either liquidity-constrained or is subject to a limited-liability constraint. As a result, the principal is unable to extract all the surplus the agent generates and must therefore provide the agent with **incentive rents** in order to motivate him. That is, offering the agent a higher-powered contract induces him to exert more effort and therefore increases the total size of the pie, but it also leaves the agent with a larger share of that pie. The principal then, in choosing a contract, chooses one that trades off the creation of surplus with her ability to extract that surplus. This is the **motivation–rent extraction trade-off.**

The third contracting friction that might arise is that the principal's objective simply cannot be written into a formal contract. Instead, the principal has to rely on imperfectly aligned performance measures. Increasing the strength of a formal contract that is based on imperfectly aligned performance measures may increase the agent's efforts toward the principal's objectives, but it may also motivate the agent to exert costly effort towards objectives that either hurt the principal or at least do not help the principal. Since the principal ultimately has to compensate the agent for whatever effort costs he incurs in order to get him to sign a contract to begin with, even the latter proves costly for the principal. Failure to account for the effects of using distorted performance measures is sometimes referred to as **the folly of rewarding A while hoping for B** (Kerr, 1975) or the **multi-task problem** (Holmstrom and Milgrom, 1991).

All three of these sources of contractual frictions lead to similar results—under the optimal contract, the agent chooses an action that is not jointly optimal from his and the principal's perspective. But in different applied settings, different assumptions regarding what is contractible and what is not are more or less plausible. As a result, it is useful to master at least elementary versions of models capturing these three sources of frictions, so that you are well-equipped to use them as building blocks.

In the elementary versions of models of these three contracting frictions that we will look at, the effort level that the Principal would induce if there were no contractual frictions would solve:

$$\max_e pe - \frac{c}{2}e^2,$$

so that $e^{FB} = p/c$. All three of these models yield *equilibrium* effort levels $e^* < e^{FB}$.

Risk-Incentives Trade-off

The exposition of an economic model usually begins with a rough (but accurate and mostly complete) description of the players, their preferences, and what they do in the course of the game. The exposition should also include a precise treatment of the timing, which includes spelling out who does what and when and on the basis of what information, and a description of the solution concept that will be used to derive predictions. Given the description of the economic environment, it is then useful to specify the program(s) that players are solving.

I will begin with a pretty general description of the standard principal-agent model, but I will shortly afterwards specialize the model quite a bit in order to focus on a single point—the risk–incentives trade-off.

Description There is a risk-neutral Principal (P) and a risk-averse Agent (A). The Agent chooses an effort level $e \in \mathbb{R}_+$ at a private cost of $c(e)$, with $c'', c' \geq 0$, and this effort level affects the distribution over output $y \in Y$, with y distributed according to cdf $F(\cdot|\, e)$. This output can be sold on the product market at price p. The Principal can write a contract $w \in W \subset \{w : Y \to \mathbb{R}\}$ that determines a transfer $w(y)$ that she is compelled to pay the Agent if output y is realized. The Agent has an outside option that provides utility $\bar{u}$ to the Agent and $\bar{\pi}$ to the Principal. If the outside option is not exercised, the Principal's and Agent's preferences are, respectively,

$$\begin{aligned} \Pi(w, e) &= \int_{y \in Y} (py - w(y))\, dF(y|\, e) = E_y[py - w|\, e] \\ U(w, e) &= \int_{y \in Y} u(w(y) - c(e))\, dF(y|\, e) = E_y[u(w - c(e))|\, e]. \end{aligned}$$

Timing The timing of the game is:

1. P offers A a contract $w(y)$, which is commonly observed.

2. A accepts the contract ($d = 1$) or rejects it ($d = 0$) and receives $\bar{u}$, and the game ends. This decision is commonly observed.

3. If A accepts the contract, A chooses effort level e and incurs cost $c(e)$. e is only observed by A.

4. Output y is drawn from distribution with cdf $F(\cdot|\, e)$. y is commonly observed.

5. P pays A an amount $w(y)$. This payment is commonly observed.

A couple remarks are in order at this point. First, behind the scenes, there is an implicit assumption that there is a third-party contract enforcer (a judge or arbitrator) who can costlessly detect when agreements have been broken and costlessly exact harsh punishments on the offender. Second, it is not necessarily important that e is unobserved by the Principal—given that the Principal takes no actions after the contract has been offered, as long as the contract cannot be conditioned directly on effort, the outcome of the game will be the same whether or not the Principal observes e. Put differently, one way of viewing the underlying source of moral-hazard problems is that contracts cannot be conditioned on relevant variables, not that the relevant variables are unobserved by the Principal. Unobservability is typically a convenient justification for why contracts cannot be conditioned on particular variables, but it is not necessary here. We will return to some of these issues when we discuss the Property Rights Theory of firm boundaries.

Solution Concept A **pure-strategy subgame-perfect equilibrium** is a contract $w^* \in W$, an acceptance decision $d^* : W \to \{0, 1\}$, and an effort choice $e^* : W \times \{0, 1\} \to \mathbb{R}_+$ such that, given the contract w^*, the Agent optimally chooses d^* and e^*, and given d^* and e^*, the Principal optimally offers contract w^*. We will say that the optimal contract **induces** effort e^*.

The Program The principal offers a contract $w \in W$ and proposes an effort level e in order to solve

$$\max_{w \in W, e \in \mathbb{R}_+} \int_{y \in Y} (py - w(y))\, dF(y|\, e)$$

subject to two constraints. The first constraint is that the agent actually prefers to choose effort level e rather than any other effort level $\hat{e}$. This is the standard **incentive-compatibility constraint**:

$$e \in \underset{\hat{e} \in \mathbb{R}_+}{\operatorname{argmax}} \int_{y \in Y} u(w(y) - c(\hat{e}))\, dF(y|\, \hat{e}).$$

The second constraint is that, given that the agent knows he will choose e if he accepts the contract, he prefers to accept the contract rather than to reject it and receive his outside utility $\bar{u}$. This is the standard **individual-rationality constraint** or **participation constraint**:

$$\int_{y \in Y} u(w(y) - c(e))\, dF(y|\, e) \geq \bar{u}.$$

CARA-Normal Case with Affine Contracts In order to establish a straightforward version of the risk-incentives trade-off, we will make a number of simplifying assumptions.

Assumption 1. The Agent has CARA preferences over wealth and effort costs, which are quadratic:

$$u\left(w\left(y\right)-c\left(e\right)\right)=-\exp\left\{-r\left(w\left(y\right)-\frac{c}{2}e^{2}\right)\right\},$$

and his outside option yields utility $-\exp\left\{-r\bar{u}\right\}$.

Assumption 2. Effort shifts the mean of a normally distributed random variable. That is, $y\sim N\left(e,\sigma^{2}\right)$.

Assumption 3. $W=\left\{w:Y\rightarrow\mathbb{R},w\left(y\right)=s+by\right\}$. That is, the contract space permits only affine contracts.

Discussion. In principle, there should be no exogenous restrictions on the functional form of $w\left(y\right)$. Applications, however, often restrict attention to affine contracts: $w\left(y\right)=s+by$. In many environments, an optimal contract does not exist if the contracting space is sufficiently rich, and situations in which the agent chooses the first-best level of effort, and the principal receives all the surplus can be arbitrarily approximated with a sequence of sufficiently perverse contracts (Mirrlees, 1974; Moroni and Swinkels, 2014). In contrast, the optimal affine contract often results in an effort choice that is lower than the first-best effort level, and the principal receives a lower payoff.

There are then at least three ways to view the exercise of solving for the optimal affine contract.

1. From an applied perspective, many pay-for-performance contracts in the world are affine in the relevant performance measure—franchisees pay a franchise fee and receive a constant fraction of the revenues their store generates, windshield installers receive a base wage and a constant piece rate, fruit pickers are paid per kilogram of fruit they pick. And so given that many practitioners seem to restrict attention to this class of contracts, why don't we just make sure they are doing what they do optimally? Put

differently, we can brush aside global optimality on purely pragmatic grounds.

2. Many pay-for-performance contracts in the world are affine in the relevant performance measure. Our models are either too rich or not rich enough in a certain sense and therefore generate optimal contracts that are inconsistent with those we see in the world. Maybe the aspects that, in the world, lead practitioners to use affine contracts are orthogonal to the considerations we are focusing on, so that by restricting attention to the optimal affine contract, we can still say something about how real-world contracts ought to vary with changes in the underlying environment. This view presumes a more positive (as opposed to normative) role for the modeler and hopes that the theoretical analogue of the omitted variables bias is not too severe.

3. Who cares about second-best when first-best can be attained? If our models are pushing us toward complicated, non-linear contracts, then maybe our models are wrong. Instead, we should focus on writing down models that generate affine contracts as the optimal contract, and therefore we should think harder about what gives rise to them. (And indeed, steps have been made in this direction—see Holmstrom and Milgrom (1987), Diamond (1998) and, more recently, Carroll (2013) and Barron, Georgiadis, and Swinkels (2017)) This perspective will come back later in the course when we discuss the Property Rights Theory of firm boundaries.

Given the assumptions, for any contract $w(y) = s + by$, the income stream the agent receives is normally distributed with mean $s + be$ and variance $b^2\sigma^2$. His expected utility over monetary compensation is therefore a moment-generating function for a normally distributed random variable, (recall that if $X \sim N(\mu, \sigma^2)$, then $E[\exp\{tX\}] = \exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}$), so his preferences can be written as

$$E[-\exp\{-r(w(y) - c(e))\}] = -\exp\left\{-r(s + be) + \frac{r^2}{2}b^2\sigma^2 + r\frac{c}{2}e^2\right\}.$$

We can take a monotonic transformation of his utility function ($-\frac{1}{r}\log(-x)$) and represent his preferences as:

$$\begin{aligned} U(e,w) &= E[w(y)] - \frac{r}{2}Var(w(y)) - \frac{c}{2}e^2 \\ &= s + be - \frac{r}{2}b^2\sigma^2 - \frac{c}{2}e^2. \end{aligned}$$

The Principal's program is then

$$\max_{s,b,e} pe - (s + be)$$

subject to incentive-compatibility

$$e \in \underset{\hat{e}}{\text{argmax}}\, b\hat{e} - \frac{c}{2}\hat{e}^2$$

and individual-rationality

$$s + be - \frac{r}{2}b^2\sigma^2 - \frac{c}{2}e^2 \geq \bar{u}.$$

Solving this problem is then relatively straightforward. Given an affine contract $s + be$, the agent will choose an effort level $e(b)$ that satisfies his first-order conditions

$$e(b) = \frac{b}{c},$$

and the Principal will choose the value s to ensure that the agent's individual-rationality constraint holds with equality (for if it did not hold with equality, the Principal could reduce s, making herself better off without affecting the Agent's incentive-compatibility constraint, while still respecting the Agent's individual-rationality constraint). That is,

$$s + be(b) = \frac{c}{2}e(b)^2 + \frac{r}{2}b^2\sigma^2 + \bar{u}.$$

In other words, the Principal has to ensure that the Agent's total expected monetary compensation, $s + be(b)$, fully compensates him for his effort costs, the risk costs he has to bear if he accepts this contract, and his opportunity cost. Indirectly, then, the Principal bears these costs when designing an optimal contract.

The Principal's remaining problem is to choose the incentive slope b to solve

$$\max_b pe(b) - \frac{c}{2}e(b)^2 - \frac{r}{2}b^2\sigma^2 - \bar{u}.$$

This is now an unconstrained problem with proper convexity assumptions, so the Principal's optimal choice of incentive slope solves her first-order condition

$$\begin{aligned} 0 &= pe'(b^*) - ce^*(b^*)e'(b^*) - rb^*\sigma^2 \\ &= \frac{p}{c} - c\frac{b^*}{c}\frac{1}{c} - rb^*\sigma^2 \end{aligned}$$

and therefore

$$b^* = \frac{p}{1 + rc\sigma^2}.$$

Also, given b^* and the individual-rationality constraint, we can back out s^*.

$$s^* = \bar{u} + \frac{1}{2}\left(rc\sigma^2 - 1\right)\frac{(b^*)^2}{c}.$$

Depending on the parameters, it may be the case that $s^* < 0$. That is, the Agent would have to pay the Principal if he accepts the job and does not produce anything.

In this setting, if the Principal could contract directly on effort, she would choose a contract that ensures that the Agent's individual-rationality constraint binds and therefore would solve

$$\max_e pe - \frac{c}{2}e^2,$$

so that

$$e^{FB} = \frac{p}{c}.$$

If the Principal wanted to implement this same level of effort using a contract on output, y, she would choose $b = p$ (since the Agent would choose $\frac{b}{c} = \frac{p}{c}$).

Why, in this setting, does the Principal not choose such a contract? Let us go back to the Principal's problem of choosing the incentive slope b.

$$\max_b pe(b) - \frac{c}{2}e(b)^2 - \frac{r}{2}b^2\sigma^2 - \bar{u}$$

Many fundamental points in models in the Organizational Economics literature can be seen as a comparison of first-order losses or gains against second-order gains or losses. Suppose the Principal chooses $b = p$, and consider a marginal reduction in b away from this value. The change in the Principal's profits would be

$$\begin{aligned} & \frac{d}{db}\left(pe(b) - \frac{c}{2}e(b)^2 - \frac{r}{2}b^2\sigma^2\right)\Bigg|_{b=p} \\ = & \underbrace{\frac{d}{db}\left(pe(b) - \frac{c}{2}e(b)^2\right)\Bigg|_{b=p}}_{=0} - rp\sigma^2 < 0 \end{aligned}$$

This first term is zero, because $b = p$ in fact maximizes $pe(b) - \frac{c}{2}e(b)^2$ (since it induces the first-best level of effort). The second term is strictly negative. That is, relative to the contract that induces first-best effort, a reduction in the slope of the incentive contract yields a first-order gain resulting from a decrease in the risk costs the Agent bears, while it yields a second-order loss in terms of profits resulting from moving away from the effort level that maximizes revenues minus effort costs. The optimal contract balances the incentive benefits of higher-powered incentives with these risk costs.

This trade-off seems first-order in some settings (e.g., insurance contracts in health care markets, some types of sales contracts in industries in which individual sales are infrequent, large, and unpredictable) and for certain types of output. There are many other environments in which contracts provide less-than-first-best incentives, but the first-order reasons for these low-powered contracts seem completely different, and we will turn to these environments shortly.

Before doing so, it is worth pointing out that many models in this course will involve trade-offs that determine the optimal way of organizing a firm. In many of the settings these models examine, results that take the form of "X is organized according to Y, because player A is more risk-averse than player B" often seem intuitively unappealing. For example, suppose a model of hierarchies generated the result that less risk-averse individuals should be at the top of an organization, and more risk-averse individuals should be at the bottom. This sounds somewhat sensible—maybe richer individuals are better able to diversify their wealth, and they can therefore behave as if they are less risk averse with respect to the income stream they derive from a particular organization. But it sounds less appealing as a general rule for who should be assigned to what role in an organization—a model that predicts that more knowledgeable or more experienced workers should be assigned to higher positions seems more consistent with experience.

Finally, connecting this analysis back to the Neoclassical view of the firm, what does the risk-incentives trade-off imply about the firm's production set? Let Y^f denote the firm's **technological possibilities set**, in which the firm's input is labor costs C, and its expected output is y. This is the set of input-output vectors that would be feasible if there were no contracting frictions.

We can write

$$C = c(e) = \frac{c}{2}e^2,$$

and since expected output is just equal to the Agent's effort choice, we have that

$$y(C) = e = \left(\frac{2C}{c}\right)^{1/2}.$$

The technological possibilities set is therefore

$$Y^f = \{(y, -C) : y \leq y(C)\}.$$

We will now augment the technological possibilities set with the contractual considerations we have just derived. Because the Principal can increase s without bound, the **contract-augmented possibilities set** can be characterized by its frontier, which is the highest level of expected output the firm can produce for a given level of costs. If the Principal puts in place an optimal contract with incentive slope b (in which case the Agent's effort choice will be $e^*(b) = \frac{b}{c}$) and an s that pins the Agent to his individual-rationality constraint, the firm's profits are

$$p\frac{b}{c} - \frac{c}{2}\left(\frac{b}{c}\right)^2 - \frac{r}{2}\sigma^2 b^2 = p\frac{b}{c} - \frac{1}{2}\frac{1 + rc\sigma^2}{c}b^2.$$

Therefore, producing expected output $\frac{b}{c}$ costs the firm

$$C = \frac{1}{2}\frac{1 + rc\sigma^2}{c}b^2.$$

Finally, we can rearrange this equation to solve for the b such that the total costs to the Principal are C:

$$b = \left(\frac{2Cc}{1 + rc\sigma^2}\right)^{1/2}.$$

In this case, the firm produces expected output $y = \tilde{y}(C)$, which is given by

$$\tilde{y}(C) = \left(\frac{2C}{c}\right)^{1/2}\left(\frac{1}{1 + rc\sigma^2}\right)^{1/2} = \left(\frac{1}{1 + rc\sigma^2}\right)^{1/2} y(C).$$

The contract-augmented possibilities set is therefore

$$\tilde{Y}^f = \{(y, -C) : y \leq \tilde{y}(C)\}.$$

Because of contractual frictions, we have that $\tilde{Y}^f \subset Y^f$, and any change in the parameters of the model for which the divergence between e^{FB} and e^* grows (such as an increase in the Agent's risk aversion or an increase in output uncertainty) will tend to increase the difference between $y(C)$ and $\tilde{y}(C)$ and therefore will shrink the contract-augmented possibilities set.

In this elementary version of this model, the contract-augmented possibilities set is a convex set. More generally, given a production-possibilities set that is convex, it need not be the case that the contract-augmented possibilities set is convex.

Further Reading Many papers restrict attention to linear contracts, even in environments in which the optimal contract (if it exists) is not linear. Holmstrom and Milgrom (1987) examines an environment in which the principal and the agent have CARA preferences and the agent controls the drift of a Brownian motion for a finite time interval. An optimal contract conditions payments only on the value of the Brownian motion at the end of the time interval. Diamond (1998) considers an environment in which the agent can choose the mean of the output distribution as well as the entire distribution itself and shows (essentially by a convexification argument) that linear contracts are optimal. Carroll (2015) shows that linear contracts can be max-min optimal when the Principal is sufficiently uncertain about the class of actions the Agent can take.

A key comparative static of the risk–incentives moral-hazard model is that incentives are optimally weaker when there is more uncertainty in the mapping between effort and contractible output, but this comparative static is inconsistent with a body of empirical work suggesting that in more uncertain environments, agency contracts tend to involve higher-powered incentives. Prendergast (2002) resolves this discrepancy by arguing that in more uncertain environments, it is optimal to assign greater responsibility to the agent and to complement this greater responsibility with higher-powered incentives. Holding responsibilities fixed, the standard risk–incentives tradeoff would arise, but the empirical studies that fail to find this relationship do not control for workers' responsibilities. Raith (2003) argues that these empirical studies examine the relationship between the risk the firm faces and the strength of the agent's incentives, while the theory is about the relationship between the risk the *agent* faces and his incentives. For an examination of several channels through which uncertainty can impact an agent's incentives, see Rantakari (2008).

Limited Liability

We saw in the previous model that the optimal contract sometimes involved upfront payments from the Agent to the Principal. To the extent that the Agent is unable to afford such payments (or legal restrictions prohibit such payments), the Principal will not be able to extract all the surplus that the Agent creates. Further, in order to extract surplus from the Agent, the Principal may have to put in place contracts that reduce the total surplus created. In equilibrium, the Principal may therefore offer a contract that induces effort below the first-best.

Description Again, there is a risk-neutral Principal (P). There is also a **risk-neutral** Agent (A). The Agent chooses an effort level $e \in \mathbb{R}_+$ at a private cost of $c(e)$, with $c'', c' > 0$, and this effort level affects the distribution over outputs $y \in Y$, with y distributed according to cdf $F(\cdot|\, e)$. These outputs can be sold on the product market for price p. The Principal can write a contract $w \in W \subset \{w : Y \to \mathbb{R}, w(y) \geq \underline{w} \text{ for all } y\}$ that determines a transfer $w(y)$ that she is compelled to pay the Agent if output y is realized. The Agent has an outside option that provides utility $\bar{u}$ to the Agent and $\bar{\pi}$ to the Principal. If the outside option is not exercised, the Principal's and Agent's preferences are, respectively,

$$
\begin{aligned}
\Pi(w, e) &= \int_{y \in Y} (py - w(y))\, dF(y|\, e) = E_y[py - w|\, e] \\
U(w, e) &= \int_{y \in Y} (w(y) - c(e))\, dF(y|\, e) = E_y[w - c(e)|\, e].
\end{aligned}
$$

There are two differences between this model and the model in the previous subsection. The first difference is that the Agent is risk-neutral (so that absent any other changes, the equilibrium contract would induce first-best effort). The second difference is that the wage payment from the Principal to the Agent has to exceed, for each realization of output, a value $\underline{w}$. Depending on the setting, this constraint is described as a liquidity constraint or a limited-liability constraint. In repeated settings, it is more naturally thought of as the latter—due to legal restrictions, the Agent cannot be legally compelled to make a transfer (larger than $-\underline{w}$) to the Principal. In static settings, either interpretation may be sensible depending on the particular application—if the Agent is a fruit picker, for instance, he may not have much liquid wealth that he can use to pay the Principal.

Timing The timing of the game is exactly the same as before.

1. P offers A a contract $w(y)$, which is commonly observed.

2. A accepts the contract $(d = 1)$ or rejects it $(d = 0)$ and receives $\bar{u}$, and the game ends. This decision is commonly observed.

3. If A accepts the contract, A chooses effort level e and incurs cost $c(e)$. e is only observed by A.

4. Output y is drawn from distribution with cdf $F(\cdot|\, e)$. y is commonly observed.

5. P pays A an amount $w(y)$. This payment is commonly observed.

Equilibrium The solution concept is the same as before. A **pure-strategy subgame-perfect equilibrium** is a contract $w^* \in W$, an acceptance decision $d^* : W \to \{0, 1\}$, and an effort choice $e^* : W \times \{0, 1\} \to \mathbb{R}_+$ such that given the contract w^*, the Agent optimally chooses d^* and e^*, and given d^* and e^*, the Principal optimally offers contract w^*. We will say that the optimal contract induces effort e^*.

The Program The principal offers a contract $w \in W$ and proposes an effort level e in order to solve

$$\max_{w \in W, e} \int_{y \in Y} (py - w(y))\, dF(y|\, e)$$

subject to three constraints: the incentive-compatibility constraint

$$e \in \underset{\hat{e} \in \mathbb{R}_+}{\operatorname{argmax}} \int_{y \in Y} (w(y) - c(\hat{e}))\, dF(y|\, \hat{e}),$$

the individual-rationality constraint

$$\int_{y \in Y} (w(y) - c(e))\, dF(y|\, e) \geq \bar{u},$$

and the limited-liability constraint

$$w(y) \geq \underline{w} \text{ for all } y.$$

Binary-Output Case Jewitt, Kadan, and Swinkels (2008) solves for the optimal contract in the general environment above (and even allows for agent risk aversion). Here, I will instead focus on an elementary case that highlights the main trade-off.

Assumption 1. Output is $y \in \{0, 1\}$, and given effort e, its distribution satisfies $\Pr[y = 1|\, e] = e$.

Assumption 2. The agent's costs have a non-negative third derivative: $c''' \geq 0$, and they satisfy conditions that ensure an interior solution: $c'(0) = 0$ and $c'(1) = +\infty$. Or for comparison across models in this module, $c(e) = \frac{c}{2}e^2$, where $p \leq c$ to ensure that $e^{FB} < 1$.

Finally, we can restrict attention to affine, nondecreasing contracts

$$\begin{aligned} W &= \{w(y) = (1 - y)\, w_0 + y w_1, w_0, w_1 \geq 0\} \\ &= \{w(y) = s + by, s \geq \underline{w}, b \geq 0\}. \end{aligned}$$

When output is binary, this restriction to affine contracts is without loss of generality. Also, the restriction to nondecreasing contracts is not restrictive (i.e., any optimal contract of a relaxed problem in which we do not impose that contracts are nondecreasing will also be the solution to the full problem). This result is something that needs to be shown and is not in general true, but in this case, it is straightforward.

In principal-agent models, it is often useful to break the problem down into two steps. The first step takes a target effort level, e, as given and solves for the set of cost-minimizing contracts implementing effort level e. Any cost-minimizing contract implementing effort level e results in an expected cost of $C(e)$ to the principal. The second step takes the function $C(\cdot)$ as given and solves for the optimal effort choice.

In general, the cost-minimization problem tends to be a well-behaved convex-optimization problem, since (even if the agent is risk-averse) the objective function is weakly concave, and the constraint set is a convex set (since given an effort level e, the individual-rationality constraint and the limited-liability constraint define convex sets, and each incentive constraint ruling out effort level $\hat{e} \neq e$ also defines a convex set, and the intersection of convex sets is itself a convex set). The resulting cost function $C(\cdot)$ need not have nice properties, however, so the second step of the optimization problem is only well-behaved under restrictive assumptions. In the present case, assumptions 1 and 2 ensure that the second step of the optimization problem is well-behaved.

Cost-Minimization Problem Given an effort level e, the cost-minimization problem is given by

$$C(e, \bar{u}, \underline{w}) = \min_{s,b} s + be$$

subject to the agent's incentive-compatibility constraint

$$e \in \operatorname*{argmax}_{\hat{e}} \{s + b\hat{e} - c(\hat{e})\},$$

his individual-rationality constraint

$$s + be - c(e) \geq \bar{u},$$

and the limited-liability constraint

$$s \geq \underline{w}.$$

I will denote a **cost-minimizing contract implementing effort level** e by (s_e^*, b_e^*).

The first step in solving this problem is to notice that the agent's incentive-compatibility constraint implies that any cost-minimizing contract implementing effort level e must have $b_e^* = c'(e)$.

If there were no limited-liability constraint, the principal would choose s_e^* to extract the agent's surplus. That is, given $b = b_e^*$, s would solve

$$s + b_e^* e = \bar{u} + c(e).$$

That is, s would ensure that the agent's expected compensation exactly equals his expected effort costs plus his opportunity cost. The resulting s, however, may not satisfy the limited-liability constraint. The question then is: given $\bar{u}$ and $\underline{w}$, for what effort levels e is the principal able to extract all the agent's surplus (i.e., for what effort levels does the limited-liability constraint not bind?), and for what effort levels is she unable to do so? Figure 1 below shows cost-minimizing contracts for effort levels e_1 and e_2. Any contract can be represented as a line in this figure, where the line represents the expected pay the agent will receive given an effort level e. The cost-minimizing contract for effort level e_1 is tangent to the $\bar{u} + c(e)$ curve at e_1 and its intercept is $s_{e_1}^*$. Similarly for e_2. Both $s_{e_1}^*$ and $s_{e_2}^*$ are greater than $\underline{w}$, which implies that for such effort levels, the limited-liability constraint is not binding.

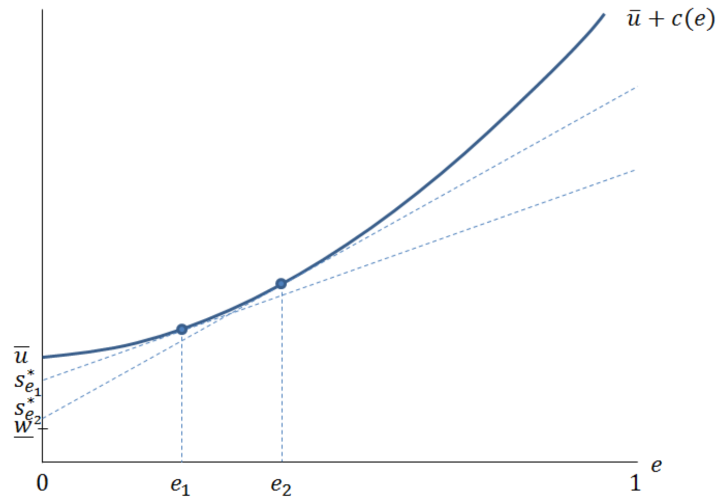

Figure 1

For effort sufficiently high, the limited-liability constraint will be binding in a cost-minimizing contract, and it will be binding for all higher effort levels. Define the threshold $\bar{e}(\bar{u}, \underline{w})$ to be the effort level such that for all $e \geq \bar{e}(\bar{u}, \underline{w})$, $s_e^* = \underline{w}$. Figure 2 illustrates that $\bar{e}(\bar{u}, \underline{w})$ is the effort level at which the contract tangent to the $\bar{u} + c(e)$ curve at $\bar{e}(\bar{u}, \underline{w})$ intersects the vertical axis at exactly $\underline{w}$. That is, $\bar{e}(\bar{u}, \underline{w})$ solves

$$c'(\bar{e}(\bar{u}, \underline{w})) = \frac{\bar{u} + c(\bar{e}(\bar{u}, \underline{w})) - \underline{w}}{\bar{e}(\bar{u}, \underline{w})}.$$

Figure 2 also illustrates that for all effort levels $e > \bar{e}(\bar{u}, \underline{w})$, the cost-minimizing contract involves giving the agent strictly positive surplus. That is, the cost to the principal of getting the agent to choose effort $e > \bar{e}(\bar{u}, \underline{w})$ is equal to the agent's opportunity costs $\bar{u}$ plus his effort costs $c(e)$ plus **incentive costs** $IC(e, \bar{u}, \underline{w})$.

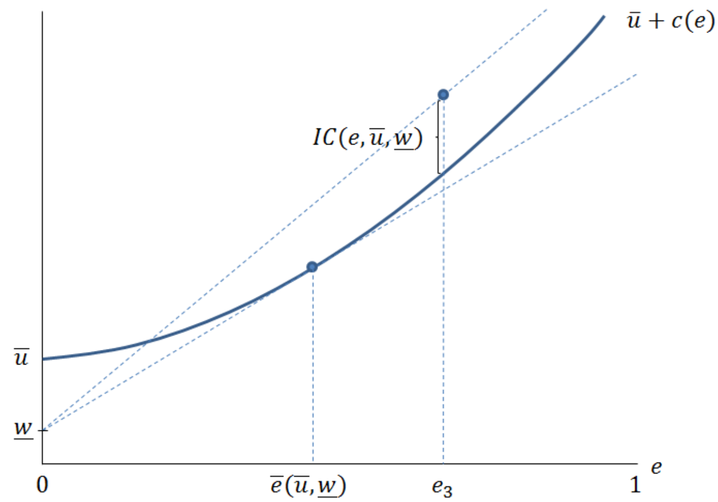

Figure 2

The incentive costs $IC(e, \bar{u}, \underline{w})$ are equal to the agent's expected compensation given effort choice e and cost-minimizing contract (s_e^*, b_e^*) minus his costs:

$$
\begin{aligned}
IC(e, \bar{u}, \underline{w}) &= \begin{cases} 0 & e \leq \bar{e}(\bar{u}, \underline{w}) \\ \underline{w} + c'(e)\,e - c(e) - \bar{u} & e \geq \bar{e}(\bar{u}, \underline{w}) \end{cases} \\
&= \max\{0, \underline{w} + c'(e)\,e - c(e) - \bar{u}\}
\end{aligned}
$$

where I used the fact that for $e \geq \bar{e}(\bar{u}, \underline{w})$, $s_e^* = \underline{w}$ and $b_e^* = c'(e)$. This incentive-cost function $IC(\cdot, \bar{u}, \underline{w})$ is the key object that captures the main contracting friction in this model. I will sometimes refer to $IC(e, \bar{u}, \underline{w})$ as the **incentive rents** required to get the agent to choose effort level e. Putting these results together, we see that

$$C(e, \bar{u}, \underline{w}) = \bar{u} + c(e) + IC(e, \bar{u}, \underline{w}).$$

That is, the principal's total costs of implementing effort level e are the sum of the agent's costs plus the incentive rents required to get the agent to choose effort level e.

Since $IC(e, \bar{u}, \underline{w})$ is the main object of interest in this model, I will describe some of its properties. First, it is continuous in e (including, in particular, at $e = \bar{e}(\bar{u}, \underline{w})$). Next, $\bar{e}(\bar{u}, \underline{w})$ and $IC(e, \bar{u}, \underline{w})$ depend on $(\bar{u}, \underline{w})$ only inasmuch as $(\bar{u}, \underline{w})$ determines $\bar{u} - \underline{w}$, so I will abuse notation and write these expressions as $\bar{e}(\bar{u} - \underline{w})$ and $IC(e, \bar{u} - \underline{w})$. Also, given that $c'' > 0$, IC is increasing in e (since $\underline{w} + c'(e)e - c(e) - \underline{u}$ is strictly increasing in e, and IC is just the max of this expression and zero). Further, given that $c''' \geq 0$, IC is convex in e. For $e \geq \bar{e}(\bar{u} - \underline{w})$, this property follows, because

$$\frac{\partial^2}{\partial e^2} IC = c''(e) + c'''(e)\, e \geq 0.$$

And again, since IC is the max of two convex functions, it is also a convex function. Finally, since $IC(\cdot, \bar{u} - \underline{w})$ is flat when $e \leq \bar{e}(\bar{u} - \underline{w})$ and it is strictly increasing (with slope independent of $\bar{u} - \underline{w}$) when $e \geq \bar{e}(\bar{u} - \underline{w})$, the slope of IC with respect to e is (weakly) decreasing in $\bar{u} - \underline{w}$, since $\bar{e}(\bar{u} - \underline{w})$ is increasing in $\bar{u} - \underline{w}$. That is, $IC(e, \bar{u} - \underline{w})$ satisfies decreasing differences in $(e, \bar{u} - \underline{w})$.

Motivation-Rent Extraction Trade-off The second step of the optimization problem takes as given the function

$$C(e, \bar{u} - \underline{w}) = \bar{u} + c(e) + IC(e, \bar{u} - \underline{w})$$

and solves for the optimal effort choice by the principal:

$$\begin{aligned} & \max_e pe - C(e, \bar{u} - \underline{w}) \\ = \quad & \max_e pe - \bar{u} - c(e) - IC(e, \bar{u} - \underline{w}). \end{aligned}$$

Note that total surplus is given by $pe - \bar{u} - c(e)$, which is therefore maximized at $e = e^{FB}$ (which, if $c(e) = ce^2/2$, then $e^{FB} = p/c$). Figure 3 below depicts the principal's expected benefit line pe, and her expected costs of implementing effort e at minimum cost, $C(e, \bar{u} - \underline{w})$. The first-best effort level, e^{FB} maximizes the difference between pe and $\bar{u} + c(e)$, while the equilibrium effort level e^* maximizes the difference between pe and $C(e, \bar{u} - \underline{w})$.

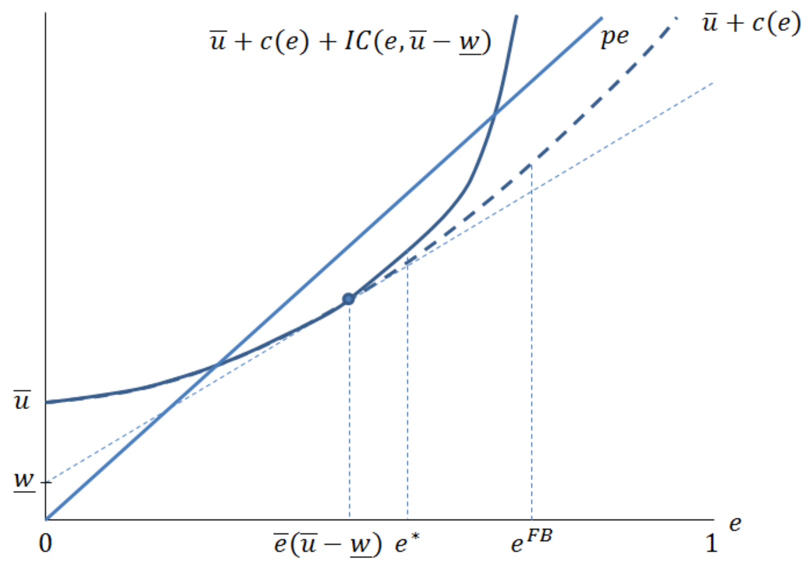

Figure 3

If $c(e) = ce^2/2$, we can solve explicitly for $\bar{e}(\bar{u} - \underline{w})$ and for $IC(e, \bar{u} - \underline{w})$ when $e > \bar{e}(\bar{u} - \underline{w})$. In particular,

$$\bar{e}(\bar{u} - \underline{w}) = \left(\frac{2(\bar{u} - \underline{w})}{c}\right)^{1/2}$$

and when $e > \bar{e}(\bar{u} - \underline{w})$,

$$IC(e, \bar{u} - \underline{w}) = \underline{w} + \frac{1}{2}ce^2 - \bar{u}.$$

If $\underline{w} < 0$ and p is sufficiently small, we can have $e^* = e^{FB}$ (i.e., these are the conditions required to ensure that the limited-liability constraint is not binding for the cost-minimizing contract implementing $e = e^{FB}$). If p is sufficiently large relative to $\bar{u} - \underline{w}$, we will have $e^* = \frac{1}{2}\frac{p}{c} = \frac{1}{2}e^{FB}$. For p somewhere in between, we will have $e^* = \bar{e}(\bar{u} - \underline{w}) < e^{FB}$. In particular, $C(e, \bar{u} - \underline{w})$ is kinked at this point.

As in the risk–incentives model, we can illustrate through a partial characterization why (and when) effort is less-than first-best. Since we know that e^{FB} maximizes $pe - \bar{u} - c(e)$, we therefore have that

$$\frac{d}{de}\left[pe - \bar{u} - c(e) - IC(e, \bar{u} - \underline{w})\right]_{e=e^{FB}} = -\frac{\partial}{\partial e} IC\left(e^{FB}, \bar{u} - \underline{w}\right) \leq 0,$$

with strict inequality if the limited-liability constraint binds at the cost-minimizing contract implementing e^{FB}. This means that, even though e^{FB} maximizes total surplus, if the principal has to provide the agent with rents at the margin, she may choose to implement a lower effort level. Reducing the effort level away from e^{FB} leads to second-order losses in terms of total surplus, but it leads to first-order gains in profits for the principal. In this model, there is a tension between total-surplus creation and rent extraction, which yields less-than-first-best effort in equilibrium.

In my view, liquidity constraints are extremely important and are probably one of the main reasons for why many jobs do not involve first-best incentives. The Vickrey-Clarke-Groves logic that first-best outcomes can be obtained if the firm transfers the entire profit stream to each of its members in exchange for a large up-front payment seems simultaneously compelling, trivial, and obviously impracticable. In for-profit firms, in order to make it worthwhile to transfer a large enough share of the profit stream to an individual worker to significantly affect his incentives, the firm would require a large up-front transfer that most workers cannot afford to pay. It is therefore not surprising that we do not see most workers' compensation tied directly to the firm's overall profits in a meaningful way. One implication of this logic is that firms have to find alternative instruments to use as performance measures, which we will turn to next. In principle, models in which firms do not motivate their workers by writing contracts directly on profits should include assumptions under which the firm optimally chooses not to write contracts directly on profits, but they almost never do.

Exercise. *Let* $\Delta \equiv (2c(\bar{u} - \underline{w}))^{1/2}$. *Part 1: Show that when* $p \leq \Delta$, *the contract-augmented possibilities set is* $\tilde{Y}^f = \left\{(y, -C) : y \leq \left(\frac{2C}{c}\right)^{1/2}\left(\frac{C-\underline{w}}{2C}\right)^{1/2}\right\}$. *Part 2: Show that when* $p \geq 2\Delta$, *the contract-augmented possibilities set is* $\tilde{Y}^f = \left\{(y, -C) : y \leq \left(\frac{2C}{c}\right)^{1/2}\right\}$. *Part 3: Solve for* $\tilde{Y}^f$ *for* $\Delta < p < 2\Delta$. *(This part is somewhat more complicated.) Part 4: In this model, the contract-augmented possibilities depend on the equilibrium price level, which implies that in a competitive-equilibrium framework, the firm's production possibilities are endogenous to the equilibrium. This was not the case for the risk-incentives trade-off model. If we define* $\tilde{Y}^f(p)$ *as the contract-augmented possibilities set given price level* p, *how does* $\tilde{Y}^f(p)$ *vary in* p? *(Note that since* $\tilde{Y}^f(p)$ *is a set, you will have to think about what it means for a set to vary in a parameter.)*

Exercise. *Holmstrom (1979) shows that in the risk-incentives model in the previous subsection, if there is a costless additional performance measure* m *that is informative about* e, *then an optimal formal contract should always put some weight on* m *unless* y *is a sufficient statistic for* y *and* m. *This is known as Holmstrom's "informativeness principle" and suggests that optimal contracts should always be extremely sensitive to the details of the environment the contract is written in. Suppose instead that the agent is risk-neutral but liquidity-constrained, and suppose there is a performance measure* $m \in \{0, 1\}$ *such that* $\Pr[m = 1 | e] = e$ *and conditional on* e, m *and* y *are independent. Suppose contracts of the form* $w(y, m) = s + b_y y + b_m m + b_{ym} ym$ *can be written but must satisfy* $w(y, m) \geq \bar{w}$ *for each realization of* (y, m). *Is it again always the case that* $b_m \neq 0$ *and/or* $b_{ym} \neq 0$?

Further Reading Jewitt, Kadan, and Swinkels (2008) derive optimal contracts in a broad class of environments with risk-averse agents and bounded payments (in either direction). Chaigneau, Edmans, and Gottlieb (2015) provide necessary and sufficient conditions for additional informative signals to have strictly positive value to the Principal. Wu (2015) shows that firms' contract-augmented possibilities sets are endogenous to the competitive environment they face when their workers are subject to limited-liability constraints.

Multiple Tasks and Misaligned Performance Measures

In the previous two models, the Principal cared about output, and output, though a noisy measure of effort, was perfectly measurable. This assumption seems sensible when we think about overall firm profits (ignoring basically everything that accountants think about every day), but as we alluded to in the previous discussion, overall firm profits are too blunt of an instrument to use to motivate individual workers within the firm if they are liquidity-constrained. As a result, firms often try to motivate workers using more specific performance measures, but while these performance measures are informative about what actions workers are taking, they may be less useful as a description of how the workers' actions affect the objectives the firm cares about. And paying workers for what is measured may not get them to take actions that the firm cares about. This observation underpins the title of the famous 1975 paper by Steve Kerr called "On the Folly of Rewarding A, While Hoping for B."

As an example, think of a retail firm that hires an employee both to make sales and to provide customer service. It can be difficult to measure the quality of customer service that a particular employee provides, but it is easy to measure that employee's sales. Writing a contract that provides the employee with high-powered incentives directly on sales will get him to put a lot of effort into sales and very little effort into customer service. And in fact, he might only be able to put a lot of effort into sales by intentionally neglecting customer service. If the firm cares equally about both dimensions, it might be optimal not to offer high-powered incentives to begin with. This is what Holmström and Milgrom (1991) refers to as the "multitask problem." We will look at a model that captures some of this intuition, although not as directly as Holmström and Milgrom's model. The model we will look at builds upon Baker (1992, 2002) and Feltham and Xie (1994).

Description Again, there is a risk-neutral Principal (P) and a risk-neutral Agent (A). The Agent chooses an effort vector $e = (e_1, e_2) \in \mathcal{E} \subset \mathbb{R}_+^2$ at a cost of $\frac{c}{2}\left(e_1^2 + e_2^2\right)$. This effort vector affects the distribution of output $y \in \mathcal{Y} = \{0, 1\}$ and a performance measure

$m \in \mathcal{M} = \{0, 1\}$ as follows:

$$\begin{aligned} \Pr[y = 1|\, e] &= f_1 e_1 + f_2 e_2 \\ \Pr[m = 1|\, e] &= g_1 e_1 + g_2 e_2, \end{aligned}$$

where it may be the case that $f = (f_1, f_2) \neq (g_1, g_2) = g$. Assume that $f_1^2 + f_2^2 = g_1^2 + g_2^2 = 1$ (i.e., the norms of the f and g vectors are unity). The output can be sold on the product market for price p. Output is noncontractible, but the performance measure is contractible. The Principal can write a contract $w \in \mathcal{W} \subset \{w : \mathcal{M} \to \mathbb{R}\}$ that determines a transfer $w(m)$ that she is compelled to pay the Agent if performance measure m is realized. Since the performance measure is binary, contracts take the form $w = s + bm$. The Agent has an outside option that provides utility $\bar{u}$ to the Agent and $\bar{\pi}$ to the Principal. If the outside option is not exercised, the Principal's and Agent's preferences are, respectively,

$$\begin{aligned} \Pi(w, e) &= f_1 e_1 + f_2 e_2 - s - b(g_1 e_1 + g_2 e_2) \\ U(w, e) &= s + b(g_1 e_1 + g_2 e_2) - \frac{c}{2}\left(e_1^2 + e_2^2\right). \end{aligned}$$

Timing The timing of the game is exactly the same as before.

1. P offers A a contract w, which is commonly observed.

2. A accepts the contract $(d = 1)$ or rejects it $(d = 0)$ and receives $\bar{u}$ and the game ends. This decision is commonly observed.

3. If A accepts the contract, A chooses effort vector e. e is only observed by A.

4. Performance measure m and output y are drawn from the distributions described above. m is commonly observed.

5. P pays A an amount $w(m)$. This payment is commonly observed.

Equilibrium The solution concept is the same as before. A **pure-strategy subgame-perfect equilibrium** is a contract $w^* \in \mathcal{W}$, an acceptance decision $d^* : \mathcal{W} \to \{0,1\}$, and an effort choice $e^* : \mathcal{W} \times \{0,1\} \to \mathbb{R}_+^2$ such that given the contract w^*, the Agent optimally chooses d^* and e^*, and given d^* and e^*, the Principal optimally offers contract w^*. We will say that the optimal contract induces effort e^*.

The Program The principal offers a contract w and proposes an effort level e to solve

$$\max_{s,b,e} p\left(f_1 e_1 + f_2 e_2\right) - \left(s + b\left(g_1 e_1 + g_2 e_2\right)\right)$$

subject to the incentive-compatibility constraint

$$e \in \underset{\hat{e} \in \mathbb{R}_+^2}{\operatorname{argmax}}\, s + b\left(g_1 \hat{e}_1 + g_2 \hat{e}_2\right) - \frac{c}{2}\left(\hat{e}_1^2 + \hat{e}_2^2\right)$$

and the individual-rationality constraint

$$s + b\left(g_1 e_1 + g_2 e_2\right) - \frac{c}{2}\left(e_1^2 + e_2^2\right) \geq \bar{u}.$$

Equilibrium Contracts and Effort Given a contract $s + bm$, the Agent will choose

$$e_1^*\left(b\right) = \frac{b}{c} g_1;\ e_2^*\left(b\right) = \frac{b}{c} g_2.$$

The Principal will choose s so that the individual-rationality constraint holds with equality

$$s + b\left(g_1 e_1^*\left(b\right) + g_2 e_2^*\left(b\right)\right) = \bar{u} + \frac{c}{2}\left(e_1^*\left(b\right)^2 + e_2^*\left(b\right)^2\right).$$

Since contracts send the Agent off in the "wrong direction" relative to what maximizes total surplus, providing the Agent with higher-powered incentives by increasing b sends the agent farther off in the wrong direction. This is costly for the Principal because in order to get the

Agent to accept the contract, she has to compensate him for his effort costs, even if they are in the wrong direction.

The Principal's unconstrained problem is therefore

$$\max_b p\left(f_1 e_1^*(b) + f_2 e_2^*(b)\right) - \frac{c}{2}\left(e_1^*(b)^2 + e_2^*(b)^2\right) - \bar{u}.$$

Taking first-order conditions,

$$pf_1 \underbrace{\frac{\partial e_1^*}{\partial b}}_{g_1/c} + pf_2 \underbrace{\frac{\partial e_2^*}{\partial b}}_{g_2/c} = c\underbrace{e_1^*(b^*)}_{b^* g_1/c}\underbrace{\frac{\partial e_1^*}{\partial b}}_{g_1/c} + c\underbrace{e_2^*(b^*)}_{b^* g_2/c}\underbrace{\frac{\partial e_2^*}{\partial b}}_{g_2/c},$$

or

$$b^* = p\frac{f_1 g_1 + f_2 g_2}{g_1^2 + g_2^2} = p\frac{f \cdot g}{g \cdot g} = p\frac{||f||}{||g||}\cos\theta = p\cos\theta,$$

where $\cos\theta$ is the angle between the vectors f and g. That is, the optimal incentive slope depends on the relative magnitudes of the f and g vectors (which in this model were assumed to be the same, but in a richer model this need not be the case) as well as how well-aligned they are. If m is a perfect measure of what the firm cares about, then g is a linear transformation of f and therefore the angle between f and g would be zero, so that $\cos\theta = 1$. If m is completely uninformative about what the firm cares about, then f and g are orthogonal, and therefore $\cos\theta = 0$. As a result, this model is often referred to as the **"cosine of theta model**." (Gibbons, 2010)

It can be useful to view this problem geometrically. Since formal contracts allow for unrestricted lump-sum transfers between the Principal and the Agent, the Principal would optimally like efforts to be chosen in such a way that they maximize total surplus:

$$\max_e p\left(f_1 e_1 + f_2 e_2\right) - \frac{c}{2}\left(e_1^2 + e_2^2\right),$$

which has the same solution as

$$\max_e -\left(e_1 - \frac{p}{c} f_1\right)^2 - \left(e_2 - \frac{p}{c} f_2\right)^2.$$

That is, the Principal would like to choose an effort vector that is collinear with the vector f:

$$\left(e_1^{FB}, e_2^{FB}\right) = \frac{p}{c} \cdot (f_1, f_2).$$

This effort vector would coincide with the first-best effort vector, since it maximizes total surplus, and the players have quasilinear preferences.

Since contracts can only depend on m and not directly on y, the Principal has only limited control over the actions that the Agent chooses. That is, given a contract specifying incentive slope b, the Agent chooses $e_1^*(b) = \frac{b}{c} g_1$ and $e_2^*(b) = \frac{b}{c} g_2$. Therefore, the Principal can only indirectly "choose" an effort vector that is collinear with the vector g:

$$\left(e_1^*(b), e_2^*(b)\right) = \frac{b}{c} \cdot (g_1, g_2).$$

The question is then: which such vector maximizes total surplus, which the Principal will extract with an ex ante lump-sum transfer? That is, which point along the $k \cdot (g_1, g_2)$ ray minimizes the mean-squared error distance to $\frac{p}{c} \cdot (f_1, f_2)$?

The following figure illustrates the first-best effort vector e^{FB} and the equilibrium effort vector e^*. The concentric rings around e^{FB} are the Principal's iso-profit curves. The rings that are closer to e^{FB} represent higher profit levels. The optimal contract induces effort vector e^*, which also coincides with the orthogonal projection of e^{FB} onto the ray $k \cdot (g_1, g_2)$.

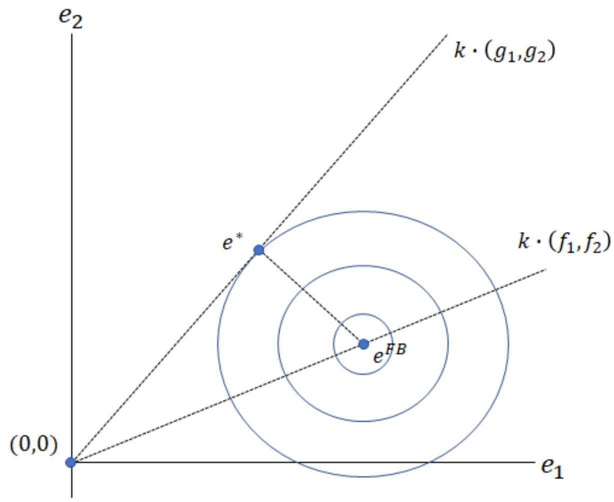

Figure 4

This is a more explicit "incomplete contracts" model of motivation. That is, we are explicitly restricting the set of contracts that the Principal can offer the Agent in a way that directly determines a subset of the effort space that the Principal can induce the Agent to choose among. And it is founded not on the idea that certain measures (in particular, y) are unobservable, but rather that they simply cannot be contracted upon.

One observation that is immediate is that it may sometimes be optimal to offer incentive contracts that provide no incentives for the Agent to choose positive effort levels (i.e., $b^* = 0$). This was essentially never the case in the model in which the Agent chose only a one-dimensional effort level, yet we often see that many employees are on contracts that look like they offer no performance-based payments. As this model highlights, this may be optimal precisely when the set of available performance measures are quite bad. As an example, suppose

$$\Pr\left[y = 1 | \, e\right] = \alpha + f_1 e_1 + f_2 e_2,$$

where $\alpha > 0$ and $f_2 < 0$, so that higher choices of e_2 reduce the probability of high output.

And suppose the performance measure is again

$$\Pr\left[m = 1 \middle| e\right] = g_1 e_1 + g_2 e_2,$$

with $g_1, g_2 > 0$.

We can think of $y = 1$ as representing whether a particular customer buys something that he does not later return, which depends on how well he was treated when he went to the store. We can think of $m = 1$ as representing whether the Agent made a sale but not whether the item was later returned. In order to increase the probability of making a sale, the Agent can exert "earnest" sales effort e_1 and "shady" sales effort e_2. Both are good for sales, but the latter increases the probability the item is returned. If the vectors f and g are sufficiently poorly aligned (i.e., if it is really easy to make sales by being shady), it may be better for the firm to offer a contract with $b^* = 0$, as the following figure illustrates.

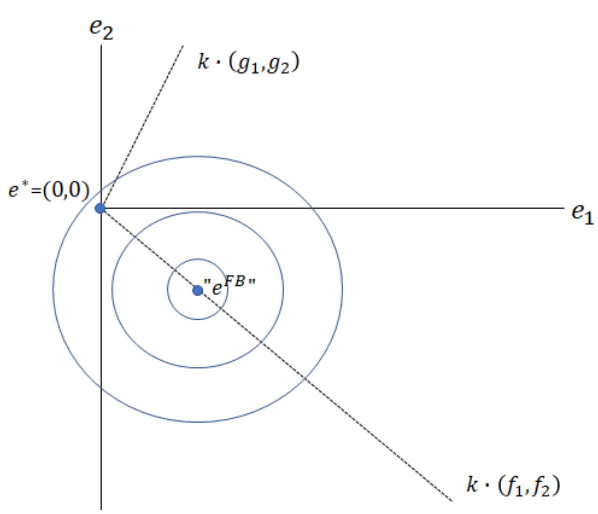

Figure 5

This example illustrates that paying the Agent for sales can be a bad idea when what the Principal wants is *sales that are not returned*. The Kerr (1975) article is filled with many colorful examples of this problem. One such example concerns the incentives offered to the managers of orphanages. Their budgets and prestige were determined largely by the number of children they enrolled and not by whether they managed to place their children with suitable families. The claim made in the article is that the managers often denied adoption applications for inappropriate reasons: they were being rewarded for large orphanages, while the state hoped for good placements.

Limits on Activities

Firms have many instruments to help address the problems that arise in multitasking situations. We will describe two of them here in a small extension to the model. Suppose now that the Principal can put some restrictions on the types of actions the Agent is able to undertake. In particular, in addition to writing a contract on the performance measure m, she can write a contract on the dummy variables $1_{e_1>0}$ and $1_{e_2>0}$. In other words, while she cannot directly contract upon, say, e_2, she can write a contract that heavily penalizes any positive level of it. The first question we will ask here is: when does the Principal want to exclude the Agent from engaging in task 2?

We can answer this question using the graphical intuition we just developed above. The following figure illustrates this intuition. If the Principal does not exclude task 2, then she can induce the Agent to choose any effort vector of the form $k \cdot (g_1, g_2)$. If she does exclude task 2, then she can induce the Agent to choose any effort vector of the form $k \cdot (g_1, 0)$. In the former case, the equilibrium effort vector will be e^*, which corresponds to the orthogonal projection of e^{FB} onto the ray $k \cdot (g_1, g_2)$. In the latter case, the equilibrium effort will be e^{**}, which corresponds to the orthogonal projection of e^{FB} onto the ray $k \cdot (g_1, 0)$.

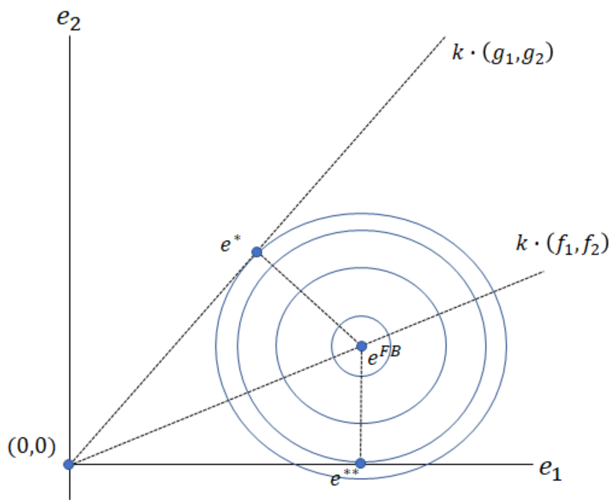

Figure 6

This figure shows that for the particular vectors f and g it illustrates, it will be optimal for the Principal to exclude e_2: e^{**} lies on a higher iso-profit curve than e^* does. This will in fact be the case whenever the angle between vector f and g is larger than the angle between f and $(g_1, 0)$—if by excluding task 2, the performance measure m acts as if it is more closely aligned with f, then task 2 should be excluded.

Job Design

Finally, we will briefly touch upon what is referred to as job design. Suppose f and g are such that it is not optimal to exclude either task on its own. The firm may nevertheless want to hire *two* Agents who each specialize in a single task. For the first Agent, the Principal could exclude task 2, and for the second Agent, the Principal could exclude task 1. The Principal could then offer a contract that gets the first Agent to choose $\left(e_1^{FB}, 0\right)$ and the second agent to choose $\left(0, e_2^{FB}\right)$. The following figure illustrates this possibility.

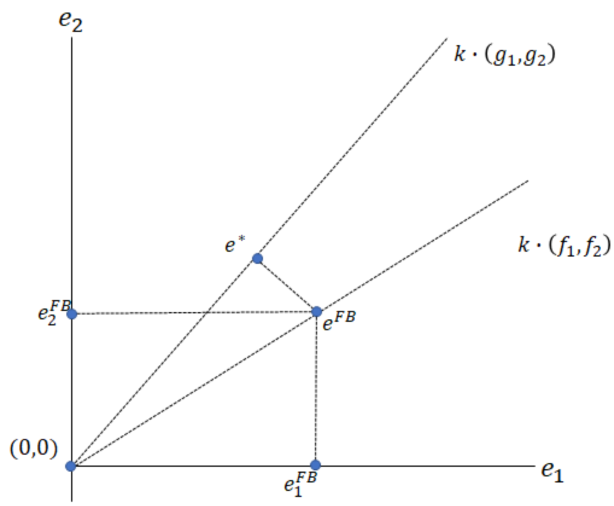

Figure 7

When is it optimal for the firm to hire two Agents who each specialize in a single task? It depends on the Agents' opportunity cost. Total surplus under a single Agent under the optimal contract will be

$$pf \cdot e^* - \frac{c}{2} e^* \cdot e^* - \bar{u},$$

and total surplus with two specialized agents under optimal contracts will be

$$pf \cdot e^{FB} - \frac{c}{2} e^{FB} \cdot e^{FB} - 2\bar{u}.$$

Adding an additional Agent in this case is tantamount to adding an additional performance measure, which allows the Principal to choose induce any $e \in \mathbb{R}_+^2$, including the first-best effort vector. She gains from being able to do this, but to do so, she has to cover the additional Agent's opportunity cost $\bar{u}$.

Further Reading Holmstrom and Milgrom (1991, 1994) explore many interesting organizational implications of misaligned performance measures in multi-task settings. In particular, they show that when performance measures are misaligned, it may be optimal to put in place rules that restrict the actions an agent is allowed to perform, it may be optimal to split up activities across agents (job design), and it may be optimal to adjust the boundaries of the firm. Job restrictions, job design, boundaries of the firm, and incentives should be designed to be an internally consistent system. The model described in this section is formally equivalent to Baker's (1992) model in which the agent receives noncontractible private information about the effectiveness of his (single) task before making his effort decision, since his contingent plan of effort choices can be viewed as a vector of effort choices that differentially affect his expected pay. This particular specification was spelled out in Baker's (2002) article, and it is related to Feltham and Xie's (1994) model.