Part II Asymmetric Information

7 Moral Hazard: Hidden Actions

7.1 Categories of Asymmetric Information Models

It used to be that the economist's first response to peculiar behavior which seemed to contradict basic price theory was "It must be some kind of price discrimination." Today, we have a new answer: "It must be some kind of asymmetric information." In a game of asymmetric information, player Smith knows something that player Jones does not. This covers a broad range of models (including price discrimination itself), so it is not surprising that so many situations come under its rubric. We will look at them in five chapters.

Moral hazard with hidden actions (Chapters 7 and 8)
Smith and Jones begin with symmetric information and agree to a contract, but then Smith takes an action unobserved by Jones. Information is complete.

Adverse selection (Chapter 9)
Nature begins the game by choosing Smith's type, unobserved by Jones. Smith and Jones then agree to a contract. Information is incomplete.

Mechanism design in adverse selection and **post- contractual hidden knowledge**) (Chapter 10)
Jones is designing a contract for Smith designed to elicit Smith's private information. This may happen under adverse selection— in which case Smith knows the information prior to contracting— or post-contractual hidden knowledge (also called moral hazard with hidden information)—in which case Smith will learn it after contracting.

Signalling and **Screening** (Chapter 11)
Nature begins the game by choosing Smith's type, unobserved by Jones. To demonstrate his type, Smith takes actions that Jones can observe. If Smith takes the action before they agree to a contract, he is signalling. If he takes it afterwards, he is being screened. Information is incomplete.

The important distinctions to keep in mind are whether or not the players agree to a contract before or after information becomes asymmetric, and whether their own actions are common knowledge. Not all the terms I used above are firmly established. In particular, some people would say that information *becomes* incomplete in a model of post-contractual hidden knowledge, even though it is complete at the start of the game. That statement runs contrary to the definition of complete information in Chapter 2, however.

We will make heavy use of the principal-agent model. Usually this term is applied to moral hazard models, since the problems studied in the law of agency usually involve an employee who disobeys orders by choosing the wrong actions, but the paradigm is useful in all four contexts listed above. The two players are the principal and the agent, who are usually representative individuals. The principal hires an agent to perform a task, and the agent acquires an informational advantage about his type, his actions, or the outside world at some point in the game. It is usually assumed that the players can make a binding **contract** at some point in the game, which is to say that the principal can commit to paying the agent an agreed sum if he observes a certain outcome. In the background of such models are courts, which will punish any player who breaks a contract in a way that can be proven with public information.

The **principal** *(or* **uninformed player***) is the player who has the coarser information partition.*

The **agent** *(or* **informed player***) is the player who has the finer information partition.*

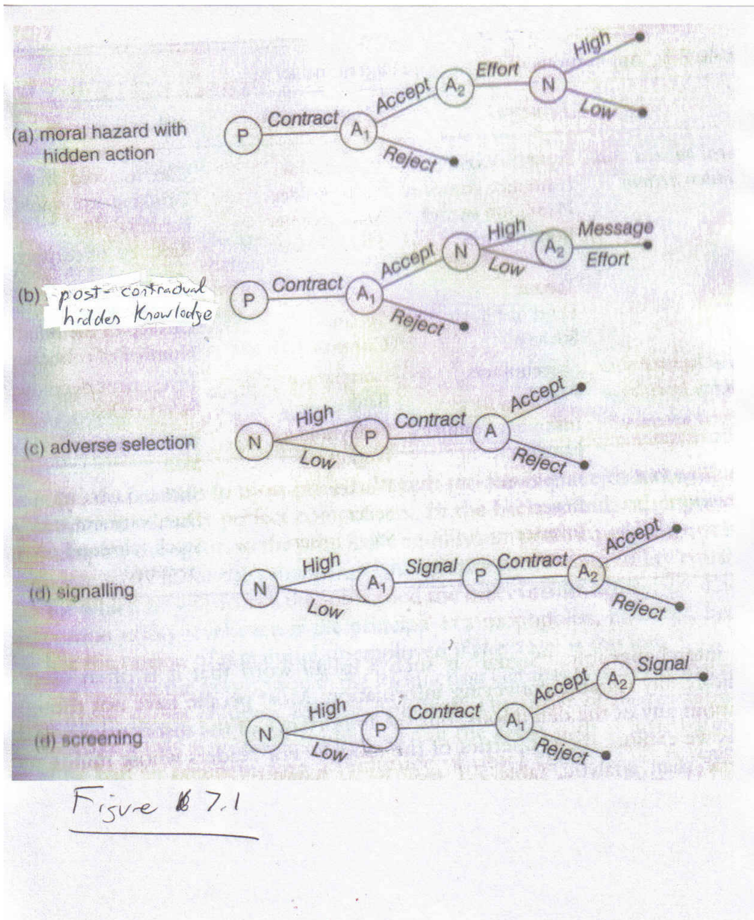

Figure 1: Categories of Asymmetric Information Models

Figure 1 shows the game trees for five principal-agent models. In each model, the principal (P) offers the agent (A) a contract, which he accepts or rejects. In some, Nature (N) makes a move, or the agent chooses an effort level, message, or signal. The moral hazard models are games of complete information with uncertainty. The principal offers a contract, and after the agent accepts, Nature adds noise to the task being performed. In moral hazard with hidden actions, Figure 1(a), the agent moves before Nature and in moral hazard with hidden knowledge, Figure 1(b), the agent moves after Nature and conveys a "message" to the principal about Nature's move.

Adverse selection models have incomplete information, so Nature moves first and picks the type of the agent, generally on the basis of his ability to perform the task. In the simplest model, Figure 1(c), the agent simply accepts or rejects the contract. If the agent can send a "signal" to the principal, as in Figures 1(d) and 1(e), the model is signalling if he sends the signal before the principal offers a contract, and screening otherwise. A "signal" is different from a "message" because it is not a costless statement, but a costly action. Some adverse selection models include uncertainty and some do not.

A problem we will consider in detail arises when an employer (the principal) hires a worker (the agent). If the employer knows the worker's ability but not his effort level, the problem is moral hazard with hidden actions. If neither player knows the worker's ability at first, but the worker discovers it once he starts working, the problem is moral hazard with hidden knowledge. If the worker knows his ability from the start, but the employer does not, the problem is adverse selection. If, in addition to the worker knowing his ability from the start he can acquire credentials before he makes a contract with the employer, the problem is signalling. If the worker acquires his credentials in response to a wage offer made by the employer, the problem is screening.

The five categories are not uniformly recognized, and in particular, some would argue that what I have called "moral hazard with hidden knowledge" and "screening" are essentially the same as adverse selection. Myerson (1991, p. 263), for example, suggests calling the problem of players taking the wrong action "moral hazard" and the problem of misreporting information "adverse selection." Or, the two problems can be called "hidden actions" versus "hidden knowledge". (The separation of asymmetric information into hidden actions and hidden knowledge is suggested in Arrow [1985] and commented on in Hart & Holmstrom [1987]). Many economists do not realize that screening and signalling are different and use the terms interchangeably. "Signal" is such a useful word that it is often used simply to indicate any variable conveying information. Most people have not thought very hard about any of the definitions, but the importance of the distinctions will become clear as we explore the properties of the models. For readers whose minds are more synthetic than analytic, Table 1 may be as helpful as anything in clarifying the categories.

Table 1: Applications of the Principal-Agent Model

	Principal	Agent	Effort or type and signal
Moral hazard with hidden actions	Insurance company	Policyholder	Care to avoid theft
	Insurance company	Policyholder	Drinking and smoking
	Plantation owner	Sharecropper	Farming effort
	Bondholders	Stockholders	Riskiness of corporate projects
	Tenant	Landlord	Upkeep of the building
	Landlord	Tenant	Upkeep of the building
	Society	Criminal	Number of robberies
Moral hazard with hidden knowledge	Shareholders	Company president	Investment decision
	FDIC	Bank	Safety of loans
Adverse selection	Insurance company	Policyholder	Infection with HIV virus
	Employer	Worker	Skill
Signalling and screening	Employer	Worker	Skill and education
	Buyer	Seller	Durability and warranty
	Investor	Stock issuer	Stock value and percentage retained

Section 7.2 discusses the roles of uncertainty and asymmetric information in a principal-agent model of moral hazard with hidden actions, called the Production Game, and Section 7.3 shows how various constraints are satisfied in equilibrium. Section 7.4 collects several unusual contracts produced under moral hazard and discusses the properties of optimal contracts using the example of the Broadway Game.

7.2 A Principal-Agent Model: The Production Game

In the archetypal principal-agent model, the principal is a manager and the agent a worker. In this section we will devise a series of these games, the last of which will be the standard principal-agent model.

Denote the monetary value of output by $q(e)$, which is increasing in effort, e. The agent's utility function $U(e, w)$ is decreasing in effort and increasing in the wage, w, while the principal's utility $V(q - w)$ is increasing in the difference between output and the wage.

The Production Game

Players

The principal and the agent.

The order of play

1 The principal offers the agent a wage w.
2 The agent decides whether to accept or reject the contract.
3 If the agent accepts, he exerts effort e.
4 Output equals $q(e)$, where $q' > 0$.

Payoffs

If the agent rejects the contract, then $\pi_{agent} = \bar{U}$ and $\pi_{principal} = 0$.
If the agent accepts the contract, then $\pi_{agent} = U(e, w)$ and $\pi_{principal} = V(q - w)$.

An assumption common to most principal-agent models is that either the principal or the agent is one of many perfect competitors. In the background, either (a) other principals compete to employ the agent, so the principal's equilibrium profit equals zero; or (b) many agents compete to work for the principal, so the agent's equilibrium utility equals the minimum for which he will accept the job, called the **reservation utility**, $\bar{U}$. There is some reservation utility level even if the principal is a monopolist, however, because the agent has the option of remaining unemployed if the wage is too low.

One way of viewing the assumption in the Production Game that the principal moves first is that many agents compete for one principal. The order of moves allows the principal to make a take-it-or-leave-it offer, leaving the agent with as little bargaining room as if he had to compete with a multitude of other agents. This is really just a modelling convenience, however, since the agent's reservation utility, $\bar{U}$, can be set at the level a principal would have to pay the agent in competition with other principals. This level of $\bar{U}$ can even be calculated, since it is the level at which the principal's payoff from profit maximization using the optimal contract is driven down to the principal's reservation utility by competition with other principals. Here the principal's reservation utility is zero, but that too can be chosen to fit the situation being modelled. As in the game of Nuisance Suits in Section 4.3, the main concern in choosing who makes the offer is to avoid the distraction of more complicated modelling of the bargaining subgame.

Refinements of the equilibrium concept will not be important in this chapter. Information is complete, and the concerns of perfect bayesian equilibrium will not arise. Subgame perfectness will be required, since otherwise the agent might commit to reject any contract that does not give him all of the gains from trade, but it will not drive the important results.

We will go through a series of eight versions of the Production Game in various chapters.

Production Game I: Full Information

In the first version of the game, every move is common knowledge and the contract is a function $w(e)$.

Finding the equilibrium involves finding the best possible contract from the point of view of the principal, given that he must make the contract acceptable to the agent and that he foresees how the agent will react to the contract's incentives. The principal must decide what he wants the agent to do and what incentive to give him to do it.

The agent must be paid some amount $\tilde{w}(e)$ to exert effort e, where $\tilde{w}(e)$ is the function that makes him just willing to accept the contract, so

$$U(e, w(e)) = \overline{U}. \tag{1}$$

Thus, the principal's problem is

$$\underset{e}{Maximize} \quad V(q(e) - \tilde{w}(e)) \tag{2}$$

The first-order condition for this problem is

$$V'(q(e) - \tilde{w}(e)) \left(\frac{\partial q}{\partial e} - \frac{\partial \tilde{w}}{\partial e} \right) = 0, \tag{3}$$

which implies that

$$\frac{\partial q}{\partial e} = \frac{\partial \tilde{w}}{\partial e}. \tag{4}$$

From condition (1), using the implicit function theorem (see section 13.4), we get

$$\frac{\partial \tilde{w}}{\partial e} = - \left(\frac{\frac{\partial U}{\partial e}}{\frac{\partial U}{\partial \tilde{w}}} \right). \tag{5}$$

Combining equations (4) and (5) yields

$$\left(\frac{\partial U}{\partial \tilde{w}} \right) \left(\frac{\partial q}{\partial e} \right) = - \left(\frac{\partial U}{\partial e} \right). \tag{6}$$

Equation (6) says that at the optimal effort level, e^*, the marginal utility to the agent which would result if he kept all the marginal output from extra effort equals the marginal disutility to him of that effort.

Figure 2 shows this graphically. The agent has indifference curves in effort- wage space that slope upwards, since if his effort rises his wage must increase also to keep his utility the same. The principal's indifference curves also slope upwards, because although he does not care about effort directly, he does care about output, which rises with effort. The principal might be either risk averse or risk neutral; his indifference curve is concave rather than linear in either case because Figure 2 shows a technology with diminishing returns to effort (that is, concave, with $q''(e) < 0$). If effort starts out being higher, extra effort yields less additional output so the wage cannot rise as much without reducing profits.

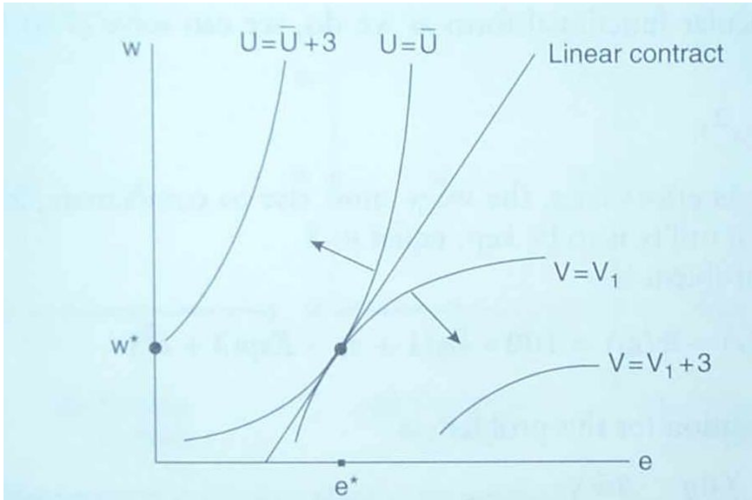

Figure 2: The Efficient Effort Level in Production Game I

Under perfect competition among the principals the profits are zero, so the reservation utility, $\overline{U}$, will be at the level such that at the profit-maximizing effort e^*, $\tilde{w}(e^*) = q(e^*)$, or

$$U(e^*, q(e^*)) = \overline{U}. \tag{7}$$

The principal selects the point on the $U = \overline{U}$ indifference curve that maximizes his profits, at effort e^* and wage w^*. He must then design a contract that will induce the agent to choose this effort level. The following three contracts, shown in Figure 3, are equally effective under full information.

1 The **forcing contract** sets $w(e^*) = w^*$ and $w(e \neq e^*) = 0$. This is certainly a strong incentive for the agent to choose exactly $e = e^*$.

2 The **threshold contract** sets $w(e \geq e^*) = w^*$ and $w(e < e^*) = 0$. This can be viewed as a flat wage for low effort levels, equal to 0 in this contract, plus a bonus if effort reaches e^*. Since the agent dislikes effort, the agent will choose exactly $e = e^*$.

3 The **linear contract**, shown in both Figure 2 and Figure 3(c), sets $w(e) = \alpha + \beta e$, where α and β are chosen so that $w^* = \alpha + \beta e^*$ and the contract line is tangent to the indifference curve $U = \bar{U}$ at e^*. In Figure 3(c), the most northwesterly of the agent's indifference curves that touch this contract line touches it at e^*.

Let's now fit out Production Game I with specific functional forms. Suppose the agent exerts effort $e \in [0, \infty]$, and output equals

$$q(e) = 100 * log(1 + e), \tag{8}$$

so $q' = \frac{100}{1+e} > 0$ and $q'' = \frac{-100}{(1+e)^2} < 0$. If the agent rejects the contract, let $\pi_{agent} = \bar{U} = 3$ and $\pi_{principal} = 0$, whereas if the agent accepts the contract, let $\pi_{agent} = U(e, w) = log(w) - e^2$ and $\pi_{principal} = q(e) - w(e)$.

The agent must be paid some amount $\tilde{w}(e)$ to exert effort e, where $\tilde{w}(e)$ is defined to be the wage that makes the agent willing to participate, i.e., as in equation (1),

$$U(e, w(e)) = \overline{U}, \quad so \;\; log(\tilde{w}(e)) - e^2 = 3. \tag{9}$$

Knowing the particular functional form as we do, we can solve (9) for the wage function:

$$\tilde{w}(e) = Exp(3 + e^2), \tag{10}$$

where we use $Exp(x)$ to mean Euler's constant (about 2.718) to the power x, since the conventional notation of e^x would be confused with e as effort.

Equation (10) makes sense. As effort rises, the wage must rise to compensate, and rise more than exponentially if utility is to be kept equal to 3.

Now that we have a necessary-wage function $\tilde{w}(e)$, we can attack the principal's problem, which is

$$\underset{e}{Maximize} \quad V(q(e) - \tilde{w}(e)) = 100 * log(1 + e) - Exp(3 + e^2) \tag{11}$$

The first-order condition for this problem is

$$V'(q(e) - \tilde{w}(e)) \left(\frac{\partial q}{\partial e} - \frac{\partial \tilde{w}}{\partial e} \right) = 0, \tag{12}$$

so for our problem,

$$\left(\frac{100}{1 + e} \right) - 2e(Exp(3 + e^2)) = 0, \tag{13}$$

which cannot be solved analytically.[1] Using the computer program Mathematica, I found that $e^* \approx 0.77$, from which, using the formulas above, we get $q^* \approx 57$ and $w^* \approx 37$. The payoffs are $\pi_{agent} = 3$ and $\pi_{principal} \approx 20$.

If $\overline{U}$ were high enough, the principal's payoff would be zero. If the market for agents were competitive, this is what would happen, since the agent's reservation payoff would be the utility of working for another principal instead of $\overline{U} = 3$.

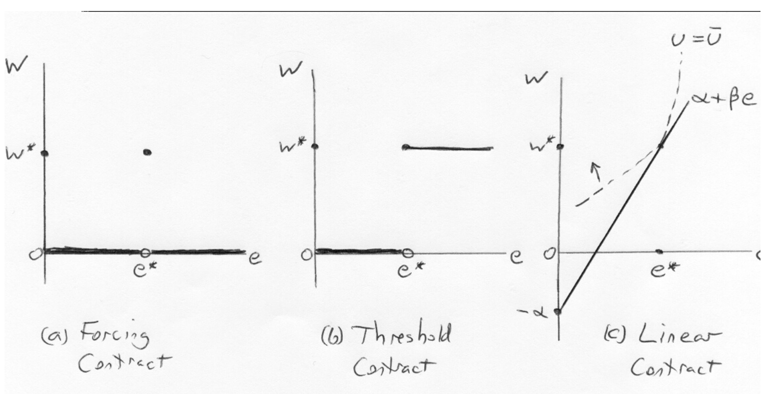

Figure 3: Three Contracts that Induce Effort e^* for Wage w^*

To obtain $e^* = 0.77$, a number of styles of contract could be used, as shown in Figure 3.

[1]Note that we did not need to use the principal's risk-neutrality— that $V' = 1$— to get to equation (13). The optimal effort does not depend on the principal's degree of risk aversion in this certainty model.

1 The **forcing contract** sets $w(e^*) = w^*$ and $w(e \neq 0.77) = 0$. Here, $w(0.77) = 37$ (rounding up) and $w(e \neq e^*) = 0$.

2 The **threshold contract** sets $w(e \geq e^*) = w^*$ and $w(e < e^*) = 0$. Here, $w(e \geq 0.77) = 37$ and $w(e < 0.77) = 0$.

3 The **linear contract** sets $w(e) = \alpha + \beta e$, where α and β are chosen so that $w^* = \alpha + \beta e^*$ and the contract line is tangent to the indifference curve $U = \bar{U}$ at e^*. The slope of that indifference curve is the derivative of the $\tilde{w}(e)$ function, which is

$$\frac{\partial \tilde{w}(e)}{\partial e} = 2e * Exp(3 + e^2). \tag{14}$$

At $e^* = 0.77$, this takes the value 56 (which only coincidentally is near the value of $q^* = 57$). That is the β for the linear contract. The α must solve $w(e^*) = 37 = \alpha + 56(0.77)$, so $\alpha = -7$.

We ought to be a little concerned as to whether the agent will choose the effort we hope for if he is given the linear contract. We constructed it so that he would be willing to accept the contract, because if he chooses $e = 0.77$, his utility will be 3. But might he prefer to choose some larger or smaller e and get even more utility? No, because his utility is concave. That makes the indifference curve convex, so its slope is always increasing and no preferable indifference curve touches the equilibrium contract line.

Quasilinearity and Alternative Functional Forms for the Production Game

Consider the following three functional forms for utility:

$$\begin{aligned} U(e,w) &= log(w) - e^2 && (a) \\ U(e,w) &= w - e^2 && (b) \\ U(e,w) &= log(w - e^2) && (c) \end{aligned} \tag{15}$$

Utility function (a) is what we just used in Production Game I. Utility function (b) is an example of **quasilinear preferences**, because utility is separable in one good— money, here— and linear in that good. This kind of utility function is commonly used to avoid wealth effects that would otherwise occur in the interactions among the various goods in the utility function. Separability means that giving an agent a higher wage does not, for example, increase his marginal disutility of effort. Linearity means furthermore that giving an agent a higher wage does not change his tradeoff between money and effort, his marginal rate of substitution, as it would in function (a), where a richer agent is less willing to accept money for higher effort. In effort-wage diagrams, quasilinearity implies that the indifference curves are parallel along the effort axis (which they are *not* in Figure 2).

Quasilinear utility functions most often are chosen to look like (b), but my colleague Michael Rauh points out that what quasilinearity really requires is just linearity in the special good (w here) for some monotonic transformation of the utility function. Utility function (c) is a logarithmic transformation of (b), which is a monotonic transformation, so it too is quasilinear. That is because marginal rates of substitution, which is what matter here, are a feature of general utility functions, not the Von Neumann-Morgenstern functions we typically use. Thus, utility function (c) is also a quasi-linear function, because it is just a monotonic function of (b). This is worth keeping in mind because utility function (c) is concave in w, so it represents a risk-averse agent.

Returning to the solution of Production Game I, let us now use a different approach to get to the same answer as we did using the principal's maximization problem (11). Instead, we will return to the general optimality condition(6), here repeated.

$$\left(\frac{\partial U}{\partial \tilde{w}}\right)\left(\frac{\partial q}{\partial e}\right) = -\frac{\partial U}{\partial e} \qquad (6)$$

For any of our three utility functions we will continue using the same output function $q(e) = 100 * log(1+e)$ from (8), which has the first derivative $q' = \frac{100}{1+e}$.

Using utility function (a), $\frac{\partial U}{\partial \tilde{w}} = 1/w$. and $\frac{\partial U}{\partial e} = -2e$, so equation (6) becomes

$$\left(\frac{1}{w}\right)\left(\frac{100}{1+e}\right) = -(-2e). \qquad (16)$$

If we substitute for w using the function $\tilde{w}(e) = Exp(3+e^2)$ that we found in equation (10), we get essentially the same equation as (13), and so outcomes are the same— $e^* \approx 0.77$, $q^* \approx 57$, and $w^* \approx 37$, $\pi_{agent} = 3$, and $\pi_{principal} \approx 20$.

Using utility function (b), $\frac{\partial U}{\partial \tilde{w}} = 1$ and $\frac{\partial U}{\partial e} = -2e$, so equation (6) becomes

$$(1)\left(\frac{100}{1+e}\right) = -(-2e) \qquad (17)$$

Notice that w has disappeared. The optimal effort no longer depends on the agent's wealth. Thus, we don't need to use the wage function to solve for the optimal effort. Solving directly, we get $e^* \approx 6.59$ and $q^* \approx 203$. The wage function will be different now, solving $w - e^2 = 3$, so $w^* \approx 43$, $\pi_{agent} = 3$, and $\pi_{principal} \approx 160$. (These numbers are not really comparable to when we used utility function (a), but they will be useful in Production Game II.)

Using utility function (c), $\frac{\partial U}{\partial \tilde{w}} = 1/(w - e^2)$ and $\frac{\partial U}{\partial e} = -2e/(w - e^2)$, so equation (6) becomes

$$\left(\frac{1}{w - e^2}\right)\left(\frac{100}{1+e}\right) = -\left(\frac{-2e}{w - e^2}\right) \qquad (18)$$

and with a little simplification,

$$\frac{100}{1+e} = 2e. \qquad (19)$$

The variable w has again disappeared, so as with utility function (b) the optimal effort does not depend on the agent's wealth. Solving for the optimal effort yields $e^* \approx 6.59$ and $q^* \approx 203$, the same as with utility function (b). The wage function is different, however. Now it solves $log(w - e^2) = 3$, so $w = e^2 + exp(3)$ and $w^* \approx 63$, $\pi_{agent} = 3$, and $\pi_{principal} \approx 140$.

Before going on to versions of the game with asymmetric information, it will be useful to look at another version of the game with full information, Production Game II, in which the agent, not the principal, proposes the contract.

Production Game II: Full Information. Agent Moves First.

In this version, every move is common knowledge and the contract is a function $w(e)$. The order of play, however, is now as follows

The Order of Play

1 The agent offers the principal a contract $w(e)$.
2 The principal decides whether to accept or reject the contract.
3 If the principal accepts, the agent exerts effort e.
4 Output equals $q(e)$, where $q' > 0$.

Now the agent has all the bargaining power, not the principal. Thus, instead of requiring that the contract be at least barely acceptable to the agent, our concern is that the contract be at least barely acceptable to the principal, who must earn zero profits so $q(e) - w(e) \geq 0$. The agent will maximize his own payoff by driving the principal to exactly zero profits, so $w(e) = q(e)$. Substituting $q(e)$ for $w(e)$ to account for this constraint, the maximization problem for the agent in proposing an effort level e at a wage $w(e)$ can therefore be written as

$$\underset{e}{Maximize} \quad U(e, q(e)) \tag{20}$$

The first-order condition is

$$\frac{\partial U}{\partial e} + \left(\frac{\partial U}{\partial q}\right)\left(\frac{\partial q}{\partial e}\right) = 0. \tag{21}$$

Since $\frac{\partial U}{\partial q} = \frac{\partial U}{\partial w}$ when the wages equals output, equation (21) implies that

$$\left(\frac{\partial U}{\partial w}\right)\left(\frac{\partial q}{\partial e}\right) = -\left(\frac{\partial U}{\partial e}\right). \tag{22}$$

Compare this with equation (6),the optimization condition in Production Game I, when the principal had the bargaining power, The optimality equation is identical in Production Games I and II. The intuition is the same in both too: since the player who proposes the contract captures all the gains from trade (for a given reservation payoff of the other player), he will choose an efficient effort level. This requires that the marginal utility of the money derived from marginal effort equal the marginal disutility of effort.

Although the form of the optimality equation is the same, however, the optimal effort might not be, because except in the special case in which the agent's reservation payoff in Production Game I equals his equilibrium payoff in Production Game II, the agent ends up with higher wealth if he has all the bargaining power. If the utility function is not quasi-linear, then the wealth effect will change the optimal effort.

We can see the wealth effect by solving out optimality equation (22) for the specific functional forms of Production Game I from expression (15).

Using utility function (a) from expression (15)

$$\left(\frac{1}{w}\right)\left(\frac{100}{1+e}\right) = -(-2e). \tag{23}$$

That is the same as in Production Game I, equation (16), but now w is different. It is not found by driving the agent to his reservation payoff, but by driving the principal to zero profits: $w = q$. Since $q = 100 * log(1+e)$, we can substitute that in for w to get

$$\left(\frac{1}{100 * log(1+e)}\right)\left(\frac{100}{1+e}\right) = 2e. \tag{24}$$

When solved numerically, this yields $e^* \approx 0.63$, and thus $q = w \approx 49$, and $\pi_{principal} = 0$ and $\pi_{agent} \approx 3.49$. In Production Game I, the optimal effort using this utility function was 0.77 and the agent's payoff was 3. The difference arises because there the agent's wealth was lower because the principal had the bargaining power. In Production Game II the agent is, in effect, wealthier, and since his marginal utility of money is lower, he chooses to convert some (but not all) of that extra wealth into what we might call leisure— working less hard.

Using the quasilinear utility functions (b) and (c) from expression (15), recall that both have the same optimality condition, the one we found in equations (17) and (19):

$$\frac{100}{1+e} = 2e \qquad (19)$$

As we observed before, w does not appear in equation (19), so the wage equation does not matter to e^*. But that means that in Production Game II, $e^* \approx 6.59$ and $q^* \approx 203$, just as in Production Game I. With quasilinear utility, the efficient action does not depend on bargaining power. Of course, the wage and payoffs do depend on who has the bargaining power. In Production Game II, $w^* = q^* \approx 203$, and $\pi_{principal} = 0$. The agent's payoff is higher than in Production Game I, but it differs, of course, depending on the payoff function. For utility function (b) it is $\pi_{agent} \approx 160$ and for utility function (c) it is $\pi_{agent} \approx 5.08$.

If utility is quasilinear, the efficient effort level is independent of which side has the bargaining power because the gains from efficient production are independent of how those gains are distributed so long as each party has no incentive to abandon the relationship. This as the same lesson as the Coase Theorem's:, under general conditions the activities undertaken will be efficient and independent of the distribution of property rights (Coase [1960]). This property of the efficient-effort level means that the modeller is free to make the assumptions on bargaining power that help to focus attention on the information problems he is studying.

There are thus three reasons why modellers so often use take-it-or-leave-it offers. The first two reasons were discussed earlier in the context of Production Game I: (1) such offers are a good way to model competitive markets, and (2) if the reservation payoff of the player without the bargaining power is set high enough, such offers lead to the same outcome as would be reached if that player had more bargaining power. Quasi-linear utility provides a third reason: (3) if utility is quasi-linear, the optimal effort level does not depend on who has the bargaining power, so the modeller is justified in choosing the simplest model of bargaining.

Production Game III: A Flat Wage Under Certainty

In this version of the game, the principal can condition the wage neither on effort nor on output. This is modelled as a principal who observes neither effort nor output, so information is asymmetric.

That a principal cannot observe effort is often realistic, but it seems less usual that he cannot observe output, since it directly affects the value of his payoff. It is not ridiculous that he cannot base wages on output, however, because a contract must be enforceable by some third party such as a court. Law professors complain about economists who speak of "unenforceable contracts." In law school, a contract is defined as an enforceable agreement, and most of a contracts class is devoted to discovering which agreements are contracts. A court cannot in practice enforce a contract in which a client agrees to pay a barber \$50 "if the haircut is especially good," but just \$10 otherwise. Similarly, an employer may be able to tell that a worker's slacking is hurting output, but that does not mean he can prove it in court. A court can only enforce contingencies it can observe. In the extreme, Production Game III is appropriate. Either output is not **contractible** (the court will not enforce a contract) or it is not **verifiable** (the court cannot observe output), which usually leads to the same outcome as when output is unobservable to the principal.

The outcome of Production Game III is simple and inefficient. If the wage is nonnegative, the agent accepts the job and exerts zero effort, so the principal offers a wage of zero.

In Production Game III, we have finally reached "moral hazard", the problem of the agent choosing the wrong action because the principal cannot use the contract to punish him. The term "moral hazard" is an old insurance term, as we will see later. A good way to think of it is that it is the danger to the principal that the agent, constrained only by his morality, not punishments, cannot be trusted to behave as he ought. Or, you might think of the situation as a temptation for the agent, a hazard to his morals.

Sometimes, as we will soon see, a clever contract can overcome moral hazard by conditioning the wage on something that is observable and correlated with effort, such as output. If there is nothing on which to condition the wage, however, the agency problem cannot be solved by designing the contract carefully. If it is to be solved at all, it will be by some other means such as reputation or repetition of the game, the solutions of Chapter 5, or by morality– which might be modelled as a part of the agent's utility function which causes him disutility if he secretly breaks an agreement. Typically, however, there is some contractible variable such as output upon which the principal can condition the wage. Such is the case in Production Game IV.

Production Game IV: An Output-Based Wage under Certainty

In this version, the principal cannot observe effort but he can observe output and specify the contract to be $w(q)$.

Unlike in Production Game III, the principal now picks not a number w but a function $w(q)$. His problem is not quite so straightforward as in Production Game I, where he picked the function $w(e)$, but here, too, it is possible to achieve the efficient effort level e^* despite the unobservability of effort. The principal starts by finding the optimal effort level e^*, as in Production Game I. That effort yields the efficient output level $q^* = q(e^*)$. To give the agent the proper incentives, the contract must reward him when output is q^*. Again, a variety of contracts could be used. The forcing contract, for example, would be any wage function such that $U(e^*, w(q^*)) = \bar{U}$ and $U(e, w(q)) < \bar{U}$ for $e \neq e^*$.

Production Game IV shows that the unobservability of effort is not a problem in itself, if the contract can be conditioned on something which is observable and perfectly correlated with effort. The true agency problem occurs when that perfect correlation breaks down, as in Production Game V.

Production Game V: An Output-Based Wage under Uncertainty.

In this version, the principal cannot observe effort but can observe output and specify the contract to be $w(q)$. Output, however, is a function $q(e, \theta)$ both of effort and the state of the world $\theta \in \mathbf{R}$, which is chosen by Nature according to the probability density $f(\theta)$ as a new move (5) of the game. Move (5) comes just after the agent chooses effort, so the agent cannot choose a low effort knowing that Nature will take up the slack. (If the agent can observe Nature's move before his own, the game becomes "moral hazard with hidden knowledge and hidden actions").

Because of the uncertainty about the state of the world, effort does not map cleanly onto observed output in Production Game V. A given output might have been produced by any of several different effort levels, so a forcing contract based on output will not necessarily achieve the desired effort. Unlike in Production Game IV, here the principal cannot deduce $e \neq e^*$ from $q \neq q^*$. Moreover, even if the contract does induce the agent to choose e^*, if it does so by penalizing him heavily when $q \neq q^*$ it will be expensive for the principal. The agent's expected utility must be kept equal to $\bar{U}$ so he will accept the contract, and if he is sometimes paid a low wage because output happens not to equal q^* despite his correct effort, he must be paid more when output does equal q^* to make up for it. If the agent is risk averse, this variability in his wage requires that his expected wage be higher than the w^* found earlier, because he must be compensated for the extra risk. There is a tradeoff between incentives and insurance against risk.

Put more technically, moral hazard is a problem when $q(e)$ is not a one-to- one function and a single value of e might result in any of a number of values of q, depending on the value of θ. In this case the output function is not invertible; knowing q, the principal cannot deduce the value of e perfectly without assuming equilibrium behavior on the part of the agent.

The combination of unobservable effort and lack of invertibility in Production Game V means that no contract can induce the agent to put forth the efficient effort level without incurring extra costs, which usually take the form of extra risk imposed on the agent. In some situations this is not actually a cost, because the agent is risk-neutral, but more often the best the principal can do is balance the benefit of extra incentive for effort against the cost of extra risk for a risk-average agent. We will still try to find a contract that is efficient in the sense of maximizing welfare given the informational constraints. The terms "first-best" and "second-best" are used to distinguish these two kinds of optimality.

A **first-best contract** *achieves the same allocation as the contract that is optimal when the principal and the agent have the same information set and all variables are contractible.*

A **second-best contract** *is Pareto optimal given information asymmetry and constraints on writing contracts.*

The difference in welfare between the first best and the second best is the cost of the agency problem.

So how do we find a second-best contract? Even to define the strategy space in a game like Production Game V is tricky, because the principal may wish to choose a very complicated function for $w(q)$. It is not very useful, for example, simply to maximize profit over all possible linear contracts, because the best contract may well not be linear. Because of the tremendous variety of possible contracts, finding the optimal contract when a forcing contract cannot be used is a hard problem without general answers. The rest of the chapter will show how the problem may be approached, if not actually solved.

7.3 The Incentive Compatibility and Participation Constraints

The principal's objective in Production Game V is to maximize his utility knowing that the agent is free to reject the contract entirely and that the contract must give the agent an incentive to choose the desired effort. These two constraints arise in every moral hazard problem, and they are named the **participation constraint** and the **incentive compatibility contraint**. Mathematically, the principal's problem is

$$\underset{w(\cdot)}{Maximize} \quad EV(q(\tilde{e},\theta) - w(q(\tilde{e},\theta))) \tag{25}$$

subject to

$$\tilde{e} = \overset{argmax}{e} \quad EU(e, w(q(e,\theta))) \quad \text{(incentive compatibility constraint) (25a)}$$

$$EU(\tilde{e}, w(q(\tilde{e},\theta))) \geq \bar{U} \qquad \text{(participation constraint) (25b)}$$

The incentive-compatibility constraint takes account of the fact that the agent moves second, so the contract must induce him to voluntarily pick the desired effort. The participation constraint, also called the **reservation utility** or **individual rationality** constraint, requires that the worker prefer the contract to leisure, home production, or alternative jobs.

Expression (25) is the way an economist instinctively sets up the problem, but setting it up is often as far as he can get with the **first- order condition approach**. The difficulty is not just that the maximizer is choosing a wage function instead of a number, because control theory or the calculus of variations can solve such problems. Rather, it is that the constraints are nonconvex— they do not rule out a nice convex set of points in the space of wage functions such as the constraint "$w \geq 4$" would, but rather rule out a very complicated set of possible wage functions.

A different approach, developed by Grossman & Hart (1983) and called the **three-step procedure** by Fudenberg & Tirole (1991a), is to focus on contracts that induce the agent to pick a particular action rather than to directly attack the problem of maximizing profits. The first step is to find for each possible effort level the set of wage contracts that induce the agent to choose that effort level. The second step is to find the contract which supports that effort level at the lowest cost to the principal. The third step is to choose the effort level that maximizes profits, given the necessity to support that effort with the costly wage contract from the second step.

To support the effort level e, the wage contract $w(q)$ must satisfy the incentive compatibility and participation constraints. Mathematically, the problem of finding the least cost $C(\tilde{e})$ of supporting the effort level $\tilde{e}$ combines steps one and two.

$$C(\tilde{e}) = \underset{w(\cdot)}{Minimum} \quad Ew(q(\tilde{e},\theta)) \tag{26}$$

subject to constraints (25a) and (25b).

Step three takes the principal's problem of maximizing his payoff, expression (25), and restates it as

$$\underset{\tilde{e}}{Maximize} \quad EV(q(\tilde{e},\theta) - C(\tilde{e})). \tag{27}$$

After finding which contract most cheaply induces each effort, the principal discovers the optimal effort by solving problem (27).

Breaking the problem into parts makes it easier to solve. Perhaps the most important lesson of the three-step procedure, however, is to reinforce the points that the goal of the contract is to induce the agent to choose a particular effort level and that asymmetric information increases the cost of the inducements.

7.4 Optimal Contracts: The Broadway Game

The next game, inspired by Mel Brooks's offbeat film *The Producers*, illustrates a peculiarity of optimal contracts: sometimes the agent's reward should not increase with his output. Investors advance funds to the producer of a Broadway show that might succeed or might fail. The producer has the choice of embezzling or not embezzling the funds advanced to him, with a direct gain to himself of 50 if he embezzles. If the show is a success, the revenue is 500 if he did not embezzle and 100 if he did. If the show is a failure, revenue is -100 in either case, because extra expenditure on a fundamentally flawed show is useless.

Broadway Game I

Players

Producer and investors.

The order of play

1 The investors offer a wage contract $w(q)$ as a function of revenue q.

2 The producer accepts or rejects the contract.

3 The producer chooses to *Embezzle* or *Do not embezzle*.

4 Nature picks the state of the world to be *Success* or *Failure* with equal probability. Table 2 shows the resulting revenue q.

Payoffs

The producer is risk averse and the investors are risk neutral. The producer's payoff is $U(100)$ if he rejects the contract, where $U' > 0$ and $U'' < 0$, and the investors' payoff is 0. Otherwise,

$$\pi_{producer} = \begin{cases} U(w(q) + 50) & if\ he\ embezzles \\ U(w(q)) & if\ he\ is\ honest \end{cases}$$

$$\pi_{investors} = q - w(q)$$

Table 2: Profits in Broadway Game I

		State of the World	
		Failure (0.5)	*Success* (0.5)
Effort	*Embezzle*	-100	$+100$
	Do not embezzle	-100	$+500$

Another way to tabulate outputs, shown in Table 3, is to put the probabilities of outcomes in the boxes, with effort in the rows and output in the columns.

Table 3: Probabilities of Profits in Broadway Game I

		Profit			
		-100	**+100**	**+500**	**Total**
	Embezzle	0.5	0.5	0	1
Effort					
	Do not embezzle	0.5	0	0.5	1

The investors will observe q to equal either $-100, +100$, or $+500$, so the producer's contract will specify at most three different wages: $w(-100), w(+100)$, and $w(+500)$. The producer's expected payoffs from his two possible actions are

$$\pi(\textit{Do not embezzle}) = 0.5U(w(-100)) + 0.5U(w(+500)) \tag{28}$$

and

$$\pi(\textit{Embezzle}) = 0.5U(w(-100)+50) + 0.5U(w(+100)+50). \tag{29}$$

The incentive compatibility constraint is $\pi(\textit{Do not embezzle}) \geq \pi(\textit{Embezzle})$, so

$$0.5U(w(-100)) + 0.5U(w(+500)) \geq 0.5U(w(-100)+50) + 0.5U(w(+100)+50), \tag{30}$$

and the participation constraint is

$$\pi(\textit{Do not embezzle}) = 0.5U(w(-100)) + 0.5U(w(+500)) \geq U(100). \tag{31}$$

The investors want the participation constraint (31) to be satisfied at as low a dollar cost as possible. This means they want to impose as little risk on the producer as possible, since he requires a higher expected wage for higher risk. Ideally, $w(-100) = w(+500)$, which provides full insurance. The usual agency tradeoff is between smoothing out the agent's wage and providing him with incentives. Here, no tradeoff is required, because of a special feature of the problem: there exists an outcome that could not occur unless the producer chooses the undesirable action. That outcome is $q = +100$, and it means that the following **boiling-in-oil contract** provides both riskless wages and effective incentives.

$w(+500) = 100$
$w(-100) = 100$
$w(+100) = -\infty$

Under this contract, the producer's wage is a flat 100 when he does not embezzle. Thus, the participation constraint is satisfied. It is also binding, because it is satisfied as an equality, and the investors would have a higher payoff if the constraint were relaxed. If the producer does embezzle, he faces a payoff of $-\infty$ with probability 0.5, so the incentive compatibility constraint is satisfied, but it is nonbinding, because it is satisfied as a strong inequality and the investors' equilibrium payoff does not fall if the constraint is tightened a little by making the producer's earnings from embezzlement slightly higher. The cost of the contract to the investors is 100 in equilibrium, so their overall expected payoff is $0.5(-100) + 0.5(+500) - 100 = 100$, an amount greater than zero and thus yielding enough return for the show to be profitable.

The boiling-in-oil contract is an application of the **sufficient statistic condition**, which says that for incentive purposes, if the agent's utility function is separable in effort and money, wages should be based on whatever evidence best indicates effort, and only incidentally on output (see Holmstrom [1979] and note N7.2). In the spirit of the three-step procedure, what the principal wants is to induce the agent to choose the appropriate effort, *Do not embezzle*, and his data on what the agent chose is the output. In equilibrium (though not out of it), the datum $q = +500$ contains exactly the same information as the datum $q = -100$. Both lead to the same posterior probability that the agent chose *Do not embezzle*, so the wages conditioned on each datum should be the same. We need to insert the qualifier "in equilibrium," because to form the posterior probabilities the principal needs to have some beliefs as to the agent's behavior. Otherwise, the principal could not interpret $q = -100$ at all.

Milder contracts would also be effective. Two wages will be used in equilibrium, a low wage $\underline{w}$ for an output of $q = 100$ and a high wage $\overline{w}$ for any other output. The participation and incentive compatibility constraints provide two equations to solve for these two unknowns. To find the mildest possible contract, the modeller must also specify a function for utility $U(w)$, something which, interestingly enough, was unnecessary for finding the first boiling-in-oil contract. Let us specify that

$$U(w) = 100w - 0.1w^2. \tag{32}$$

A quadratic utility function like this is only increasing if its argument is not too large, but since the wage will not exceed $w = 1000$, it is a reasonable utility function for this model. Substituting (32) into the participation constraint (31) and solving for the full-insurance high wage $\overline{w} = w(-100) = w(+500)$ yields $\overline{w} = 100$ and a reservation utility of 9000. Substituting into the incentive compatibility constraint, (30), yields

$$9000 \geq 0.5U(100 + 50) + 0.5U(\underline{w} + 50). \tag{33}$$

When (33) is solved using the quadratic equation, it yields (with rounding error), $\underline{w} \leq 5.6$. A low wage of $-\infty$ is far more severe than what is needed.

If both the producer and the investors were risk averse, risk sharing would change the part of the contract that applied in equilibrium. The optimal contract would then provide for $w(-100) < w(+500)$ to share the risk. The principal would have a lower marginal utility of wealth when output was $+500$, so he would be better able to pay an extra dollar of wages in that state than when output was -100.

One of the oddities of Broadway Game I is that the wage is higher for an output of -100 than for an output of $+100$. This illustrates the idea that the principal's aim is to reward input, not output. If the principal pays more simply because output is higher, he is rewarding Nature, not the agent. People usually believe that higher pay for higher output is "fair," but Broadway Game I shows that this ethical view is too simple. Higher effort usually leads to higher output, but not always. Thus, higher pay is usually a good incentive, but not always, and sometimes low pay for high output actually punishes slacking.

The decoupling of reward and result has broad applications. Becker (1968) in criminal law and Polinsky & Che (1991) in tort law note that if society's objective is to keep the amount of enforcement costs and harmful behavior low, the penalty applied should not simply be matched to the harm. Very high penalties seldom inflicted will provide the proper incentives and keep enforcement costs low, even though a few unlucky offenders will receive penalties out of proportion to the harm they caused.

A less gaudy name for a boiling-in-oil contract is the alliterative "**shifting support scheme**," so named because the contract depends on the support of the output distribution being different when effort is optimal than when effort is other than optimal. The set of possible outcomes under optimal effort must be different from the set of possible outcomes under any other effort level. As a result, certain outputs show without doubt that the producer embezzled. Very heavy punishments inflicted only for those outputs achieve the first-best because a non-embezzling producer has nothing to fear.

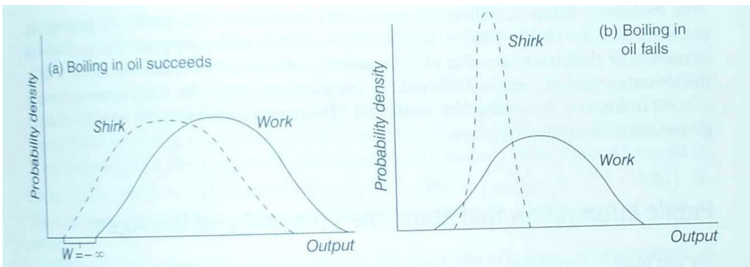

Figure 4: Shifting Supports in an Agency Model

Figure 4a shows shifting supports in a model where output can take not three but a continuum of values. If the agent shirks instead of working, certain low outputs become possible and certain high outputs become impossible. In a case like this, where the support of the output shifts when behavior changes, boiling-in-oil contracts are useful: the wage is $-\infty$ for the low outputs possible only under shirking. In Figure 4b, on the other hand, the support just shrinks under shirking, so boiling in oil is inappropriate. When there is a limit to the amount the agent can be punished, or the support under shirking is a subset of the first-best support, the threat of boiling-in-oil might not achieve the first-best. Sometimes, however, similar contracts can still be used. The conditions favoring such contracts are

1 The agent is not very risk averse.

2 There are outcomes with high probability under shirking that have low probability under optimal effort.

3 The agent can be severely punished.

4 It is credible that the principal will carry out the severe punishment.

Selling the Store

Another first-best contract that can sometimes be used is **selling the store.** Under this arrangement, the agent buys the entire output for a flat fee paid to the principal, becoming the **residual claimant**, since he keeps every additional dollar of output that his extra effort produces. This is equivalent to fully insuring the principal, since his payoff becomes independent of the moves of the agent and of Nature.

In Broadway Game I, selling the store takes the form of the producer paying the investors 100 ($= 0.5[-100] + 0.5[+500] - 100$) and keeping all the profits for himself. The drawbacks are that (1) the producer might not be able to afford to pay the investors the flat price of 100; and (2) the producer might be risk averse and incur a heavy utility cost in bearing the entire risk. These two drawbacks are why producers go to investors in the first place.

Public Information that Hurts the Principal and the Agent

We can modify Broadway Game I to show how having more public information available can hurt both players. This will also provide a little practice in using information sets. Let us split *Success* into two states of the world, *Minor Success* and *Major Success*, which have probabilities 0.3 and 0.2 as shown in Table 4.

Table 4: Profits in Broadway Game II

		State of the World		
		Failure (0.5)	*Minor Success* (0.3)	*Major Success* (0.2)
Effort	*Embezzle*	-100	-100	$+400$
	Do not embezzle	-100	$+450$	$+575$

Under the optimal contract,

$$w(-100) = w(+450) = w(+575) > w(+400) + 50. \tag{34}$$

This is so because the producer is risk averse and only the datum $q = +400$ is proof that the producer embezzled. The optimal contract must do two things: deter embezzlement and pay the producer as predictable a wage as possible. For predictability, the wage is made constant unless $q = +400$. To deter embezzlement, the producer must be punished if $q = +400$. As in Broadway Game I, the punishment would not have to be infinitely severe, and the minimum effective punishment could be calculated in the same way as in that game. The investors would pay the producer a wage of 100 in equilibrium and their expected payoff would be 100 ($= 0.5(-100) + 0.3(450) + 0.2(575) - 100$). Thus, a contract can be found for Broadway Game II such that the agent would not embezzle.

But consider what happens when the information set is refined so that before the agent takes his action both he and the principal can tell whether the show will be a major success or not. Let us call this Broadway Game III. Under the refinement, each player's information partition is

$$(\{Failure,\ Minor\ Success\},\ \{Major\ Success\}),$$

instead of Broadway Game I and II's coarse partition

$$(\{Failure, \textit{Minor Success}, \textit{Major Success}\}).$$

If the information sets were refined all the way to singletons, this would be very useful to the investors, because they could abstain from investing in a failure and they could easily determine whether the producer embezzled or not. As it is, however, the refinement does not help the investors decide when to finance the show. If they could still hire the producer and prevent him from embezzling at a cost of 100, the payoff from investing in a major success would be 475 (= 575 − 100). But the payoff from investing in a show given the information set $\{Failure, \textit{Minor Success}\}$ would be about 6.25, which is still positive $\left(\left(\frac{0.5}{0.5+0.3}\right)(-100) + \left(\frac{0.3}{0.5+0.3}\right)(450) - 100\right)$. So the improvement in information is no help with respect to the decision of when to invest.

The refinement does, however, ruin the producer's incentives. If he observes $\{Failure, \textit{Minor Success}\}$, he is free to embezzle without fear of the oil-boiling output of +400. He would still refrain from embezzling if he observed $\{\textit{Major Success}\}$, but no contract that does not impose risk on a nonembezzling producer can stop him from embezzling if he observes $\{Failure, \textit{Minor Success}\}$. Whether a risky contract can be found that would prevent the producer from embezzling at a cost of less than 6.25 to the investors depends on the producer's risk aversion. If he is very risk averse, the cost of the incentive is more than 6.25, and the investors will give up investing in shows that might be minor successes. Better information reduces welfare, because it increases the producer's temptation to misbehave.

Notes

N7.1 Categories of Asymmetric Information Models

- Three books that cover the theory of contracts are Macho-Stadler and Perez- Castillo's 1997 *An Introduction to the Economics of Information: Incentives and Contracts*, Salanie's 1997 *The Economics of Contracts: A Primer* and Bolton and Dewatripont's 2005 *Contract Theory.* For a survey of recent empirical work on contracts, see Chiappori, P.A. & B. Salanie (2003).

- A large literature of nonmathematical theoretical papers looks at organizational structure in the light of the agency problem. See Alchian & Demsetz (1972), Fama (1980), and Klein, Crawford, & Alchian (1978). Milgrom & Roberts (1992) have written a book on organization theory that describes what has been learned about the principal-agent problem at a technical level that MBA students can understand. There may be much to be learned from intelligent economists of the past also; note that part III, chapter 8, section 12 of Pigou's *Economics of Welfare* (1932/1920) has an interesting discussion of the advantage of piece-rate work, which can more easily induce each worker to choose the correct effort when abilities differ as well as efforts. I recommend Chapter 1 ("Incentives in Economic Thought") of Laffont & Martimort (2002) as a short survey of the history of incentives in economic theory.

- An important consideration in real-world contracts is manipulation and fraud by the agent. If an executive is rewarded for higher earnings per share, and he controls the accountants who measure earnings, for example, he may tell them to measure them so that they surpass the threshold in his threshold contract. Even worse, he might cut prices so as to speed up sales before January 1 and meet his yearly target. Jensen's 2003 "Paying People To Lie: The Truth about the Budgeting Process," points this out and notes that linear contracts are less vulnerable to this type of manipulation.

- We have lots of "prinsipuls" in economics. I find this paradigm helpful for remembering spelling: "The principal's principal principle was to preserve his principal." More simply, "The principal is my pal," takes care of the "al" case.

- "Principal" and "Agent" are legal terms, and agency is an important area of the law. Economists have focussed on quite different questions than lawyers. Economists focus on effort: how the principal induces the agent to do things. Lawyers focus on malfeasance and third parties: how the principal stops the agent from doing the wrong things and who bears the burden if he fails. If, for example, the manager of a tavern enters into a supply contract against the express command of the owner, who must be disappointed— the owner or the third party supplier?

- *Double-sided moral hazard.* The text described one-sided moral hazard. Moral hazard can also be double-sided, as when each player takes actions unobservable by the other that affect the payoffs of both of them. An example is tort negligence by both plaintiff and defendant: if a careless auto driver hits a careless pedestrian, and they go to law, the court must try to allocate blame, and the legislature must try to set up laws to induce the proper amount of care. Landlords and tenants also face double moral hazard, as implied in Table 1.

- A common convention in principal-agent models is to make one player male and the other female, so that "his" and "her" can be used to distinguish between them. I find this distracting, since gender is irrelevant to most models and adds one more detail for the reader to keep track of. If readers naturally thought "male" when they saw "principal," this would not be a problem— but they do not.

N7.2 A Principal-Agent Model: The Production Game

- The model in the text uses "effort" as the action taken by the agent, but effort is used to represent a variety of real-world actions. The cost of pilferage by employees is an estimated \$8 billion a year in the USA. Employers have offered rewards for detection, one even offering the option of a year of twice-weekly lottery tickets instead of a lump sum. The Chicago department store Marshall Field's, with 14,000 workers, in one year gave out 170 rewards of \$500 each, catching almost 500 dishonest employees. ("Hotlines and Hefty Rewards: Retailers Step Up Efforts to Curb Employee Theft," *Wall Street Journal*, September 17,1987, p. 37.)

  For an illustration of the variety of kinds of "low effort," see "Hermann Hospital Estate, Founded for the Poor, Has Benefited the Wealthy, Investigators Allege," *Wall Street Journal*, March 13, 1985, p. 4, which describes such forms of misbehavior as pleasure trips on company funds, high salaries, contracts for redecorating awarded to girlfriends, phony checks, kicking back real estate commissions, and investing in friendly companies. Nonprofit enterprises, often lacking both principles and principals, are especially vulnerable, as are governments, for the same reason.

- The Production Game assumes that the agent dislikes effort. Is this realistic? People differ. My father tells of his experience in the navy when the sailors were kept busy by being ordered to scrape loose paint. My father found it a way to pass the time but says that other sailors would stop chipping when they were not watched, preferring to stare into space. *De gustibus non est disputandum* ("About tastes there can be no arguing"). But even if effort has positive marginal utility at low levels, it has negative marginal utility at high enough levels— including, perhaps, at the efficient level. This is as true for professors as for sailors.

- Suppose that the principal does not observe the variable θ (which might be effort), but he does observe t and x (which might be output and profits). From Holmstrom (1979) and Shavell (1979) we have, restated in my words,

  The Sufficient Statistic Condition. *If t is a sufficient statistic for θ relative to x, then the optimal contract needs to be based only on t if both principal and agent have separable utility functions.*

  The variable t *is a* **sufficient statistic** *for θ relative to x if, for all t and x,*

$$Prob(\theta|t,x) = Prob(\theta|t). \tag{35}$$

  This implies, from Bayes' Rule, that $Prob(t,x|\theta) = Prob(x|t)Prob(t|\theta)$; that is, x depends on θ only because x depends on t and t depends on θ.

  The sufficient statistic condition is closely related to the Rao- Blackwell Theorem (see Cox & Hinkley [1974] p. 258), which says that the decision rule for nonstrategic decisions ought not to be random.

Gjesdal (1982) notes that if the utility functions are not separable, the theorem does not apply and randomized contracts may be optimal. Suppose there are two actions the agent might take. The principal prefers action X, which reduces the agent's risk aversion, to action Y, which increases it. The principal could offer a randomized wage contract, so the agent would choose action X and make himself less risk averse. This randomization is not a mixed strategy. The principal is not indifferent to high and low wages; he prefers to pay a low wage, but we allow him to commit to a random wage earlier in the game.

N7.3 The Incentive Compatibility and Participation Constraints

- The term, "individual rationality constraint," is perhaps more common, but "participation constraint" is more sensible. Since in modern modelling every constraint requires individuals to be rational, the name is ill-chosen.

- **Paying the agent more than his reservation wage.** If agents compete to work for principals, the participation constraint is binding whenever there are only two possible outcomes or whenever the agent's utility function is separable in effort and wages. Otherwise, it might happen that the principal picks a contract giving the agent more expected utility than is necessary to keep him from quitting. The reason is that the principal not only wants to keep the agent working, but to choose a high effort.

- If the distribution of output satisfies the **monotone likelihood ratio property** (MLRP), the optimal contract specifies higher pay for higher output. Let $f(q|e)$ be the probability density of output. The MLRP is satisfied if

$$\forall e' > e, \text{ and } q' > q, \quad f(q'|e')f(q|e) - f(q'|e)f(q|e') > 0, \tag{36}$$

  or, in other words, if when $e' > e$, the ratio $f(q|e')/f(q|e)$ is increasing in q. Alternatively, f satisfies the MLRP if $q' > q$ implies that q' is a more favorable message than q in the sense of Milgrom (1981b). Less formally, the MLRP is satisfied if the ratio of the likelihood of a high effort to a low effort rises with observed output. The distributions in The Broadway Game of Section 7.4 violate the MLRP, but the normal, exponential, Poisson, uniform, and chi-square distributions all satisfy it. Stochastic dominance does not imply the MLRP. If effort of 0 produces outputs of 10 or 12 with equal probability, and effort of 1 produces outputs of 11 or 13 also with equal probability, the second distribution is stochastically dominant, but the MLRP is not satisfied. See Chapter 13 (Auctions) for more on the MLRP.

N7.4 Optimal Contracts: The Broadway Game

- Daniel Asquith suggested the idea behind Broadway Game II.

- Franchising is one compromise between selling the store and paying a flat wage. See Mathewson & Winter (1985), Rubin (1978), and Klein & Saft (1985).

- Mirrlees (1974) is an early reference on the idea of the boiling-in-oil contract.

- Broadway Game II shows that improved information could reduce welfare by increasing a player's incentive to misbehave. This is distinct from the nonstrategic insurance reason why improved information can be harmful. Suppose that Smith is insuring Jones against hail ruining Jones' wheat crop during the next year, increasing Jones' expected utility and giving a profit to Smith. If someone comes up with a way to forecast the weather before the insurance contract is agreed upon, both players will be hurt. Insurance will break down, because if it is known that hail will ruin the crop, Smith will not agree to share the loss, and if it is known there will be no hail, Jones will not pay a premium for insurance. Both players prefer not knowing the outcome in advance.

Problems

7.1: First-Best Solutions in a Principal-Agent Model (easy)

Suppose an agent has the utility function of $U = \sqrt{w} - e$, where e can assume the levels 0 or 1. Let the reservation utility level be $\overline{U} = 3$. The principal is risk neutral. Denote the agent's wage, conditioned on output, as $\underline{w}$ if output is 0 and $\overline{w}$ if output is 100. Table 5 shows the outputs.

Table 5: A Moral Hazard Game

Effort	Probability of Output of 0	100	Total
Low ($e = 0$)	0.3	0.7	1
High ($e = 1$)	0.1	0.9	1

(a) What would the agent's effort choice and utility be if he owned the firm?

(b) If agents are scarce and principals compete for them, what will the agent's contract be under full information? His utility?

(c) If principals are scarce and agents compete to work for them, what would the contract be under full information? What will the agent's utility and the principal's profit be in this situation?

(d) Suppose that $U = w - e$. If principals are the scarce factor and agents compete to work for principals, what would the contract be when the principal cannot observe effort? (Negative wages are allowed.) What will be the agent's utility and the principal's profit be in this situation?

7.2: The Principal-Agent Problem (medium)

Suppose the agent has a utility function of $U = \sqrt{w} - e$, where e can assume the levels 0 or 7, and a reservation utility of $\overline{U} = 4$. The principal is risk neutral. Denote the agent's wage, conditioned on output, as $\underline{w}$ if output is 0 and $\overline{w}$ if output is 1,000. Only the agent observes his effort. Principals compete for agents. Table 6 shows the output.

Table 6: Output from Low and High Effort

Effort	Probability of output of 0	1,000	Total
Low($e = 0$)	0.9	0.1	1
High ($e = 7$)	0.2	0.8	1

(a) What are the incentive compatibility, participation, and zero- profit constraints for obtaining high effort?

(b) What would utility be if the wage were fixed and could not depend on output or effort?

(c) What is the optimal contract? What is the agent's utility?

(d) What would the agent's utility be under full information? Under asymmetric information, what is the agency cost (the lost utility) as a percentage of the utility the agent receives?

7.3. Why Entrepreneurs Sell Out (medium)
Suppose an agent has a utility function of $U = \sqrt{w} - e$, where e can assume the levels 0 or 2.4, and his reservation utility is $\overline{U} = 7$. The principal is risk neutral. Denote the agent's wage, conditioned on output, as $w(0)$, $w(49)$, $w(100)$, or $w(225)$. Table 7 shows the output.

Table 7: Entrepreneurs Selling Out

	Probability of output of				
Method	0	49	100	225	Total
$Safe$ $(e = 0)$	0.1	0.1	0.8	0	1
$Risky$ $(e = 2.4)$	0	0.5	0	0.5	1

(a) What would the agent's effort choice and utility be if he owned the firm?

(b) If agents are scarce and principals compete for them, what will the agent's contract be under full information? His utility?

(c) If principals are scarce and agents compete to work for principals, what will the contract be under full information? What will the agent's utility and the principal's profit be in this situation?

(d) If agents are the scarce factor, and principals compete for them, what will the contract be when the principal cannot observe effort? What will the agent's utility and the principal's profit be in this situation?

7.4. Authority (medium)
A salesman must decide how hard to work on his own time on getting to know a potential customer. If he exerts effort X, he incurs a utility cost $X^2/2$. With probability X, he can then go to customer X and add V to his own earnings. With probability (1-X), he offends the customer, and on going to him would subtract L from his earnings. The boss will receive benefit B from the sale in either case. The ranking of these numbers is $V > L > B > 0$. The boss and the salesman have equal bargaining power, and are free to make side payments to each other.

(a) What is the first-best level of effort, X_a?

(b) If the boss has the authority to block the salesman from selling to this customer, but cannot force him to sell, what value will X take?

(c) If the salesman has the authority over the decision on whether to sell to this customer, and can bargain for higher pay, what will his effort be?

(d) Rank the effort levels X_a, X_b, and X_c in the previous three sections.

7.5. Worker Effort (easy)

A worker can be $Careful$ or $Careless$, efforts which generate mistakes with probabilities 0.25 and 0.75. His utility function is $U = 100 - 10/w - x$, where w is his wage and x takes the value 2 if he is careful, and 0 otherwise. Whether a mistake is made is contractible, but effort is not. Risk-neutral employers compete for the worker, and his output is worth 0 if a mistake is made and 20 otherwise. No computation is needed for any part of this problem.

(a) Will the worker be paid anything if he makes a mistake?

(b) Will the worker be paid more if he does not make a mistake?

(c) How would the contract be affected if employers were also risk averse?

(d) What would the contract look like if a third category, "slight mistake," with an output of 19, occurs with probability 0.1 after $Careless$ effort and with probability zero after $Careful$ effort?

7.6. The Supercomputer Salesman (medium)

If a salesman exerts high effort, he will sell a supercomputer this year with probability 0.9. If he exerts low effort, he will succeed with probability 0.5. The company will make a profit of 2 million dollars if the sale is made. The salesman would require a wage of \$50,000 if he had to exert low effort, but \$70,000 if he had to exert high effort, he is risk neutral, and his utility is separable in effort and money. (Let's just use payoffs in thousands of dollars, so 70,000 dollars will be written as 70, and 2 million dollars will be 2000)

(a) Prove that high effort is first-best efficient.

(b) How high would the probability of success with low effort have to be for high effort to be inefficient?

(c) If you cannot monitor the programmer and cannot pay him a wage contingent on success, what should you do?

(d) Now suppose you can make the wage contingent on success. Let the wage be S if he makes a sale and F if he does not. S and F will have to satisfy two conditions: a participation constraint and an incentive compatibility constraint. What are they?

(e) What is a contract that will achieve the first best?

(f) Now suppose the salesman is risk averse, and his utility from money is $log(w)$. Set up the participation and incentive compatibility constraints again.

(g) You do not need to solve for the optimal contract. Using the $log(w)$ utility function assumption, however, will the expected payment by the firm in the optimal contract rise, fall, or stay the same, compared with what it was in part (e) for the risk neutral salesman?

(h) You do not need to solve for the optimal contract. Using the $log(w)$ utility function assumption, however, will the gap between S and F in the optimal contract rise, fall, or stay the same, compared with what it was in part (e) for the risk neutral salesman?

7.7. Optimal Compensation (easy)

An agent's utility function is U =(log(wage) - effort). What should his compensation scheme be if different (output,effort) pairs have the probabilities in Table 8?

(a) The agent should be paid exactly his output.

(b) The same wage should be paid for outputs of 1 and 100.

(c) The agent should receive more for an output of 100 than of 1, but should receive still lower pay if output is 2.

(d) None of the above.

Table 8: Output Probabilities

		Output		
		1	2	100
Effort	*High*	0.5	0	0.5
	Low	0.1	0.8	0.1

7.8. Effort and Output, Multiple Choices (easy)

The utility function of the agents whose situation is depicted in Table 9 is $U = w + \sqrt{w} - \alpha e$, and his reservation utility is 0. Principals compete for agents, and have reservation profits of zero. Principals are risk neutral.

Table 9: Output Probabilities

		Effort	
		Low ($e = 0$)	High ($e = 5$)
Output	$y = 0$	0.9	0.5
	$y = 100$	0.1	0.5

(a) If $\alpha = 2$, then if the agent's action can be observed by the principal, his equilibrium utility is in the interval

(a) $[-\infty, 0.5]$

(b) $[0.5, 5]$

(c) $[5, 10]$

(d) $[10, 40]$

(e) $[40, \infty]$

(b) If $\alpha = 10$, then if the agent's action can be observed by the principal, his equilibrium utility is in the interval
(a) $[-\infty, 0.5]$
(b) $[0.5, 5]$
(c) $[5, 10]$
(d) $[10, 40]$
(e) $[40, \infty]$

(c) If $\alpha = 5$, then if the agent's action can be observed by the principal, his equilibrium effort level is
(a) Low
(b) High
(c) A mixed strategy effort, sometimes low and sometimes high

(d) If $\alpha = 2$, then if the agent's action cannot be observed by the principal, and he must be paid a flat wage, his wage will be in the interval
(a) $[-\infty, 2]$
(b) $[2, 5]$
(c) $[5, 8]$
(d) $[8, 12]$
(e) $[12, \infty]$

(e) If the agent owns the firm, and $\alpha = 2$, will his utility be higher or lower than in the case where he works for the principal and his action can be observed?
(a) Higher
(b) Lower
(c) Exactly the same.

(f) If the agent owns the firm, and $\alpha = 2$, his equilibrium utility is in the interval
(a) $[-\infty, 0.5]$
(b) $[0.5, 5]$
(c) $[5, 10]$
(d) $[10, 40]$
(e) $[40, \infty]$

(g) If the agent owns the firm, and $\alpha = 8$, his equilibrium utility is in the interval
(a) $[-\infty, 0.5]$
(b) $[0.5, 5]$
(c) $[5, 10]$
(d) $[10, 40]$
(e) $[40, \infty]$

7.9. Hiring a Lawyer (easy)

A one-man firm with concave utility function $U(X)$ hires a lawyer to sue a customer for breach of contract. The lawyer is risk-neutral and effort averse, with a convex disutility of effort. What can you say about the optimal contract? What would be the practical problem with such a contract, if it were legal?

7.10. Constraints (medium)

An agent has the utility function $U = log(w) - e$, where e can take the levels 0 and 4, and his reservation utility is $\overline{U} = 4$. His principal is risk-neutral. Denote the agent's wage conditioned on output as $\underline{w}$ if output is 0 and $\overline{w}$ if output is 10. Only the agent observes his effort. Principals compete for agents. Output is as shown in Table 10.

Table 10: Effort and Outputs

	Probability of Outputs		
Effort	0	10	Total
Low$(e = 0)$	0.9	0.1	1
High $(e = 4)$	0.2	0.8	1

What are the incentive compatibility and participation constraints for obtaining high effort?

7.11. Constraints Again (medium)

Suppose an agent has the utility function $U = log(w) - e$, where e can take the levels 1 or 3, and a reservation utility of $\overline{U}$. The principal is risk-neutral. Denote the agent's wage conditioned on output as $\underline{w}$ if output is 0 and $\overline{w}$ if output is 100. Only the agent observes his effort. Principals compete for agents, and outputs occur according to Table 11.

Table 11: Efforts and Outputs

	Probability of Outputs	
Effort	0	100
Low$(e = 1)$	0.9	0.1
High $(e = 3)$	0.5	0.5

What conditions must the optimal contract satisfy, given that the principal can only observe output, not effort? You do not need to solve out for the optimal contract– just provide the equations which would have to be true. Do not just provide inequalities– if the condition is a binding constraint, state it as an equation.

Moral Hazard: A Classroom Game for Chapter 7

Each student works as a salesman for Apex, Brydox, or neither firm, choosing anew each year. Each year you also pick your effort level, which is unobserved by the firms. Your sales equal

$$Q = 2 + e + u,$$

where u takes the values -2 and $+2$ with equal probability.

Your payoff is 600 if you work for neither firm and otherwise is a function of your wage and effort:

$$Payoff = V(w) - e^2, \tag{37}$$

where $V(w)$ is shown in the following table:

w	<0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	≥ 14
$V(w)$	0	0	100	190	370	540	690	830	930	1000	1020	1030	1038	1044	1046	1046

You will have limited time to make your choices of contract and effort. If you do not hand in a card with your choice by the deadline, your default choices are to work for neither firm (if you fail to choose an employer) and $e = 0$ (if you don't choose your effort).

A **linear contract** takes the form,

$$w(Q) = \alpha + \beta Q, \tag{38}$$

where if the value of w from the equation is not an integer it is rounded up.

A **threshold contract** takes the form,

$$w(Q) = \alpha \; if \; Q \geq \beta; \quad w = 0 \text{ otherwise} \tag{39}$$

A **monitoring contract** takes the form

$$w(Q) = \alpha \text{ unless you are caught with } e < \beta; \quad w = 0 \text{ otherwise; probability of monitoring} = \gamma \tag{40}$$