September 4, 1999

7 Moral Hazard: Hidden Actions

7.1 Categories of Asymmetric Information Models

(1) **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.

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

(3) **Mechanism Design in Adverse Selection and in Moral Hazard with Hidden Information**) (Chapter 9A)
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 moral hazard with hidden information–in which case Smith will learn it after contracting.

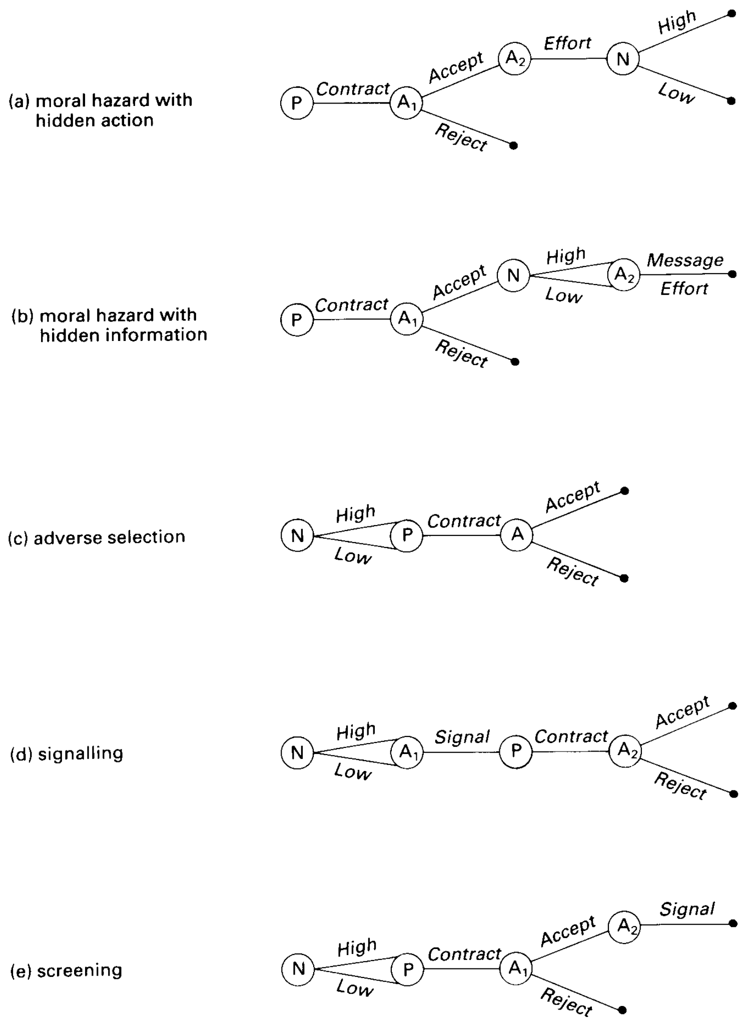
(4, 5) **Signalling** and **Screening** (Chapter 10)
Nature begins the game by choosing Smith's type, unobserved by Jones. To demonstrate his type, Smith takes actions that Jones can observe. If Smith takes the action before they agree to a contract, he is signalling; if he takes it afterwards, he is being screened. Information is incomplete.

The **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 7.1 Categories of Asymmetric Information Models

Figure 7.1 Categories of Asymmetric Information Models



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 $\pi_{principal} = V(q - w)$.

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 how to give him incentives to do it as cheaply as possible.

The agent must be paid some amount $\tilde{w}(e)$ to exert effort e, where $\tilde{w}(e)$ is defined to be the w that solves the participation constraint

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

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 the implicit function theorem (see Section 13.4) and the participation constraint,

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

Combining equations (7.4) and (7.5) yields

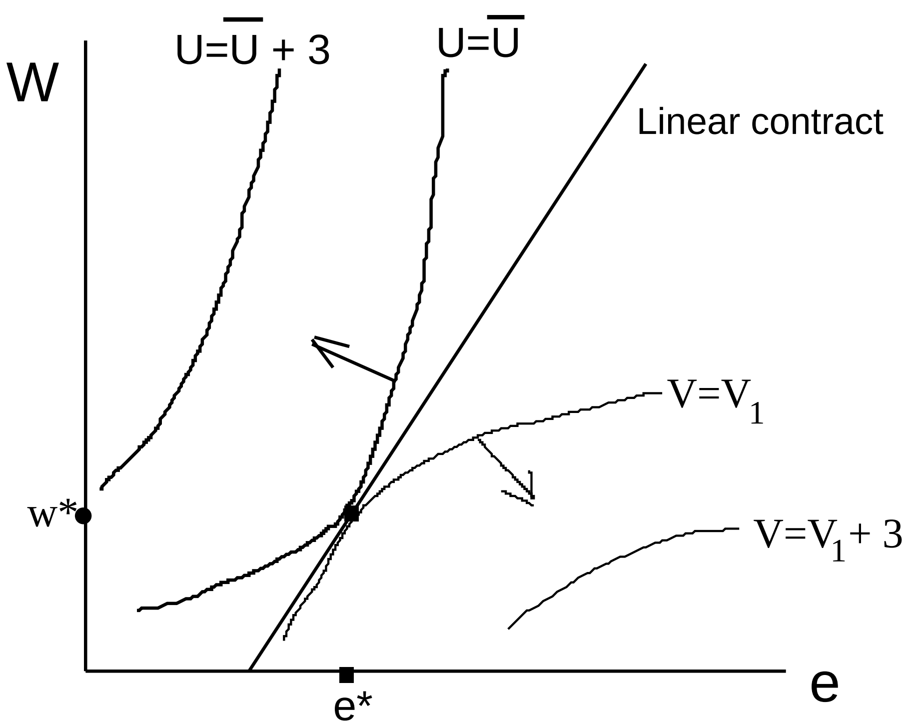
$$\frac{\partial U}{\partial \tilde{w}} \frac{\partial q}{\partial e} = - \frac{\partial U}{\partial e}. \tag{6}$$

Equation (7.6) 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. As usual, the outcome can be efficient even if the agent does not actually keep the extra output, since who keeps the output is a distributional question.

Figure 7.2 shows this graphically. The agent has indifference curves in effort-wage space that slope upwards, since if effort is increased the wage must be increased to keep 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 7.2 shows a technology with diminishing returns to effort. If effort starts out being higher, extra effort yields less additional output so the wage cannot rise as much without reducing profits.

Figure 7.2 The Efficient Effort Level in Production Game I

Figure 7.2 The Efficient Effort Level in "Production Game I"



Under perfect competition among the principals the profits are zero, so the reservation utility $\overline{U}$ is chosen so 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, which is at the 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 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 Figure 7.2 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 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 $q(e) = 100 * log(1+e)$. If the agent rejects the contract, then $\pi_{agent} = \bar{U} = 3$ and $\pi_{principal} = 0$, whereas if the agent accepts the contract, then $\pi_{agent} = U(e, w) = log(w) - e^2$ 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 solve the participation constraint,

$$U(e, w(e)) = \overline{U}, \quad so \;\; log(\tilde{w}(e)) - e^2 = 3. \tag{8}$$

Knowing the particular functional form as we do, we can solve (7.8) for the wage function:

$$\tilde{w}(e) = Exp(3 + e^2). \tag{9}$$

This 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.

The principal's problem is

$$\underset{e}{Maximize} \;\; V(q(e) - \tilde{w}(e)) = 100 * log(1+e) - Exp(3 + e^2) \tag{10}$$

The first order condition for this problem is

$$V'(q(e) - \tilde{w}(e)) \left(\frac{\partial q}{\partial e} - \frac{\partial \tilde{w}}{\partial e} \right) = 0, \tag{11}$$

or, for our problem, since the firm is risk-neutral and $V' = 1$,

$$\frac{100}{1+e} - 2e(Exp(3 + e^2)) = 0, \tag{12}$$

We can simplify the first order condition a little to get

$$(2e + 2e^2)Exp(3 + e^2) = 100, \tag{13}$$

but this cannot be solved analytically. Using the computer program Mathematica, I found that $e^* \approx 0.84934$, from which, using the formulas above, we get $q^* \approx 100 * log(1 + .84934) \approx 61.48$ and $w^* \approx 41.32$.

To implement the contract, a number of types of contracts could be used.

(1) The **forcing contract** sets $w(e^*) = w^*$ and $w(e \neq .84) = 0$. Here, $w(.84) = 41$ and $w(e \neq e^*) = 0$.

(2) The **threshold contract** sets $w(e \geq e^*) = w^*$ and $w(e < e^*) = 0$. Here, $w(e \geq .84) = 41$ and $w(e < .84) = 0$.

(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

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

At $e^* = .84$, this takes the value 253. That is the β for the linear contract. The α must solve $w(e^*) = 41 = \alpha + 253 * .84$, so $\alpha = -171.52$.

You ought to be a little concerned that the linear contract satisfies the incentive compatibility constraint. We constructed it so that it satisfied the participation constraint, because if the agent chooses $e = .84$, his utility will be 3. But might he prefer to choose some larger or smaller e and get even more utility?

He will not, 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, as the diagram here shows.

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 $q' > 0$.

In this game, the agent has all the bargaining power, not the principal. The participation constraint is now that the principal 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 the participation constraint, the maximization problem for the agent in proposing an effort level e at a wage $w(e)$ can therefore be written as

$$\underset{e}{Maximize} \;\; U(e, q(e)) \tag{15}$$

The first-order condition is

$$\frac{\partial U}{\partial e} + \left(\frac{\partial U}{\partial q}\right)\left(\frac{\partial q}{\partial e}\right) = 0. \tag{16}$$

Since $\frac{\partial U}{\partial q} = \frac{\partial U}{\partial w}$ when the wages equals output, equation (7.16) implies that

$$\frac{\partial U}{\partial w}\frac{\partial q}{\partial e} = -\frac{\partial U}{\partial e}. \tag{17}$$

Comparing this with equation (7.6), the equation when the principal had the bargaining power, it is clear that e^* is identical in Production Game I and Production Game II. It does not matter who has the bargaining power; the efficient effort level stays the same.

Figure 7.2 can be used to illustrate this game as well. Suppose that $V_1 = 0$. The agent must choose a point on the $V_1 = 0$ indifference curve that maximizes his own utility, and then provide himself with contract incentives to choose that point. The agent's payoff is highest at effort e^* given that he must make $V_1 = 0$, and all three contracts described in Production Game I provide him with the correct incentives.

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 that of the Coase Theorem, which says that 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.

Production Game III: 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.

Output is not **contractible** or **verifiable**, which leads to the same outcome as when it is unobservable in a contracting model. 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.

Production Game IV: Output-Based Wage under Certainty. In this version, the principal cannot observe effort but can observe output and specify the contract to be $w(q)$.

Now the principal picks not a number w but a function $w(q)$. 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 V: 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 the 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). Moral hazard becomes a problem only when $q(e)$ is not a one-to-one function because 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.

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 world and the second-best world is the cost of the agency problem.

7.3 The Incentive Compatibility, Participation, and Competition 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

$$\underset{w(\cdot)}{Maximize} \;\; EV(q(\tilde{e},\theta) - w(q(\tilde{e},\theta))) \tag{18}$$

subject to

$$\tilde{e} = \overset{argmax}{e} \;\; EU(e, w(q(e,\theta))) \quad \text{(incentive compatibility constraint) (7. 18a)}$$

$$EU(\tilde{e}, w(q(\tilde{e},\theta))) \geq \bar{U} \qquad \text{(participation constraint) (7. 18b)}$$

To support the effort level e, the wage contract $w(\cdot)$ 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.

$$C(\tilde{e}) = \underset{w(\cdot)}{Minimum} \;\; Ew(q(\tilde{e},\theta)) \tag{19}$$

subject to constraints (7.11a) and (7.11b).

Step three takes the principal's problem of maximizing his payoff, (7.18), and restates it as

$$\underset{\tilde{e}}{Maximize} \;\; EV(q(\tilde{e},\theta) - C(\tilde{e})). \tag{20}$$

After finding which contract most cheaply induces each effort, the principal discovers the optimal effort by solving problem (7.20).

7.4 Optimal Contracts: The Broadway Game

"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 7.2 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 7.2 Profits in Broadway Game I

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

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

$$\pi(Do\ not\ embezzle) = 0.5U(w(-100)) + 0.5U(w(+500)) \quad (21)$$

and

$$\pi(Embezzle) = 0.5U(w(-100) + 50) + 0.5U(w(+100) + 50). \quad (22)$$

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

$$0.5U(w(-100)) + 0.5U(w(+500)) \geq 0.5U(w(-100)+50) + 0.5U(w(+100)+50), \quad (23)$$

and the participation constraint is

$$\pi(Do\ not\ embezzle) = 0.5U(w(-100)) + 0.5U(w(+500)) \geq U(100). \quad (24)$$

The investors want the participation constraint (7.24) 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 value for his wage if the risk is higher. 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.

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

Milder contracts than this 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 $U(w)$ which, interestingly enough, was unnecessary for finding the first boiling-in-oil contract. Let us specify that

$$U(w) = 100w - 0.1w^2. \tag{25}$$

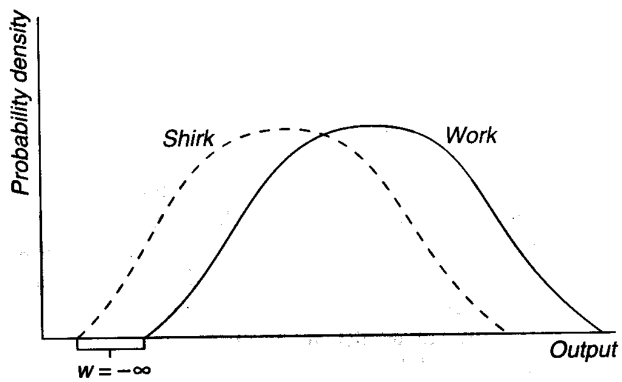
A quadratic utility function like this is only increasing if its argument is not too large, but since the wage will not exceed $w = 1000$, it is a reasonable utility function for this model. Substituting (7.25) into the participation constraint (7.24) and solving for the full-insurance high wage $\overline{w} = w(-100) = w(+500)$ yields $\overline{w} = 100$ and a reservation utility of 9000. Substituting into the incentive compatibility constraint, (7.23), yields

$$9000 \geq 0.5U(100 + 50) + 0.5U(\underline{w} + 50). \tag{26}$$

When (7.26) 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.

Figure 7.3 Shifting Supports in an Agency Model

Figure 7.3 Shifting Supports in an Agency Model



Boiling in Oil is useful when:

(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.

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 7.4.

Table 7.4 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,

$$w(-100) = w(+450) = w(+575) > w(+400) + 50. \qquad (27)$$

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 initial information partition is

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

instead of the original coarse partition

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

As it is, 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 $\{Failure,\ Minor\ Success\}$ 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 ruins the producer's incentives. If he observes $\{Failure,\ Minor\ Success\}$, he is free to embezzle without fear of the oil-boiling output of +400. He would still refrain from embezzling if he observed $\{Major\ Success\}$, but no contract that does not impose risk on a nonembezzling producer can stop him from embezzling if he observes $\{Failure,\ Minor\ Success\}$. 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.