Chapter 10

mechanism design and postcontractual hidden knowledge

10.1 Mechanisms, Unravelling, Cross Checking, and the Revelation Principle

This chapter looks at mechanism design. A mechanism is a set of rules that one player constructs and another freely accepts in order to convey information from the second player to the first. The mechanism contains an information report by the second player and a mapping from each possible report to some action by the first.

Adverse selection models can be viewed as problems of mechanism design. Insurance Game III was about an insurance company which wanted to know whether a customer was safe or not. In equilibrium it offers two contracts, an expensive full-insurance contract preferred by the safe customers and a cheap partial-insurance contract preferred by the unsafe. In the language of mechanism design, the insurance company sets up a game in which a customer reports his type as Safe or Unsafe, whichever he prefers to report, and the company then assigns him either partial or full insurance as a consequence. The contract offers are a mechanism for getting the agents to truthfully report their types.

Mechanism design goes beyond simple adverse selection. It can be useful even when players begin a game with symmetric information or when both players have hidden information that they would like to exchange.

Section 10.1 introduces moral hazard with hidden knowledge using Production Game VIII. It shows how to construct an optimal mechanism, how cross checking can attain the first-best when information is observable by the players but unverifiable by courts, and how "unravelling" can reveal private information if lying can be prevented but silence cannot. Section 10.2 designs a mechanism for a product quality game invented by Roger Myerson. Section 10.3 uses diagrams to apply the model to sales quotas. Section 10.4 introduces a mechanism that is multi-player and "dominant-strategy," the Groves Mechanism, for use when the problem is to elicit truthful reports from not one but N agents who need to decide whether to invest in a public good. Section 10.5 applies the principles of mechanism design to price discrimination. Section 10.6 lays out a more complicated model, of rate-of-return regulation by a government that constructs a mechanism to induce a regulated company to reveal how easily it could reduce its costs.

Postcontractual Hidden Knowledge

Information is complete in moral hazard games, but in **moral hazard with hidden knowledge**, also called **postcontractual adverse selection**, the agent, but not the principal, observes a move of Nature after the game begins, but before he takes his action. Information is symmetric at the time of contracting – thus the "moral hazard" – but becomes asymmetric later – thus the "hidden knowledge." From the principal's point of view, agents are identical at the beginning of the game but develop private types midway through. His chief concern is to give them incentives to disclose their types later, which gives the games a flavor close to that of adverse selection. The agent might also have to exert effort in the game, but effort's contractibility is less important when the principal does not know which effort is appropriate because he is ignorant of the state of the world chosen by Nature. The main difference technically is that if information is symmetric at the start and only becomes asymmetric after a contract is signed, the participation constraint is based on the agent's expected payoffs across the different types of agent he might become. Thus, there is just one participation constraint even if there are eventually n possible types of agents in the model, rather than the n participation constraints that would be required in a standard adverse selection model.

What makes postcontractual hidden knowledge an ideal setting for the paradigm of mechanism design is that the problem is to set up a contract that (1) induces the agent to make a truthful report to the principal, and (2) is acceptable to both principal and agent. There is more hope for obtaining efficient outcomes than in adverse selection. The advantage is that information is symmetric at the time of contracting, so neither player can use private information to extract surplus from the other by choosing inefficient contract terms.

For a comparison between the two types of moral hazard, let us modify Production Game VII from chapter 9 and turn it into a slightly different game of hidden knowledge.

Production Game VIII: Mechanism Design

PLAYERS
The principal and the agent.

THE ORDER OF PLAY

1. The principal offers the agent a wage contract of the form $w(q,m)$, where q is output and m is a message to be sent by the agent.
2. The agent accepts or rejects the principal's offer.
3. Nature chooses the state of the world s, according to probability distribution $F(s)$, where the state s is *good* with probability 0.5 and *bad* with probability 0.5. The agent observes s, but the principal does not.
4. If the agent accepted, he exerts effort e unobserved by the principal, and sends message $m \in \{good, bad\}$ to him.
5. Output is $q(e,s)$, where $q(e,good) = 3e$ and $q(e,bad) = e$, and the wage is paid.

PAYOFFS
If the agent rejects the contract, $\pi_{agent} = \overline{U} = 0$ and $\pi_{principal} = 0$.
If the agent accepts the contract, $\pi_{agent} = U(e,w,s) = w - e^2$ and $\pi_{principal} = V(q-w) = q - w$.

This game is almost the same as Production Game VII. The big difference is that now the agent does not know his type at the point in time at which he must accept or reject the contract. A smaller difference is that we have added the message m which the agent sends to the principal. This message is cheap talk – it does not affect payoffs directly and there is no penalty for lying. It is useful as a modelling convenience, to indicate which output-wage combination the agent chooses.

The principal would like to know s so he can tell which effort level is appropriate. In an ideal world he would employ an honest agent who always chose $m = s$, but in noncooperative games we ordinarily assume that agents have no moral sense. Since the agent's words are worthless, the principal must try to design a contract that either provides incentive for truth-telling or takes lying into account. He **implements** a **mechanism** to extract the agent's information.

In Production Game VII, the adverse selection version of the game, the optimal contracts had to satisfy two participation constraints and two incentive compatibility constraints.

In Production Game VIII, the moral hazard with hidden information version, the optimal contract must satisfy just one participation constraint, with the same two incentive compatibility constraints.

The first-best is unchanged from Production Game VII. The optimal effort and output in the good state of the world are $e_g = 1.5$ and $q_g = 4.5$, and in the bad state they are $e_b = 0.5$ and $q_b = 0.5$. Also unchanged is that the principal must solve the problem:

$$\underset{q_g,q_b,w_g,w_b}{Maximize}\ [0.5(q_g - w_g) + 0.5(q_b - w_b)], \tag{10.1}$$

where the agent is paid under one of two forcing contracts, (q_g, w_g) if he reports $m = good$ and (q_b, w_b) if he reports $m = bad$, where producing the wrong output for a given contract results in boiling in oil.

The contracts must induce participation and self-selection. We can write the constraints in terms of the agent's payoff function from in effect choosing one of the two (q, w) contracts by his choice of the report *good* or *bad*. The effort he will choose under those contracts will be $e = q/3$ for the good state and $e = q$ for the bad state, the same as in Production Game VII.

The self-selection constraints are the same as in Production Game VII. In the good state, the agent must choose the good-state contract, so

$$\pi_{agent}(q_g, w_g|good) = w_g - \left(\frac{q_g}{3}\right)^2 \geq \pi_{agent}(q_b, w_b|good) = w_b - \left(\frac{q_b}{3}\right)^2 \tag{10.2}$$

and in the bad state he must choose the bad-state contract, so

$$\pi_{agent}(q_b, w_b|bad) = w_b - q_b^2 \geq \pi_{agent}(q_g, w_g|bad) = w_g - q_g^2. \tag{10.3}$$

The participation constraints of Production Game VII now merge, though, because at the time of contracting the agent does not know what the state will be. The single participation constraint is therefore

$$0.5\pi_{agent}(q_g,w_g|good)+0.5\pi_{agent}(q_b,w_b|bad)=0.5\left(w_g-\left(\frac{q_g}{3}\right)^2\right)+0.5(w_b-q_b^2)\geq 0. \tag{10.4}$$

This single participation constraint is binding, since the principal wants to pay the agent as little as possible. The good state's self-selection constraint will be binding. In the good state the agent will be tempted to take the easier contract appropriate for the bad state (due to the "single-crossing property" to be discussed in a later section) and so the principal has to increase the agent's payoff from the good-state contract to yield him at least as much as in the bad state. He does not want to increase the surplus any more than necessary, though, so the good state's self-selection constraint will be exactly satisfied. This gives us two equations,

$$\begin{aligned} &0.5\left(w_g - \left(\frac{q_g}{3}\right)^2\right) + 0.5(w_b - q_b^2) = 0, \\ &w_g - \left(\frac{q_g}{3}\right)^2 = w_b - \left(\frac{q_b}{3}\right)^2. \end{aligned} \tag{10.5}$$

Solving them out yields $w_b = (5/9)q_b^2$ and $w_g = (1/9)q_g^2 + (4/9)q_b^2$.

Returning to the principal's maximization problem in (10.1) and substituting for w_b and w_g, we can rewrite it as

$$\underset{q_g, q_b}{Maximize}\ \pi_{principal} = \left[0.5\left(q_g - \frac{q_g^2}{9} - \frac{4q_b^2}{9}\right) + 0.5\left(q_b - \frac{5q_b^2}{9}\right)\right] \tag{10.6}$$

with no constraints. The first-order conditions are

$$\frac{\partial \pi_{principal}}{\partial q_g} = 0.5\left(1 - \left[\frac{2}{9}\right] q_g\right) = 0, \tag{10.7}$$

so $q_g = 4.5$, and

$$\frac{\partial \pi_{principal}}{\partial q_b} = 0.5\left(-\frac{8q_b}{9}\right) + 0.5\left(1 - \frac{10q_b}{9}\right) = 0, \tag{10.8}$$

so $q_b = 9/18 = 0.5$. We can then find the wages that satisfy the constraints, which are $w_g \approx 2.36$ and $w_b \approx 0.14$.

As in Production Game VII, in the good state effort is at the first-best level while in the bad state it is less. Unlike in Production Game VII the agent does not earn informational rents, because at the time of contracting he has no private information. In Production Game VII the wages were $w_g' \approx 2.32$ and $w_b' \approx 0.07$. Both wages are higher in Production Game VIII, but so is the effort and output required of the agent in the bad state. The principal in Production Game VIII is less constrained, and thus able to (1) come closer to the first-best when the state is bad, and (2) reduce the rents to the agent. Those are general features of moral hazard with hidden knowledge.

Observable but Nonverifiable Information and the Maskin Matching Scheme

If the state or type is public information, then it is straightforward to obtain the first-best using forcing contracts. What if the state is observable by both principal and agent, but is not public information?

The problem is that there are really three players involved in the contracting situation: the principal who offers the contract, the agent who accepts it – and the court that enforces it. If the courts cannot observe the state, a contract conditioning the wage on the state is unenforceable, no better than having no contract at all. We say that the variable s is **nonverifiable** if contracts based on it cannot be enforced. Most simply, the variable would one whose value the court cannot measure accurately enough to be useful – a worker's intensity of effort, for example, as opposed to the number of hours he was on the job. Or, the variable could be one which the court could observe but which for some reason it will not allow to be used in a legal contract – the amount of cocaine in a package, for example.

It does seem, however, that even if the courts will not enforce a contract based on a variable, if both the principal and the agent observe it they should be able to come up with a more efficient contract than if just the agent observes it. And indeed, mutual observability can help. Maskin (1977) suggested **cross checking**, as idea which would take the following two-part form for Production Game VIII:

1 Principal and agent simultaneously send messages m_p and m_a to the court saying whether the state is good or bad. If $m_p \neq m_a$, then no contract is chosen and both players earn zero payoffs. If $m_p = m_a$, the court enforces part (2) of the scheme.
2 The agent is paid the wage $(w|q)$ with either the good-state forcing contract $(2.25|4.5)$ or the bad-state forcing contract $(0.25|0.5)$, depending on his report m_a, or is boiled in oil if the output is inappropriate to his report.

There exists an equilibrium in which both players are willing to send truthful messages, because a deviation would result in zero payoffs. The agent earns a payoff of zero anyway (the numbers in the forcing contract come from making his participation constraint binding), but that is due to the open-set problem and to our assumption that the principal has all of the bargaining power. The principal's payoff is positive and efforts are at the first-best level.

Usually this kind of scheme has multiple equilibria, however, perverse ones in which both players send false messages which match and inefficient actions result. Here, in a perverse equilibrium the principal and agent would always send the message $m_p = m_a = bad$. Even when the state was actually good, the payoffs would be $\pi_{principal}(good) = 0.5 - 0.25 > 0$ and $\pi_{agent}(good) = 0.25 - (0.5)^2 = 0$, so neither player would have incentive to deviate unilaterally and drive payoffs to zero.[1]

Perhaps a bigger problem than the multiplicity of equilibria is renegotiation due to players' inability to commit to the mechanism. Suppose the equilibrium says that both players will send truthful messages, but the agent deviates and reports $m_a = bad$ even though the state is good. The court will say that the contract is void. I told you that the payoffs would be zero is that case, but the court's decision is not really the end of the story. The agent could negotiate a new contract with the principal, something the principal would be willing to do rather than give up the gains from trade. In cross checking the punishment for deviation is costly to both players, and so they will agree to bypass the scheme rather than inflict the punishment.

[1] By the way, "$m_p = m_a = good$ always" could not happen in equilibrium because the agent would prefer the zero payoff from mismatching to being forced by the threat of boiling oil to try to attain q_g when the state is bad.

In this, cross checking is like the Holmstrom Teams contract, where if output was even a little too small, it was destroyed rather than divided among the team members. There, a solution was to have a third party who would receive the output if it was slightly too small, and so would refuse to renegotiate it. The analogy here would be to write a contract in which both principal and agent paid a third party if their announcements disagreed. This sounds possible, yet we rarely observe this in practice.

Or do we? In his book *Wise Guys* (p. 57), Nicholas Pileggi quotes low-level gangster Henry Hill as saying that criminals need the protection of mafia "wiseguys" because they can't go to the police when they have a dispute over illegal activities: "For instance, say I've got a fifty-thousand-dollar hijack load, and when I make my delivery, instead of getting paid, I get stuck up. What am I supposed to do? Go to the cops? Not likely. Shoot it out? I'm a hijacker, not a cowboy. No. The only way to guarantee that I'm not going ripped off by anybody is to be established with a member, like Paulie. Somebody who is a made man. A member of a crime family. A soldier. Then … that's the end of the ball game. Goodbye. They're dead … Of course, problems can arise when the guys sticking you up are associated with wiseguys too. Then there has to be a sit-down between your wiseguys and their wiseguys. What usually happens then is that the wiseguys divide whatever you stole for their own pocket, send you and the guy who robbed you home with nothing. And if you complain, you're dead." The low-level gangsters have a strong incentive to report the same story, or the higher-ups take away the property under dispute. This may sound familiar to parents too – "If we can't resolve this, the toy is going in the closet for a whole week." It is perhaps even the wisdom of Solomon – see Glazer & Ma (1989) or Baliga (2002) (which also has a nice exposition of mechanism design using the author's own work to illustrate).

Unravelling: Information Disclosure When Lying Is Prohibited

There is another special case in which hidden information can be forced into the open: when the agent is prohibited from lying and only has a choice between telling the truth or remaining silent.

In Production Game VIII, this set-up would give the agent two possible message sets. If the state were good, the agent's message would be taken from $m \in \{good, silent\}$. If the state were *bad*, the agent's message would be taken from $m \in \{bad, silent\}$.

The agent would have no reason to be silent if the true state were bad (which means low output would be excusable), so his message then would be *bad*. But then if the principal hears the message *silent* he knows the state must be good – *good* and *silent* both would occur only when the state was good. So the option to remain silent is worthless to the agent.

This can be generalized. Suppose Nature uses the uniform distribution to assign the variable s some value in the interval $[0, 10]$ and the agent's payoff is increasing in the principal's estimate of s. Usually we assume that the agent can lie freely, sending a message m taking any value in $[0, 10]$, but let us assume instead that he cannot lie but he can conceal information. Thus, if $s = 2$, he can send the uninformative message $m \geq 0$ (equivalent to no message), or the message $m \geq 1$, or $m = 2$, but not the lie that $m \geq 4.36$.

When $s = 2$ the agent might as well send a message that is the exact truth: "$m = 2$." Loosely speaking, if he were to choose the message "$m \geq 1$" instead, the principal's first thought might be to estimate s as the average value in the interval $[1, 10]$, which is 5.5. But the principal would realize that no agent with a value of s greater than 5.5 would want to send the message "$m \geq 1$" if 5.5 was the resulting deduction. This realization restricts the possible interval to [1, 5.5], which in turn has an average of 3.25. But then no agent with $s > 3.25$ would send the message "$m \geq 1$." The principal would continue this process of logical **unravelling** to conclude that $s = 1$. The message "$m \geq 0$" would be even worse, making the principal believe that $s = 0$. In this model, no news is bad news. The agent would therefore not send the message "$m \geq 1$" and he would be indifferent between "$m = 2$" and "$m \geq 2$" because the principal would make the same deduction from either message.

More precisely, and perhaps more simply, use the Nash equilibrium approach of looking for profitable deviations. The equilibrium is either fully separating or it has some pooling. If it is fully separating, the agent's type is revealed, so he might as well send $m = s$. If it has some pooling, there exists some type s_{best} which is the best in the pool that sends a given signal $\widetilde{m}$. The principal's estimate of s on observing $\widetilde{m}$ would be the average in the pool, which is less than s_{best}. Agent type s_{best} would therefore deviate to $m = s_{best}$ and reveal his type. Thus, somebody would deviate from any partially pooling equilibrium. The unique equilibrium must be fully separating.

Perfect unravelling is paradoxical, but that is because the assumptions behind the last paragraph's reasoning are rarely satisfied in the real world. In particular, either unpunishable lying or genuine ignorance allow information to be concealed. If the seller is free to lie without punishment then in the absence of other incentives he always pretends that his information is extremely favorable, so nothing he says conveys any information, good or bad. If he really is ignorant in some states of the world, then his silence could mean either that he has nothing to say or that he has nothing favorable to report. The unravelling argument fails because if he sends an uninformative message the buyers will attach some probability to "I don't know" instead of "unfavorable news." Problem 10.1 explores unravelling further. For a careful discussion of the idea, see Milgrom (1981b).

The Revelation Principle

A principal might choose to offer a contract that induces his agent to lie in equilibrium, since he can take lying into account when he designs the contract, but this complicates the analysis. Each state of the world has a single truth, but a continuum of lies. Generically speaking, almost everything is false. The following principle helps us simplify contract design. Recall that w is the wage, q is output, m is the message, and s is the agent type.

> **The Revelation Principle:** *For every contract $w(q, m)$ that leads to lying (i.e., to $m \neq s$), there is a contract $w^*(q, m)$ with the same payoff for every s but no incentive for the agent to lie.*

Many possible contracts make false messages profitable for the agent because when the state of the world is a he receives a reward of x_1 for the true report of a and $x_2 > x_1$ for the false report of b. A contract which gives the agent the same reward of x_2 regardless of whether he reports a or b would lead to exactly the same payoffs for each player while giving the agent no incentive to lie. The revelation principle notes that a truth-telling contract like this can always be found by imitating the relation between states of the world and payoffs in the equilibrium of a contract with lying. There are two levels of simplification in mechanism design problems. First, if there are n possible types of agent, we can restrict the agent's message to take only n values (or, similarly, if the type is a value t on an interval $[a, b]$, we can restrict messages to that interval). If we do so, it is called a **direct mechanism**; if we allow more possible messages than types, it is an **indirect mechanism**. Second, we can require the mechanism to be constructed to elicit truthful messages from the agent. The revelation principle says that this requirement does not change the payoffs a mechanism might achieve.

In Production Game VIII, the revelation principle is not very useful since there are only two possible states and the equilibrium is perfectly separating – agents get different rewards in each state. Where the principle has bite is when the equilibrium involves some pooling. Suppose we are trying to design a mechanism to make people with higher incomes pay higher taxes, but anyone who makes $70,000 a year can claim he makes $50,000 and we do not have the resources to catch him, or any other variables to use for leverage to end up with a fully separating equilibrium. We could design a mechanism in which higher reported incomes pay higher taxes, but reports of $50,000 would come both from people who truly have that income and people whose income is $70,000. The revelation principle says that we can rewrite the tax code to set the tax to be the same for taxpayers earning $70,000 and for those earning $50,000, and the same amount of taxes will be collected without anyone having incentive to lie. This would be better if we are concerned with the effect on the moral climate of cheating on income taxes. Similarly, applied to children's education, the principle says that the mother who agrees never to punish her daughter if she tells her all her escapades will never hear any untruths. That is fine for deterring lying, but not very useful for deterring other misbehavior.

Clearly, the principle's usefulness is not so much to improve outcomes as to simplify contracts. The principal (and the modeller) need only look at contracts which induce truth-telling instead of more complex schemes in which the principal knows lying is going on but adjusts rewards accordingly. Thus, the relevant strategy space is shrunk, and we can add a third constraint to the incentive compatibility and participation constraints to help calculate the equilibrium:

Truth-telling: *The equilibrium contract makes the agent willing to choose* $m = s$.

The revelation principle says that a truth-telling equilibrium exists, but not that it is unique. It may well happen that the equilibrium is a weak Nash equilibrium in which the optimal contract gives the agent no incentive to lie but also no incentive to tell the truth. This is similar to section 4.3's open-set problem; the optimal contract satisfies the agent's participation constraint but makes him indifferent between accepting and rejecting the contract. If agents derive the slightest utility from telling the truth, then truth-telling becomes a strong equilibrium, but if their utility from telling the truth is really significant, it should be made an explicit part of the model. If the utility of truth-telling is strong enough, in fact, agency problems and the costs associated with them disappear. This is one reason why morality is useful to business.

The revelation principle does depend heavily on an implicit assumption we have made: the principal cannot breach his contract. In the example of tax design, for example, if the government can freely modify the tax schedule after the agents report their incomes, the direct mechanism under which agents pay the same tax for income of either $50,000 or $70,000 will break down, because the rich agents know that if they report their incomes of $70,000 truthfully, the government will change the tax schedule ex post and make them pay more. Instead, the government will have to use an indirect mechanism in which taxes are always higher for higher incomes and those people with incomes of $70,000 untruthfully report $50,000. Thus, throughout this chapter we will be assuming that the principal can commit to his mechanism – can commit to not using all the information he receives from the agent.

The Sender–Receiver Game of Crawford and Sobel: Coarse Information Transmission

Even if the informed and uninformed players have different incentives, can lie, and can't commit to a mechanism, if their incentives are close enough, truthful if imperfect messages can be sent in equilibrium – unlike in the unravelling games, where incentives are entirely divergent. Let us call the informed player "the sender" and the uninformed player "the receiver." That is common in this kind of model, instead of calling the two players the agent and the principal. Let the sender's type, t, be uniformly distributed on $[0, 10]$. The sender sends a message, m, and the receiver chooses an action, a, in response, where a and m are also in $[0, 10]$. At the extremes of payoff function similarity, it is clear what happens. Suppose the sender wants a to be as close to t as possible. If the sender also wants a to be close to t, he will tell the truth: $m = t$. If he wants a to be as big as possible regardless of t, on the other hand, because his ideal action is $a = 10$, then he will lie and the receiver will ignore any message.

But what happens if, say, the sender's ideal action is $(t + 1)$, so he doesn't want a to be too large, but he does want a to be bigger than a fully-informed receiver would choose? Crawford & Sobel (1982) discovered the answer in a famous article titled "Strategic Information Transmission." We will see what happens in the game below, which I adapted from Gibbons (1992, section 4.3.A), a good place to go if you want to learn more.

The Crawford–Sobel Sender–Receiver Game

PLAYERS
The sender and the receiver.

THE ORDER OF PLAY
0 Nature chooses the sender's type to be $t \sim U[0, 10]$.
1 The sender chooses message $m \in [0, 10]$.
2 The receiver chooses action $a \in [0, 10]$.

PAYOFFS
The payoffs are quadratic loss functions in which each player has an ideal point and wants a to be close to that ideal point.

$$\begin{aligned} \pi_{sender} &= \alpha - (a - [t + 1])^2, \\ \pi_{receiver} &= \alpha - (a - t)^2. \end{aligned} \tag{10.9}$$

First, let's see why perfect truthtelling cannot happen in equilibrium. Suppose the receiver believed that the sender always sent $m = t$ and so chooses $a = m$. Would the sender indeed be willing to tell the truth?

He would not. The sender would not always report $m = 10$, because his ideal point is $a = t + 1$, rather than a being as big as possible. If, however, the sender thinks the receiver will believe him, he will deviate to reporting $m = t + 1$, always exaggerating his type slightly.

What if the receiver adapts to this, and chooses $a = m - 1$? Then the sender would change too, and send $m = t + 2$, exaggerating more. This unravelling away from the truth (as opposed to the unravelling *toward* the truth when outright lying is forbidden) continues until the only message the sender reports is $m = 10$, regardless of his type and the receiver ignores it. This explanation is heuristic, and does not prove that there is no fully separating equilibrium, in which each type of sender reports a different message, but Crawford & Sobel (1982) prove that this is the case.

Thus, one equilibrium is the pooling equilibrium in which the sender's message is ignored and the receiver chooses $a = Et = 5$. This equilibrium could take either of two forms:

Pooling Equilibrium 1

Sender: Send $m = 10$ regardless of t.

Receiver: Choose $a = 5$ regardless of m.

Out-of-equilibrium belief: If the sender sends $m < 10$, the receiver uses passive conjectures and still believes that $t \sim U[0, 10]$.

Pooling Equilibrium 2

Sender: Send m using a mixed-strategy distribution independent of t that has the support $[0, 10]$ with positive density everywhere.

Receiver: Choose $a = 5$ regardless of m.

Out-of-equilibrium belief: Unnecessary, since any message might be observed in equilibrium.

In each of these two equilibria, the sender's action conveys no information and is ignored by the receiver. The sender is happy about this if it happens that $t = 4$, and the receiver is if $t = 5$, but averaging over all possible t, both their payoffs are lower than if the sender could commit to truthtelling. There also, however, exists a partial pooling equilibrium in which the sender truthfully reports whether his type is in the low interval $[0, x]$ or the high interval $[x, 10]$, with $x = 3$.

Partial Pooling Equilibrium 3

Sender: Send $m = 0$ if $t \in [0, 3]$ or $m = 10$ if $t \in [3, 10]$.

Receiver: Choose $a = 1.5$ if $m < 3$ and $a = 6.5$ if $m \geq 3$.

Out-of-equilibrium belief: If m is something other than 0 or 10, then $t \sim U[0, 3]$ if $m \in [0, 3)$ and $t \sim U[3, 10]$ if $a \in [3, 10]$.

In effect, the Sender has reduced his message space to two messages, LOW (=0) and HIGH (=10), in Equilibrium 3. Rather than just testing that this is an equilibrium, let us derive it, to show why the equilibrium interval-splitting type is $x = 3$.

First, note that the receiver's optimal strategy in a a partially pooling equilibrium is to choose his action to equal the expected value of the type in the interval the sender has chosen. Thus, if $m = 0$, the receiver will choose $a = x/2$ and if $m = 10$ he will choose $a = (x + 10)/2$.

The receiver's equilibrium response determines the sender's payoffs from his two messages. The payoffs between which he chooses are:

$$\begin{aligned} \pi_{sender,\, m=0} &= \alpha - \left([t+1] - \frac{x}{2}\right)^2, \\ \pi_{sender,\, m=10} &= \alpha - \left(\frac{10+x}{2} - [t+1]\right)^2. \end{aligned} \tag{10.10}$$

There exists a value x such that if $t = x$, the sender is indifferent between $m = 0$ and $m = 10$, but if t is lower he prefers $m = 0$ and if t is higher he prefers $m = 10$. To find x, equate the two payoffs in expression (10.10) and simplify to obtain

$$[t+1] - \frac{x}{2} = \frac{10+x}{2} - [t+1]\,. \tag{10.11}$$

We set $t = x$ at the point of indifference, and solving for x yields $x = 3$.

Thus, the divergence in preferences of the sender and receiver coarsens the message space, in effect. The sender will not send a truthful precise message, but if expectations are right (so we have the partially pooling equilibrium) he will send a truthful coarse message. If the true value of t is small, the sender will report the fairly precise information that t lies in [0,3]. If t is larger, it is harder to induce a truthful report, since the sender has a tendency to exaggerate and report t larger than it is, but the message can at least rule out the interval [0,3).

If instead of wanting $(t + 1)$ to be the action, the preferences of sender and receiver diverged more – say, to $(t + 8)$ – then there would only be the uninformative pooling equilibrium. If they diverged less – say, to $(t + 0.1)$ – then there would exist other partially pooling equilibria that had more than just two effective messages and would distinguish between three or more intervals instead of between just two.

In the Sender–Receiver Game, the receiver cannot commit to the way he reacts to the message, so this is not a mechanism design problem. The sender is not punished for lying, so the unravelling argument for truthtelling does not apply. Nor do the players' payoffs depend directly on the message, which might permit the signalling we will study in chapter 11 to operate. Instead, this is a **cheap-talk game**, so called because of these absences: m does not affect the payoff directly, the players cannot commit to future actions, and lying brings no direct penalty. In chapter 3 I alluded to how cheap talk might help select among equilibria in the Battle of the Sexes. In particular, knowing what the other player was thinking of doing would help to avoid the mixed-strategy equilibrium, with its low payoff. Our sender and receiver are in a similar situation here: their interests are similar but not identical, and they could both benefit from some transfer of information. If expectations are appropriate, they do so, in the partially pooling equilibrium. If they do not expect the cheap talk to be informative, however, it will not be, and coordination will fail.

10.2 Myerson Mechanism Design

The classic example of mechanism design is Roger Myerson's, who uses a version of it in sections 6.4 and 10.3 of his 1991 book. A seller has 100 units of a good. If it is high quality, he values it at 40 dollars per unit; if it is low quality, at 20 dollars. The buyer, who cannot observe quality before purchase, values high quality at 50 dollars per unit and low quality at 30 dollars. For efficiency, all of the good should be transferred from the seller to the buyer. The only way to get the seller to truthfully reveal the quality of the good, however, is for the buyer to say that if the seller admits the quality is bad, he will buy more units than if the seller claims it is good. Let us see how this works out.

Depending on who offers the contract and when it is offered, various games result. We will look at one in which the seller makes the offer, and does so before he knows whether his quality is high or low.

The Myerson Trading Game: Postcontractual Hidden Knowledge

PLAYERS
A buyer and a seller.

THE ORDER OF PLAY

1. The seller offers the buyer a contract $\{q_h, p_h, q_l, p_l\}$ under which the seller will declare his quality m to be high or low, and the buyer will then buy q_l or q_h units of the 100 the seller has available, at price p_l or p_h. The contract is $\{w(m) = q(m)p(m), q(m)\}$. Zero is paid if the wrong output is delivered.
2. The buyer accepts or rejects the contract.
3. Nature chooses whether the type of the seller's good, s, is High quality (probability 0.2) or Low (probability 0.8), unobserved by the buyer.
4. If the contract was accepted by both sides, the seller declares his type to be L or H and sells at the appropriate quantity and price as stated in the contract.

PAYOFFS
If the buyer rejects the contract, $\pi_{buyer} = 0$, $\pi_{seller\ H} = 40 * 100$, and $\pi_{seller\ L} = 20 * 100$.
If the buyer accepts the contract and the seller declares a type that has price p and quantity q then

$$\pi_{buyer|L} = (30 - p)q \quad \text{and} \quad \pi_{buyer|H} = (50 - p)q \tag{10.12}$$

and

$$\pi_{seller\ H} = 40(100 - q) + pq \quad \text{and} \quad \pi_{seller\ L} = 20(100 - q) + pq. \tag{10.13}$$

because the seller has an opportunity cost (a personal value or production cost) of 40 per high-quality unit and 20 per low-quality unit.

The seller wants to design a contract subject to two sets of constraints. First, the buyer must accept the contract. Thus, the participation constraint is[2]

$$\begin{aligned}&(0.8)\pi_{buyer|seller\ L} + (0.2)\pi_{buyer|seller\ H} \geq 0,\\ &0.8[(30-p_l)q_l] + 0.2[(50-p_h)q_h] \geq 0.\end{aligned} \tag{10.14}$$

This constraint will be binding, since the seller has no reason to leave the buyer any surplus. Notice, however, that if it is binding and both q_l and q_h are positive, we can conclude that $p_l = 30$ and $p_h = 50$.

There might also be a participation constraint for the seller himself, because it might be that even when he designs the contract that maximizes his payoff, his payoff is no higher than when he refuses to offer a contract. He can always offer the acceptable (if vacuous) null contract ($q_l = 0, p_l = 0, q_h = 0, p_h = 0$), however, so we do not need to write out the seller's participation constraint separately.

Second, the seller must design a contract that will induce himself to tell the truth later once he discovers his type. This is, of course a bit unusual – the seller is like a principal designing a contract for himself as agent. That is why things are different in this chapter than in the chapters on moral hazard. What is happening is that the seller is trying to sell not just a good, but a contract, and so he must make the contract attractive to the buyer. Thus, he faces two incentive compatibility constraints: one for when he is low quality,

$$\begin{aligned}&\pi_{seller\ L}(q_l, p_l) \geq \pi_{seller\ L}(q_h, p_h),\\ &20(100-q_l) + p_l q_l \geq 20(100-q_h) + p_h q_h,\\ &20(100-q_l) + 30q_l \geq 20(100-q_h) + 50q_h,\\ &q_l \geq 3q_h.\end{aligned} \tag{10.15}$$

and one for when he has high quality,

$$\begin{aligned}&\pi_{seller\ H}(q_h, p_h) \geq \pi_{seller\ H}(q_l, p_l),\\ &40(100-q_h) + p_h q_h \geq 40(100-q_l) + p_l q_l,\\ &40(100-q_h) + 50q_h \geq 40(100-q_l) + 30q_l,\\ &q_h \geq -q_l.\end{aligned} \tag{10.16}$$

The low-quality incentive compatibility constraint, inequality (10.15), tells us that $q_l > q_h$. The price of 50 that comes from claiming to have high quality dominates the price of 30 that comes from claiming to have low quality. If the seller cannot sell as great a quantity when he claims the quality is high, though, he might admit to low quality (think of an extreme case such as $q_l = 100, q_h = 5$).

On the other hand, the high-quality incentive compatibility constraint is satisfied for all possible q_l and q_h, because the high-quality seller's opportunity cost of 40 is greater than the price of 30 he could get by pretending to have low quality.

[2] Another kind of participation constraint would apply if the buyer had the option to reject purchasing anything, after accepting the contract and hearing the seller's type announcement. That would not make a difference here.

Thus, we know from the low-quality incentive compatibility constraint that we need $q_l > q_h$, and in fact that $q_l = 3q_h$ at the optimum. The seller's payoff function is

$$\begin{aligned}\pi_s &= 0.8\pi_{seller\ L}(q_l, p_l) + 0.2\pi_{seller\ H}(q_h, p_h),\\ &= 0.8[(20)(100 - q_l) + p_l q_l] + 0.2[(40)(100 - q_h) + p_h q_h],\\ &= 0.8[(20)(100 - q_l) + 30q_l] + 0.2[(40)(100 - q_h) + 50q_h]. \qquad (10.17)\end{aligned}$$

This is increasing in both q_l and q_h, so the seller would like to choose them to be as big as possible, subject to the constraints that $q_l = 3q_h$, $q_l \leq 100$, and $q_h \leq 100$. Thus, q_l will be the maximum possible, the first-best level $q_l = 100$, and q_h will be $q_h = q_l/3 = 33(1/3)$.

The equilibrium follows the general pattern for these games, though it has a twist because the informed player (the seller) has all the bargaining power, so it is the uninformed player (the buyer) whose participation constraint is binding. The incentive compatibility constraint is binding for the type with the most temptation to lie, and not for the other type. Using the two binding constraints, we can solve out for the values of some of the choice variables in terms of other choice variables, and then we can maximize the payoff of the player making the offer (the seller) to solve for values of those remaining variables. That is a useful general method, even though different games will have their own special features.

The mechanism will not work if further offers can be made after the end of the game. The mechanism is not first-best efficient; if the seller is high-quality, then he only sells 33(1/3) units to the buyer instead of all 100, even though both realize that the buyer's value is 50 and the seller's is only 40. If they could agree to sell the remaining 66(2/3) units, then the mechanism would not be incentive compatible in the first place, though, because then the low-quality seller would pretend to be high-quality, first selling 33(1/3) units and then selling the rest. The importance of commitment is a general feature of mechanisms.

A technical observation is that although we specified the contract in terms of (p, q), a price per unit p and a quantity q at that price, we could have set it up instead as (w, q), a total price amount w for the quantity q. That would be more in the style of mechanism design, with its emphasis on setting up the outcome of every outcome as the total transfers paid from one player to another and how an allocation or other decision is made.

10.3 An Example of Postcontractual Hidden Knowledge: The Salesman Game

Suppose the manager of a company has told his salesman to investigate a potential customer, who is either a *Pushover* or a *Bonanza*. If he is a *Pushover*, the efficient sales effort is low and sales should be moderate. If he is a *Bonanza*, the effort and sales should be higher. This is similar to Production Game VIII, but not the same. In the Salesman Game, the principal can perfectly deduce effort, even out of equilibrium. Also, there will be pooling as well as separating equilibria, and for the analysis we will make use of diagrams.

The Salesman Game

PLAYERS
A manager and a salesman.

THE ORDER OF PLAY
1 The manager offers the salesman a contract of the form $[w(m), q(m)]$, where w is the wage, q is sales, and m is a message.
2 The salesman decides whether or not to accept the contract.
3 Nature chooses whether the customer type t is a *Bonanza* or a *Pushover* with probabilities 0.2 and 0.8. The salesman observes the type, but the manager does not.
4 If the salesman has accepted the contract, he chooses his effort e. His sales level is $q = e$, so his sales perfectly reveal his effort.
5 The salesman's wage is $w(m)$ if he chooses $e = q(m)$ and zero otherwise.

PAYOFFS
The manager is risk-neutral and the salesman is risk-averse. If the salesman rejects the contract, his payoff is $\overline{U} = 8$ and the manager's is zero. If he accepts the contract, then

$$\pi_{manager} = q - w$$

$$\pi_{salesman} = U(e, w, \theta), \text{ where } \frac{\partial U}{\partial e} < 0, \ \frac{\partial^2 U}{\partial e^2} < 0, \ \frac{\partial U}{\partial w} > 0, \ \frac{\partial^2 U}{\partial w^2} < 0.$$

Figure 10.1 shows the indifference curves of manager and salesman, labelled with numerical values for exposition. The manager's indifference curves are straight lines with slope 1 because he is acting on behalf of a risk-neutral company. If the wage and the quantity both rise by a dollar, profits are unchanged, and the profits do not depend directly on whether s takes the value *Pushover* or *Bonanza*.

The salesman's indifference curves also slope upwards, because he must receive a higher wage to compensate for the extra effort that makes q greater. They are convex because the marginal utility of dollars is decreasing and the marginal disutility of effort is increasing. As figure 10.1 shows, the salesman has two sets of indifference curves, solid for *Pushovers* and dashed for *Bonanzas*, since the effort that secures a given level of sales depends on the state.

Because of the participation constraint, the manager must provide the salesman with a contract giving him at least his reservation utility of 8, which is the same in both states. If the true state is that the customer is a *Bonanza*, the manager would like to offer a contract that leaves the salesman on the dashed indifference curve $\widetilde{U_S} = 8$, and the efficient outcome is (q_2, w_2), the point at which the salesman's indifference curve is tangent to one of the manager's indifference curves. At that point, if the salesman sells an extra dollar he requires an extra dollar of compensation.

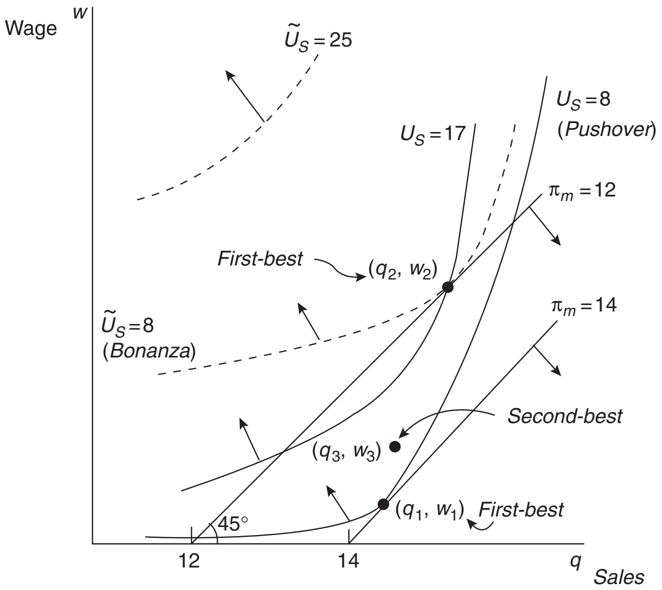

Figure 10.1 The Salesman Game with curves for pooling equilibria.

If it were common knowledge that the customer was a *Bonanza*, the principal could choose w_2 so that $U(q_2, w_2, Bonanza) = 8$ and offer the forcing contract

$$w = \begin{cases} 0 & \text{if } q < q_2, \\ w_2 & \text{if } q \geq q_2. \end{cases} \tag{10.18}$$

The salesman would accept the contract and choose $e = q_2$. But if the customer were actually a *Pushover*, the salesman would still choose $e = q_2$, an inefficient outcome that does not maximize profits. High sales would be inefficient because the salesman would be willing to give up more than a dollar of wages to escape having to make his last dollar of sales. Profits would not be maximized, because the salesman achieves a utility of 17, and he would have been willing to work for less.

The revelation principle says that in searching for the optimal contract we need only look at contracts that induce the agent to truthfully reveal what kind of customer he faces. If it required more effort to sell any quantity to the *Bonanza*, as shown in figure 10.1, the salesman would always want the manager to believe that he faced a *Bonanza*, so he could extract the extra pay necessary to achieve a utility of 8 selling to *Bonanzas*. The optimal truth-telling contract is the pooling contract that pays the intermediate wage of w_3 for the intermediate quantity of q_3, and zero for any other quantity, regardless of the message. The pooling contract is a second-best contract, a compromise between the optimum for *Pushovers* and the optimum for *Bonanzas*. The point (q_3, w_3) is closer to (q_1, w_1) than to (q_2, w_2), because the probability of a *Pushover* is higher and the contract must satisfy the participation constraint,

$$0.8U(q_3, w_3, Pushover) + 0.2U(q_3, w_3, Bonanza) \geq 8. \tag{10.19}$$

The nature of the equilibrium depends on the shapes of the indifference curves. If they are shaped as in figure 10.2, the equilibrium is separating, not pooling, and there does exist

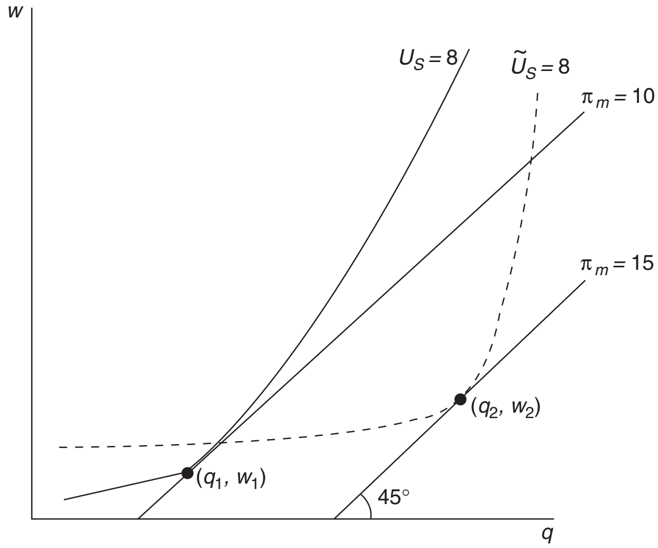

Figure 10.2 Indifference curves for a separating equilibrium.

a first-best, fully revealing contract.

$$\text{Separating contract} \begin{cases} \text{Agent announces } \textit{Pushover}\text{:} \quad w = \begin{cases} 0 & \text{if } q < q_1, \\ w_1 & \text{if } q \geq q_1, \end{cases} \\ \text{Agent announces } \textit{Bonanza}\text{:} \quad w = \begin{cases} 0 & \text{if } q < q_2, \\ w_2 & \text{if } q \geq q_2. \end{cases} \end{cases} \tag{10.20}$$

Again, we know from the revelation principle that we can narrow attention to contracts that induce the salesman to tell the truth. With figure 10.2's indifference curves, contract (10.20) induces the salesman to be truthful and the incentive compatibility constraint is satisfied. If the customer is a *Bonanza*, but the salesman claims to observe a *Pushover* and chooses q_1, his utility is less than 8 because the point (q_1, w_1) lies below the $\widetilde{U}_S = 8$ indifference curve. If the customer is a *Pushover* and the salesman claims to observe a *Bonanza*, then although (q_2, w_2) does yield the salesman a higher wage than (q_1, w_1), the extra income is not worth the extra effort, because (q_2, w_2) is far below the indifference curve $U_S = 8$.

Another way to look at a separating equilibrium is to think of it as a choice of contracts rather than as one contract with different wages for different outputs. The salesman agrees to work for the manager, and after he discovers the customer's type he chooses either the contract (q_1, w_1) or the contract (q_2, w_2), where each is a forcing contract that pays him 0 if after choosing the contract (q_i, w_i) he produces output of $q \neq q_i$. In this interpretation, the manager offers a **menu of contracts** and the salesman selects one of them after learning his type.

Sales contracts in the real world are often complicated because it is easy to measure sales and hard to measure efforts when workers are out in the field away from direct supervision. The Salesman Game is a real problem. Gonik (1978) describes hidden knowledge contracts used by IBM's subsidiary in Brazil. Salesmen were first assigned quotas. They then announced their own sales forecasts as a percentage of quota and chose from among a set of contracts, one for each possible forecast. I'll invent some numbers for illustration. If

Smith were assigned a quota of 400 and he announced 100 percent, he would get $w = 70$ if he sold 400 and $w = 80$ if he sold 450. If he had announced 120 percent, he would have gotten $w = 60$ for 400 and $w = 90$ for 450. The contract encourages extra effort when the extra effort is worth the extra sales. The idea here, as in the Salesman Game, is to reward salesmen not just for high effort, but for appropriate effort.

The Salesman Game illustrates a number of ideas. It can have either a pooling or a separating equilibrium, depending on the utility function of the salesman. The revelation principle can be applied to avoid having to consider contracts in which the manager must interpret the salesman's lies. It also shows how to use diagrams when the algebraic functions are intractable or unspecified, a problem that does not arise in most of the two-valued numerical examples in this book.

*10.4 The Groves Mechanism

Hidden knowledge is particularly important in public economics, the study of government spending and taxation. In Mirrlees (1971), a classic article in the optimal taxation literature, citizens differ in their income-producing ability and the government wishes to demand higher taxes from the more able citizens. Since the government cannot observe ability directly, this is a problem of hidden knowledge. An even purer hidden knowledge problem is choosing the level of public goods based on private preferences. The government must decide whether it is worthwhile to buy a public good based on the combined preferences of all the citizens, but it needs to discover those preferences. Unlike so far in this chapter, a group of agents will now be involved, not just one agent. Moreover, here the principal is an altruistic government that cares directly about the utility of the agents rather than a car buyer or an insurance seller who cares about the agents' utility only in order to satisfy self-selection and participation constraints.

Our example is adapted from Varian (1992, p. 426). The mayor of a town is considering installing a streetlight costing \$100. Each of the five houses near the light would be taxed exactly \$20, but the mayor will only install it if he decides that the sum of the residents' valuations for it is greater than the cost.

The mayor's problem is to discover the valuations. If he could observe them directly, he would simply builds the streetlight if $\sum_{i=1}^{5} v_i > 100$. Otherwise, he has a problem. If he simply asks the householders and tells them that each must pay a tax of \$20 if the streetlight is built, pro-light householder Smith might say that his valuation is \$5,000, and anti-light householder Brown might say that he likes darkness and would pay \$5,000 to *not* have a streetlight, but all the mayor could conclude would be that Smith's valuation exceeded \$20 and Brown's did not. Talk is cheap, and the dominant strategy would be to overreport or underreport.

The flawed mechanism just described can be written as

$$M_1: \quad \left(w(m_i) = 20, \textit{Build iff } \sum_{i=1}^{5} m_i \geq 100 \right); \tag{10.21}$$

that is, each resident pays \$20, and the light is installed if the sum of the valuations exceeds 100.

The Streetlight Game

PLAYERS
The mayor and five householders.

THE ORDER OF PLAY

0 Nature chooses the value v_i that householder i places on having a streetlight installed, using distribution $f_i(v_i)$. Only Householder i observes v_i.
1 The mayor announces a mechanism, M, which requires a householder who reports m to pay $w(m)$ if the streetlight is installed and installs the streetlight if $g(m_1, m_2, m_3, m_4, m_5) \geq 0$.
2 Householder i reports value m_i simultaneously with all other householders.
3 If $g(m_1, m_2, m_3, m_4, m_5) \geq 0$, the streetlight is built and householder i pays $w(m_i)$.

PAYOFFS
The mayor tries to maximize social welfare, including the welfare of taxpayers besides the 5 householders. His payoff is zero if the streetlight is not built. Otherwise, it is

$$\pi_{mayor} = \left(\sum_{i=1}^{5} v_i \right) - 100, \tag{10.22}$$

subject to the constraint that $\sum_{i=1}^{5} w(m_i) \geq 100$ so he can raise the taxes to pay for the light.
The payoff of householder i is zero if the streetlight is not built. Otherwise it is

$$\pi_i(m_1, m_2, m_3, m_4, m_5) = v_i - w(m_i). \tag{10.23}$$

An alternative is to make resident i pay the amount of his message, or pay zero if it is negative. This mechanism is

$$M_2\text{:} \quad \left(w(m_i) = Max\{m_i, 0\}, Build\ iff\ \sum_{j=1}^{5} m_j \geq 100 \right). \tag{10.24}$$

Mechanism M_2 has no dominant strategy. Householder i would announce $m_i = 0$ if he thought the project would go through without his support, based on his estimates of other people's values, but he would announce up to his valuation if necessary.

If all the householders knew each others' values perfectly, then there would be a continuum of Nash equilibria that attained the efficient result, much as in the Holmstrom Teams Game of chapter 8. If, for example, the values were known to be $(10, 30, 30, 30, 80)$, one equilibrium would be to report $(0, 25, 25, 25, 25)$. Since typically equilibria would be asymmetric, though, it is problematic how the equilibrium to be played out would become common knowledge, as well as how the householders know the v's in the first place. M_2 is a simple mechanism, however, and it already teaches a lesson: people are more likely to report their true political preferences if they must bear part of the costs themselves.

It turns out, however, that not only can a mechanism be found which makes truth-telling a Nash equilibrium, one can be found which makes truth-telling the best strategy for a player regardless of what the other players do – a **dominant-strategy mechanism**. Consider the mechanism M_3.

$$M_3: \quad \left(w(m_i) = 100 - \sum_{j \neq i} m_j, \textit{Build iff} \sum_{j=1}^{5} m_j \geq 100 \right). \tag{10.25}$$

Under mechanism M_3, player i's message does not affect his tax bill except by its effect on whether or not the streetlight is installed. If player i's valuation is v_i, his full payoff is $v_i - 100 + \sum_{j \neq i} m_j$ if $m_i + \sum_{j \neq i} m_j \geq 100$, and zero otherwise. It is not hard to see that he will be truthful in a Nash equilibrium in which the other players are truthful, but we can go further: truthfulness is weakly dominant. Moreover, the players are strictly better off telling the truth whenever lying would alter the mayor's decision.

Consider a numerical example. Suppose Smith's valuation is 40 and the sum of the valuations is 110, so the project is indeed efficient. If the other players report their truthful sum of 70, Smith's payoff from truthful reporting is his valuation of 40 minus his tax of 30. Reporting more would not change his payoff, while reporting less than 30 would reduce it to 0.

If we are wondering whether Smith's strategy is dominant, we must also consider his best response when the other players lie. If they underreported, announcing 50 instead of the truthful 70, then Smith could make up the difference by overreporting 60, but his payoff would be -10 $(=40 + 50 - 100)$ so he would do better to report the truthful 40, killing the project and leaving himself with a payoff of 0. If the other players overreported, announcing 80 instead of the truthful 70, then Smith would benefit if the project went through, and he should report at least 20. Whether he reports 20, 21, 40, or 400, the streetlight is built and he pays a tax of 20 under mechanism M_3, leaving him with payoff of $20 (=40 - 20)$. In particular, he is willing to report exactly 40, so it is a weakly best response to the other players' lies.

The problem with a dominant-strategy mechanisms like M_3 is that it is not budget balancing. This is not so bad if the budget had a surplus, as required in our game rules above, but it turns out to have a deficit except in special cases where it is perfectly balanced (e.g., $m_i = 20$ for all 5 householders). The first part of the tax $w(m)$ would collect 100 from each player, for 500, which leaves a surplus of 400 once we pay for the streetlight. The second part of $w(m)$, however, would subtract each player's value four times (1 for each other player), subtracting $4(\sum_{i=1}^{5} m_i)$. If the project goes through though, then $\sum_{i=1}^{5} m_i > 100$, so the budget would be left in deficit.

In fact, the total tax revenue could easily be negative too, because the "taxes" under M_3 are sometimes negative. If $v = 60$ for all five players, for example, then $m = 60$ and $w(m) = -140 (=100 - 4(60))$.

Vickrey (1961) first suggested the non-budget-balancing mechanism for revelation of preferences, but that was in the context of auctions, not public economics. We will see in chapter 13 that the second-price auction has the same distinctive feature that a player's own report affects the allocative decision (whether the light is built, who wins the auction) but not the amount he pays conditional on the decision being made (the tax, the price the auction winner pays). The idea was rediscovered later and became known as the Groves Mechanism from Groves (1973).

10.5 Price Discrimination

Now let's go to algebra to do a more conventional mechanism design problem, where the agents not only select an action that reveals their information, but must choose to play the game in the first place.

When a firm has market power – most simply when it is a monopolist – it would like to charge different prices to different consumers. To the consumer who would pay up to $45,000 for a car, the firm would like to charge $45,000; to the consumer who would pay up to $36,000, the profit-maximizing price is $36,000. But how does the car dealer know how much each consumer is willing to pay?

He does not, and that is what makes this a problem of mechanism design under adverse selection. The consumer who would be willing to pay $45,000 can hide under the guise of being a less intense consumer, and despite facing a monopolist he can end up retaining consumer surplus – an **informational rent**, a return to the consumer's private information about his own type.[3]

Pigou was a contemporary of Keynes at Cambridge who usefully divided price discrimination into three types in 1920 but named them so obscurely that I relegate his names to the endnotes and use better ones here:

1 **Interbuyer price discrimination:** This is when the seller can charge different prices to different buyers. Smith's price for a hamburger is $4 per burger, but Jones's is $6.
2 **Interquantity price discrimination or nonlinear pricing:** This is when the seller can charge different unit prices for different quantities. A consumer can buy a first sausage for $9, a second sausage for $4, and a third sausage for $3. Rather than paying the "linear" total price of $9 for one sausage, $18 for two, and $27 for three, he thus pays the nonlinear price of $9 for one sausage, $13 for two, and $16 for three, the concave price path shown in figure 10.3.
3 **Perfect price discrimination:** This combines interbuyer and interquantity price discrimination. When the seller does have perfect information and can charge each buyer that buyer's reservation price for each unit bought, Smith might end up paying $50 for his first hot dog and $20 for his second, while next to him Jones pays $4 for his first and $3 for his second.

To illustrate price discrimination as mechanism design we will use a modified version of an example in chapter 14 of Hal Varian's third edition (Varian [1992]).

[3] A standard opening ploy of car salesman is to ask. "So, how much are you able to spend on a car today?" My recommendation: don't tell him. This may sound obvious, but remember it the next time your department chairman asks you how high a salary it would take to keep you from leaving for another university.

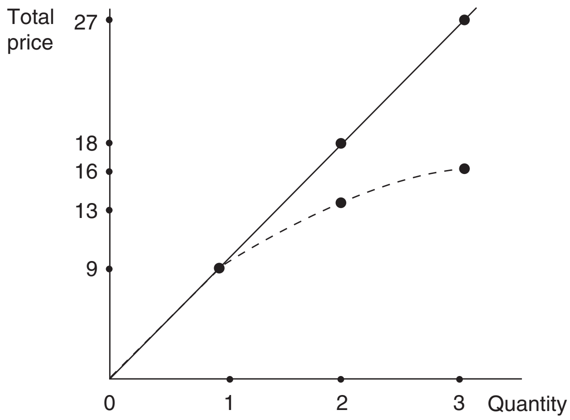

Figure 10.3 Linear and nonlinear pricing.

Varian's Nonlinear Pricing Game

PLAYERS
One seller and one buyer.

THE ORDER OF PLAY

0 Nature assigns the buyer a type, s. The buyer is "unenthusiastic" with utility function u or "valuing" with utility function v, with equal probability. The seller does not observe Nature's move, but the buyer does.

1 The seller offers mechanism $\{w_m, q_m\}$ under which the buyer can announce his type as m and buy amount q_m for lump sum w_m.

2 The buyer chooses a message m or rejects the mechanism entirely and does not buy at all.

PAYOFFS
The seller has a zero marginal cost, so his payoff is

$$w_u + w_v. \tag{10.26}$$

The buyers' payoffs are $\pi_u = u(q_u) - w_u$ and $\pi_v = v(q_v) - w_v$ if q is positive, and 0 if $q = 0$, with $u', v' > 0$ and $u'', v'' < 0$. The marginal willingness to pay is greater for the valuing buyer: for any q,

$$u'(q) < v'(q) \tag{10.27}$$

Condition (10.27) is an example of **the single-crossing property**, which we will discuss at the end of this section. Combined with the assumption that $v(0) = u(0) = 0$, it also implies that

$$u(q) < v(q) \tag{10.28}$$

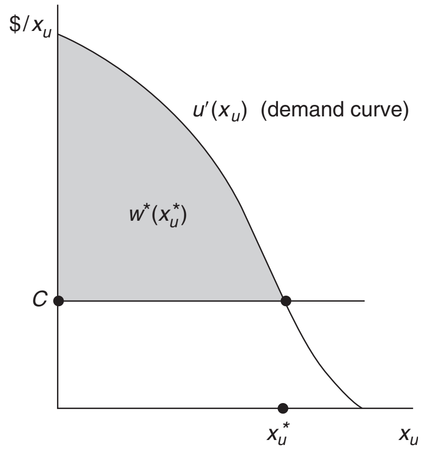

Figure 10.4 Perfect price discrimination.

for any value of q. Figure 10.7a, some pages below, shows an example of functions which satisfy it.

To ease into the difficult problem of solving for the equilibrium mechanism, let us start with two simpler versions of the game that limit it to (a) perfect price discrimination and (b) interbuyer discrimination.

Perfect Price Discrimination

The game would allow perfect price discrimination if the seller did know which buyer had which utility function. He can then just maximize profit subject to the participation constraints for the two buyers:

$$\underset{w_u, w_v, q_u, q_v}{Maximize} \; w_u + w_v, \tag{10.29}$$

subject to

$$\begin{aligned} &(a)\; u(q_u) - w_u \geq 0 \;\; \text{and} \\ &(b)\; v(q_v) - w_v \geq 0. \end{aligned} \tag{10.30}$$

The constraints will be satisfied as equalities, since the seller will charge all that the buyers will pay. Substituting for w_u and w_v into the maximand, the first order conditions become

$$\begin{aligned} &(a)\; u'(q_u^*) - c = 0 \quad \text{and} \\ &(b)\; v'(q_v^*) - c = 0. \end{aligned} \tag{10.31}$$

Thus, the seller will choose quantities so that each buyer's marginal utility equals the marginal cost of production, and will choose prices so that the entire consumer surpluses are eaten up: $w^*(q_u^*) = u(q_u^*)$ and $w^*(q_v^*) = v(q_v^*)$. Figure 10.4 shows this for the unenthusiastic buyer.

Interbuyer Price Discrimination

The interbuyer price discrimination problem arises when the seller knows which utility functions Smith and Jones have and can sell to them separately. If he can choose w_u and w_v as before and use forcing contracts, this is the same as the perfect price discrimination problem we just solved. If the seller must charge each buyer a single price per unit and let the buyer choose the quantity, however, the problem is quite different:

$$\underset{q_u,q_v,p_u,p_v}{Maximize} \; p_u q_u + p_v q_v, \tag{10.32}$$

subject to the participation constraints

$$u(q_u) - p_u q_u \geq 0 \quad \text{and} \quad v(q_v) - p_v q_v \geq 0 \tag{10.33}$$

and the incentive compatibility constraints

$$q_u = argmax[u(q_u) - p_u q_u] \quad \text{and} \quad q_v = argmax[v(q_v) - p_v q_v]. \tag{10.34}$$

This should remind you of moral hazard. It is very like the problem of a principal designing two incentive contracts for two agents to induce appropriate effort levels given their different disutilities of effort.
The agents will solve their quantity choice problems in (10.34), yielding

$$u'(q_u) - p_u = 0 \quad \text{and} \quad v'(q_v) - p_v = 0. \tag{10.35}$$

Thus, we can simplify the original problem in (10.32) to

$$\underset{q_u,q_v}{Maximize} \; u'(q_u)q_u + v'(q_v)q_v, \tag{10.36}$$

subject to the participation constraints

$$u(q_u) - u'(q_u)q_u \geq 0 \quad \text{and} \quad v(q_v) - v'(q_v)q_v \geq 0. \tag{10.37}$$

The participation constraints will not be binding. If they were, then $u(q)/q = u'(q)$, but since $u'' < 0$ there is diminishing utility of consumption and the average utility, $U(q)/q$, will be greater than the marginal utility, $u'(q)$. Thus we can solve problem (10.36) as if there were no constraints. The first-order conditions are

$$u''(q_u)q_u + u' = 0 \quad \text{and} \quad v''(q_v)q_v + v' = 0. \tag{10.38}$$

This is just the "marginal revenue equals marginal cost" condition that any monopolist uses, but one for each buyer instead of one for the entire market.

The assumption of constant marginal cost (equal to zero here) helps make this problem easier, because it makes it two independent problems, really. Choosing a contract for the valuing customer is completely separate from choosing one for the unenthusiastic customer. If the cost function were a more general convex $c(q_u + q_v)$, on the other hand, the two first-order conditions in (10.38) would have to be solved together, because each condition would depend on both q_u and q_v.

Back to Nonlinear Pricing

Neither the perfect price discrimination nor the interbuyer problems are mechanism design problems, since the seller is perfectly informed about the types of the buyers and has no need to worry about designing incentives to separate them. In the original game, however, separation is the seller's main concern. He must satisfy not just the participation constraints, but self-selection constraints. The seller's problem is

$$\underset{q_u, q_v, w_u, w_v}{Maximize} \; w_u + w_v, \tag{10.39}$$

subject to the participation constraints,

$$\begin{aligned} &(a) \quad u(q_u) - w_u \geq 0 \quad \text{and} \\ &(b) \quad v(q_v) - w_v \geq 0, \end{aligned} \tag{10.40}$$

and the self-selection constraints,

$$\begin{aligned} &(a) \quad u(q_u) - w_u \geq u(q_v) - w_v, \\ &(b) \quad v(q_v) - w_v \geq v(q_u) - w_u. \end{aligned} \tag{10.41}$$

Not all of these constraints will be binding. If neither type had a binding participation constraint, however, the principal would be losing a chance to increase his profits. In a mechanism design problem like this, what always happens is that the contracts are designed so that one type of agent is pushed down to his reservation utility.

Suppose the optimal contract is in fact separating, and also that both types accept a contract. At least one type will have a binding participation constraint. Since the valuing consumer gets more consumer surplus from a given w and q than an unenthusiastic consumer, it must be the unenthusiastic consumer who is driven down to zero surplus for (w_u, q_u). The valuing consumer would get positive surplus from accepting that same contract, so his participation constraint is not binding. To persuade the valuing consumer to accept (w_v, q_v) instead, the seller must give him that same positive surplus from it. The seller will not be any more generous than he has to, though, so the valuing consumer's self-selection constraint will be binding.

Rearranging our two binding constraints and setting them out as equalities yields:

$$w_u = u(q_u) \tag{10.42}$$

and

$$w_v = w_u - v(q_u) + v(q_v). \tag{10.43}$$

This allows us to reformulate the seller's problem from (10.39) as

$$\underset{q_u, q_v}{Maximize} \; u(q_u) + u(q_u) - v(q_u) + v(q_v), \tag{10.44}$$

which has the first-order conditions

$$\begin{aligned} &(a) \quad u'(q_u) + [u'(q_u) - v'(q_u)] = 0, \\ &(b) \quad v'(q_v) = 0. \end{aligned} \tag{10.45}$$

The first-order conditions in (10.45) could be solved for exact values of q_u and q_v if we chose particular functional forms, but they are illuminating even if we do not. Equation (10.45b) tells us that the valuing type of buyer buys a quantity such that his last unit's marginal utility exactly equals the marginal cost of production; his consumption is at the efficient level. The unenthusiastic type, however, buys less than his first-best amount, something we can deduce using the single-crossing property, assumption (10.27b), that $u'(q) < v'(q)$, which implies from (10.45a) that $u'(q_u) > 0$ and the unenthusiastic type has not bought enough to drive his marginal utility down to marginal cost. The intuition is that the seller must sell less than first-best optimal to the unenthusiastic type so as not to make that contract too attractive to the valuing type. On the other hand, making the valuing type's contract more valuable to him actually helps separation, so q_v is chosen to maximize social surplus.

The single-crossing property has another important implication. Substituting from first-order condition (10.45b) into first-order condition (10.45a) yields

$$[u'(q_u) - v'(q_v)] + [u'(q_u) - v'(q_u)] = 0. \tag{10.46}$$

The second term in square brackets is negative by the single-crossing property. Thus, the first term must be positive. But since the single-crossing property tells us that $[u'(q_u) - v'(q_u)] < 0$, it must be true, since $v'' < 0$, that if $q_u \geq q_v$ then $[u'(q_u) - v'(q_v)] < 0$ – that is, that the first term is negative. We cannot have that without contradiction, so it must be that $q_u < q_v$. The unenthusiastic buyer buys strictly less than the valuing buyer. This accords with our intuition, and also lets us know that the equilibrium is separating, not pooling (though we still have not proven that the equilibrium involves both players buying a positive amount, something hard to prove elegantly since one player buying zero would be a corner solution to our maximization problem).

A Graphical Approach to the Same Problem

Under perfect price discrimination, the seller would charge $w_u = A + B$ and $w_v = A + B + J + K + L$ to the two buyers for quantities q_u^* and q_v^*, as shown in figure 10.5a. An attempt to charge $w_u^* = A + B$ and $w_v^* = A + B + J + K + L$, however, would simply lead to both buyers choosing to buy q_u^*, which would yield the valuing buyer a payoff of $J + K$ rather than the zero he would get as a payoff from buying q_v^*. The seller's payoff from this pooling equilibrium (which is the best pooling contract possible for him, since it drives the unenthusiastic type to a payoff of zero) is $2(A + B)$.

The seller could separate the two buyers by charging $w_u^* = A+B$ for q_u^* and $w_v^* = A+B+L$ for q_v^*, since the unenthusiastic buyer would have no reason to switch to the greater quantity, and that would increase his profits over pooling by amount L.

Figure 10.5b shows, however, that the seller would do better to slightly reduce the quantity sold to the unenthusiastic buyer, to below q_u^*, and reduce the price to him by the amount of the dark shading. He could then sell q_v^* to the valuing buyer and raise the price to him by the light shaded area. The valuing buyer will not be tempted to buy the smaller quantity at the lower price, and the seller will have gained profit by, loosely speaking, increasing the size of the L triangle.

Continuing this process until profit is maximized is what we did earlier using algebra. Our profit-maximizing mechanism is shown in figure 10.5a as $w_u' = A$ for q_u' and $w_v^* = A + B + K + L$ for q_v^*. The unenthusiastic buyer is left with a binding participation constraint and inefficiently low consumption, because $w_u' = A = u(q_u')$. The valuing

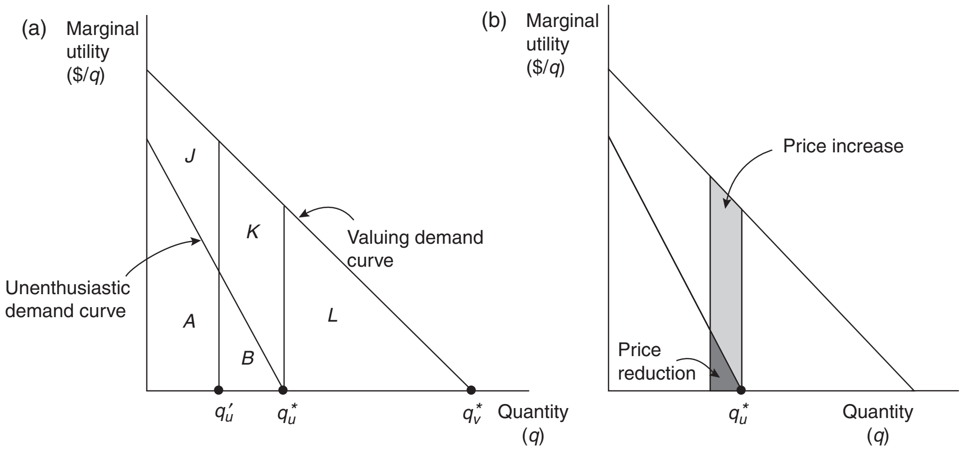

Figure 10.5 The Varian Nonlinear Pricing Game.

buyer has a nonbinding participation constraint, because $w_v^* = A + B + K + L < v(q_v^*) = A+B+J+K+L$; he is left with a surplus of J. Moreover, he consumes the efficient amount for him, which is q_v^*. He also has a binding self-selection constraint, because he is exactly indifferent between buying q_u' and q_v^*. His choice is between a payoff of $\pi_v(U) = (A+J)-A$ and $\pi_v(V) = (A + B + J + K + L) - (A + B + K + L)$. Thus, the diagram replicates the algebraic conclusions.

The Single-crossing Property

Condition (10.27) is an example of **the single-crossing property**, since it implies that the indifference curves of the two agents cross at most one time. Combined with the assumption that $v(0) = u(0) = 0$, it implies that $u(q) < v(q)$ for any value of q, as stated in inequality (10.28) earlier. Thus, in Varian's Nonlinear Pricing Game it is unambiguous that the valuing buyer has stronger demand than the unenthusiastic buyer.

When we say that Buyer V's demand is stronger than Buyer U's, however, there are two things we might mean:

1 Buyer V's *average demand* is stronger: $v(q)/q > u(q)/q$. Buyer V would pay more for quantity q than Buyer U would.
2 Buyer V's *marginal demand* is stronger: $v'(q) > u'(q)$. Buyer V would pay more for an additional unit than Buyer U would.

Definitions (1) and (2) are not equivalent. In figure 10.6a, Buyer U is willing to pay 5 per unit up to $q = 4$, but only 1 per unit thereafter. Buyer V is willing to pay only 2 per unit up to $q = 10$, and 1 per unit thereafter. As a result $u(q) > v(q)$, and U has the stronger demand by definition (1). But for $q \in [4, 10]$, Buyer V is willing to pay 2 per new unit while Buyer U is only willing to pay 1, so in that interval Buyber V has the stronger demand by definition (2).

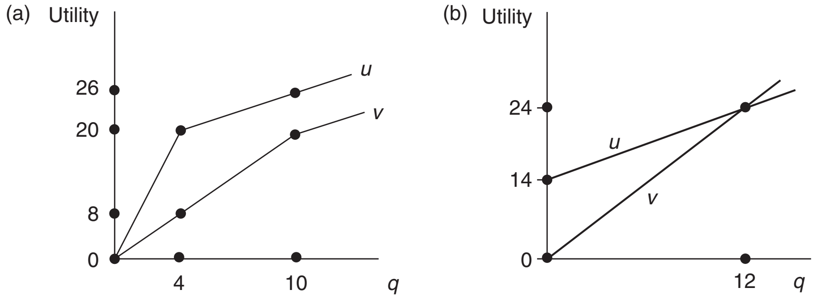

Figure 10.6 Marginal versus average demand.

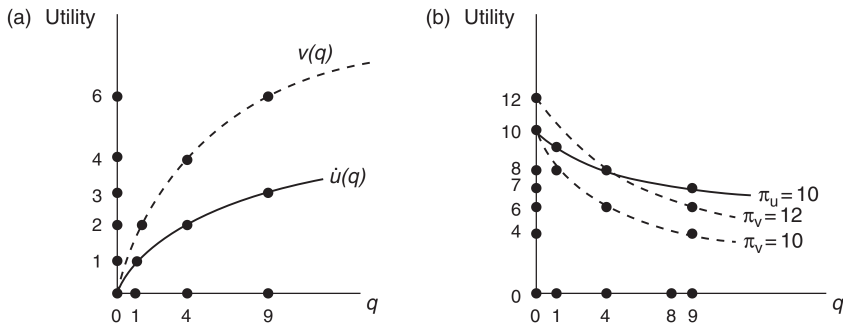

Figure 10.7 Two depictions of the single-crossing property.

It is the marginal demand that is more important to economic behavior and which forms the basis for the single-crossing property. The utility functions in figure 10.6b satisfy the single-crossing property, even though it is not true that $u(q) < v(q)$ (because they don't satisfy $u(0) = v(0) = 0$). Buyer V is always willing to pay 2 per unit, but Buyer U is only willing to pay 1 per unit. To be sure, $u(q) > v(q)$ for $u < 12$, but that is only because Buyer U starts with $u(0) = 14$, a "fixed benefit" which is irrelevant to his behavior. Buyer U may be the happier person at a consumption level of zero, but since that happiness is unaffected by his material circumstances, it is not going to affect any economic predictions we might want to make. Thus, definition (2) is best: when we say strong demand we should mean greater marginal demand. Indeed, that is what having a higher demand curve means in our usual diagrams, since the Marshallian demand curve is a marginal curve, showing not the total amount a consumer spends on quantity Q but the amount extra he willing to pay to buy a little more.

Figure 10.7a depicts functions which satisfy the assumptions of Varian's Nonlinear Pricing Game: $u = \sqrt{q}$ and $v = 2\sqrt{q}$. The two curves satisfy the single-crossing property, condition (10.27), because $v'(q) > u'(q)$ for all q and $u(0) = 0$ and $v(0) = 0$.

Another way to think about the single-crossing property, a way that makes the "single crossing" clearly visible, is using indifference curves. Since utility is a more theoretical concept than the idea of indifference between two consumption bundles, it is more reliable not to use graphs in utility space. The two goods in Varian's Nonlinear Pricing Game are (1) the commodity being, sold and (2) money, which enters linearly in the form of $-w$, the amount paid for the commodity. Another way to write the payoff functions would have been as $\pi_u(q, money) = money + u(q)$, where $money = wealth - w(q)$. Figure 10.7b shows how the buyers trade off money and the commodity. One comparison is between the curves for which $\pi = 10$, which both pass through the point (0, 10) in $(q, money)$ space. The $\pi_u = 10$ indifference curve then descends more slowly than the $\pi_v = 10$ curve because the commodity is not so valued by Buyer U. Another comparison is between the two curves which contain the point (4, 8), which are $\pi_u = 10$ and $\pi_v = 12$. These two curves also cross only once, at that point. In fact, if you pick any one indifference curve for Buyer U and any one for Buyer V, those curves will cross either not at all, or once.

It is often natural to assume that the single-crossing property holds, and it is a useful sufficient condition for separation to be possible, but it is not a necessary condition. If there are just two types of agents, as in many of our models, what matters is that the incentive compatibility constraints hold at the outputs that the principal specifies in the mechanism, not at all outputs. If the V player actually has a smaller marginal than the U player over a small range of consumption near zero, that will not hurt separation if the two contractual outputs are much larger.

*10.6 Rate-of-return Regulation and Government Procurement

The central idea in both government procurement and regulation of natural monopolies is that the government is trying to induce a private firm to efficiently provide a good to the public while covering the cost of production. If information is symmetric, this is an easy problem; the government simply pays the firm the cost of producing the good efficiently, whether the good be a missile or electricity. Usually, however, the firm has better information about costs and demand than the government does.

The variety of ways the firm might have better information and the government might extract it has given rise to a large literature in which moral hazard with hidden actions, moral hazard with hidden knowledge, adverse selection, and signalling all put in appearances. Suppose the government wants a firm to provide cable television service to a city. The firm knows more about its costs before agreeing to accept the franchise (adverse selection), discovers more after accepting it and beginning operations (moral hazard with hidden knowledge), and exerts greater or smaller effort to keep costs low (moral hazard with hidden actions). The government's problem is to acquire cable service at the lowest cost. It wants to be generous enough to induce the firm to accept the franchise in the first place but no more generous than necessary. It cannot simply agree to cover the firm's costs, because the firm would always claim high costs and exert low effort. Instead, the government might auction off the right to provide the service, might allow the firm a maximum price (a **price cap**), or might agree to compensate the firm to varying degrees for different levels of cost (**rate-of-return regulation**).

The problems of regulatory franchises and government procurement are the same in many ways. If the government wants to purchase a missile, it also has the problem of how much to offer the firm. Roughly speaking, the equivalent of a price cap is a flat price, and the equivalent of rate-of-return regulation is a cost-plus contract, although the details differ in interesting ways. (A price cap allows downwards flexibility in prices, and rate-of-return regulation allows an expected but not guaranteed profit, for example.)

Procurement I: Full Information[4]

PLAYERS
The government and the firm.

THE ORDER OF PLAY

0 Nature assigns the firm expensive problems with the project, which add costs of x, with probability θ. A firm is thus "normal," with type N and $s = 0$, or "expensive," with type X and $s = x$. The government and the firm both observe the type.

1 The government offers a contract $\{w(m) = c(m) + p(m), c(m)\}$ which pays the firm its observed cost c and a profit p if it announces its type to be m and incurs cost $c(m)$, and pays the firm zero otherwise.

2 The firm accepts or rejects the contract.

3 If the firm accepts, it chooses effort level e, unobserved by the government.

4 The firm finishes the missile at a cost of $c = \overline{c} + s - e$, which is observed by the government, plus an additional unobserved cost[5] of $f(e - \overline{c})$. The government reimburses $c(m)$ and pays $p(m)$.

PAYOFFS
Both firm and government are risk-neutral and both receive payoffs of zero if the firm rejects the contract. If the firm accepts, its payoff is

$$\pi_{firm} = p - f(e - \overline{c}) \qquad (10.47)$$

where $f(e - \overline{c})$, the cost of effort, is increasing and convex, so $f' > 0$ and $f'' > 0$. Assume for technical convenience that f is increasingly convex, so $f''' > 0$.[6] The government's payoff is

$$\pi_{government} = B - (1 + t)c - tp - f, \qquad (10.48)$$

where B is the benefit of the missile and t is the deadweight loss from the taxation needed for government spending. This is substantial. Hausman & Poterba (1987) estimate the loss to be around \$0.30 for each \$1 of tax revenue raised at the margin for the United States.

[4] I have changed the notation from the third edition of this book. The expensiveness variable x replaces the ability variable a; p replaces s; the type L firm becomes an expensive firm.

[5] The reader may ask why this disutility is specified as $f(e - \overline{c})$ rather than just $f(e)$. The reason is that we will later find an equilibrium cost level of $(\overline{c} - e^*)$, which would be negative if $c_0 = 0$.

[6] The argument of f is normalized to be $(\overline{c} - e)$ rather than just e to avoid clutter in the algebra later. The assumption that $f''' > 0$ allows the use of first-order conditions by making concave the maximand in (10.59), which is a difference of two concave functions. It will also make deterministic contracts superior to stochastic ones. See Laffont & Tirole (1993, p. 58).

Many of these situations are problems of moral hazard with hidden knowledge, because one player is trying to design a contract that the other will accept that will then induce him to use his private information properly.

Although the literature on mechanism design can be traced back to Mirrlees (1971) and in 1979 Loeb and Magat suggested using a Groves Mechanism to extract information from regulated firms, the extensive application to regulation began with Baron & Myerson's 1982 article, "Regulating a Monopolist with Unknown Costs." McAfee & McMillan (1988), and Laffont & Tirole (1993) provide 168-page, and 702-page treatments of the confusing array of possible models and policies in their books on government regulation. Here, we will look at a version of the model Laffont and Tirole use to introduce their book (pp. 55–62). This is a two-type model in which a special cost characteristic and the effort of a firm is its private information but its realized cost is public and nonstochastic. The model combines moral hazard and adverse selection, but it will behave more like an adverse selection model. The government will reimburse the firm's costs, but also fixes a price (which if negative becomes a tax) that depends on the level of the firm's costs. The questions the model hopes to answer are (1) whether effort will be too high or too low and (2) whether the price is positive and rises with costs.

The first version of the model will be one in which the government can observe the firm's type and so the first-best can be attained. It will be a benchmark for our later versions.

The model differs from most other principal–agent models in this book (though not from the Streetlight Game) because the principal cares about the welfare of the agent. If the government cared only about the value of the missile and the cost to taxpayers, its payoff would be $[B-(1+t)c-(1+t)p]$. Instead, the payoff function maximizes social welfare, the sum of the welfares of the taxpayers and the firm. The welfare of the firm is $(p-f)$, and summing the two welfares yields equation (10.48). Either kind of government payoff function may be realistic, depending on the political balance in the country being modelled, and the model will have similar properties whichever one is used. In the end, though, this model behaves in the same way as one with a selfish principal, because though the government does care about the welfare of the agent, the fact that taxation has deadweight loss means that the government will want to pay the firm as little as possible.

Assume for the moment that B is large enough that the government definitely wishes to build the missile (how large will become apparent later). Cost, not output, is the focus of this model. The optimal output is one missile regardless of agency problems, but the government wants to minimize the cost of producing the missile.

In Procurement I, whether the firm has expensive problems is observed by the government, which can therefore specify a contract conditioned on the type of the firm. The government pays p_N to a normal firm with the cost c_N, p_X to an expensive firm with the cost c_X, and $p=0$ to a firm that does not achieve its appropriate cost level. The government thus maximizes its payoff, equation (10.48), by choice of p_X, p_N, c_X, and c_N, subject to participation and incentive compatibility constraints.

The expensive firm exerts effort $e=\overline{c}+x-c_X$, achieves $c=c_X$, generating unobserved effort disutility $f(e-\overline{c})=f(x-c_X)$, so its participation constraint, that type X's payoff from reporting that it is type X, is

$$\begin{aligned} &\pi_X(X) \geq 0, \\ &p_X - f(x-c_X) \geq 0. \end{aligned} \tag{10.49}$$

Similarly, in equilibrium the normal firm exerts effort $e = \overline{c} - c_N$, so its participation constraint is

$$\begin{aligned} &\pi_N(N) \geq 0, \\ &p_N - f(-c_N) \geq 0. \end{aligned} \tag{10.50}$$

The incentive compatibility constraints are trivial here: the government can use a forcing contract that pays a firm zero if it generates the wrong cost for its type, since types are observable.

To make a firm's payoff zero and reduce the deadweight loss from taxation, the government will provide prices that do no more than equal the firm's disutility of effort. Since there is no uncertainty, we can invert the cost equation and write it as $e = \overline{c} + x - c$ or $e = \overline{c} - c$. The prices will be $p_X = f(e - \overline{c}) = f(x - c_X)$ and $p_N = f(e - \overline{c}) = f(-c_N)$.

Suppose the government knows the firm has expensive problems. Substituting the price p_X into the government's payoff function, equation (10.48), yields

$$\pi_{government} = B - (1+t)c_X - tf(x - c_X) - f(x - c_X). \tag{10.51}$$

Since $f'' > 0$, the government's payoff function is concave, and standard optimization techniques can be used. The first-order condition for c_X is

$$\frac{\partial \pi_{government}}{\partial c_X} = -(1+t) + (1+t)f'(x - c_X) = 0, \tag{10.52}$$

so

$$f'(x - c_X) = 1. \tag{10.53}$$

Equation (10.53) is the crucial efficiency condition for effort. Since the argument of f is $(e - \overline{c})$, whenever $f' = 1$ the effort level is efficient. At the optimal effort level, the marginal disutility of effort equals the marginal reduction in cost because of effort. This is the first-best efficient effort level, which we will denote by $e^* \equiv e:\{f'(e - \overline{c}) = 1\}$.

If we derived the first-order condition for the normal firm we would find $f'(-c_N) = 1$ in the same way, so $c_N = c_X - x$. Also, if the equilibrium disutility of effort is the same for both firms, then both must choose the same effort, e^*, though the normal firm can reach a lower cost target with that effort. The cost targets assigned to each firm are $c_X = \overline{c} + x - e^*$ and $c_N = \overline{c} - e^*$. Since both types must exert the same effort, e^*, to achieve their different targets, $p_X = f(e^* - \overline{c}) = p_N$. The two firms exert the same efficient effort level and are paid the same price to compensate for the disutility of effort. Let us call this price level p^*.

The assumption that B is sufficiently large can now be made more specific: it is that $B - (1+t)c_X - tf(e^* - \overline{c}) - f(e^* - \overline{c}) \geq 0$, which requires that $B - (1+t)(\overline{c} + x - e^*) - (1+t)p^* \geq 0$. If that were not true, then the government would not want to build the missile at all if the firm had an expensive cost function, as we will not treat of here.

Procurement II: Incomplete Information (Adverse Selection)

In the second variant of the game, the existence of expensive problems is not observed by the government, which must therefore provide incentives for the firm to volunteer its type if the normal firm is to produce at lower cost than the expensive firm.

If the government offered the two contracts of Procurement I, both types of firm would accept the expensive-cost contract, which has a price of p^* for a cost of $c = \overline{c} + x - e^*$, enough to compensate the expensive firm for its effort, and $p = 0$ for any other cost. That is the cheapest pooling contract, since any contract that paid less would violate the expensive-cost firm's participation constraint. It is inefficient, though, because the normal firm can reduce costs to $c = \overline{c}+x-e^*$ by exerting effort lower than e^*. The government would still be willing to build the missile, since the social cost of having the normal firm build the missile inefficiently is still lower than of having the expensive-cost firm build it efficiently. But it will turn out that separating contracts will yield higher welfare than the pooling contract.

First, let us establish that *some* pair of separating contracts is better than the pooling contract, and then we will find the *optimal separating contract*. A separating contract menu superior to the pooling contract would be a choice of (1) the old pooling contract ($p^*, c = \overline{c} + x - e^*$), and (2) a new contract that offers a slightly higher price p but requires reimbursable costs c to be slightly lower. By definition of e^* in first-order condition (10.53), $f'(e^* - \overline{c}) = 1$, so $f'(e' - \overline{c}) < 1$ for the effort the normal firm exerts in the old pooling contract. If the normal firm increased its effort from e' by some small amount Δe, costs would fall by $(1)\Delta e$ but the firm would only have to be paid $f'(e' - \overline{c})\Delta e$ more to compensate for its extra disutility. Thus, there is a new contract that would draw the normal firm away from the old pooling contract and be preferred by the government.

We have shown that there is a pair of separating contracts that the government likes better than the pooling contract, but not whether that pair is optimal. We will therefore proceed to find the optimal pair of contracts (c_N, p_N) (c_X, p_X) for firms that announce *Normal* or *Expensive* (with $p = 0$ for other cost levels). Adapting the government's payoff in (10.48) to the probability θ of a expensive firm and probability $1-\theta$ of a normal firm, the government's maximization problem under incomplete information is

$$\underset{c_N, c_X, p_N, p_X}{Maximize}\ \theta[B - (1+t)c_X - tp_X - f(x - c_X)] + [1-\theta][B - (1+t)c_N - tp_N - f(-c_N)]. \quad (10.54)$$

A separating contract must satisfy participation constraints and incentive compatibility constraints for each type of firm. The participation constraints are the same as in Procurement I: inequalities (10.49) and (10.50):

$$\pi_X(X) = p_X - f(x - c_X) \geq 0 \quad (10.49)$$

and

$$\pi_N(N) = p_N - f(-c_N) \geq 0. \quad (10.50)$$

The incentive compatibility constraint for the expensive firm is

$$\pi_X(X) = p_X - f(x - c_X) \geq \pi_X(N) = p_N - f(x - c_N), \quad (10.55)$$

and the incentive compatibility constraint for the normal firm is

$$\pi_N(N) = p_N - f(-c_N) \geq \pi_N(X) = p_X - f(-c_X). \tag{10.56}$$

Since the normal firm can achieve the same cost level as the expensive firm with less effort, inequality (10.56) tells us that if we are to have $c_N < c_X$, as is necessary for us to have a separating equilibrium, we need $p_N > p_X$. The second half of inequality (10.56) must be positive. If the expensive firm's participation constraint, inequality (10.49), is satisfied, then $p_X - f(-c_X) > 0$. This, in turn implies that (10.50) is a strong inequality; the normal firm's participation constraint is nonbinding.

The expensive firm's participation constraint (10.49), will be binding (and therefore satisfied as an equality), because the government wishes to keep the price p low to reduce the deadweight loss of extra taxation, the $-tp_X$ term in problem (10.54). The normal firm's incentive compatibility constraint must also be binding, because if the pair (c_N, p_N) were strictly more attractive for the normal firm, the government could reduce the price p_N and save on the $-tp_N$ term in problem (10.54). Constraint (10.56) is therefore satisfied as an equality. Knowing that constraints (10.49) and (10.56) are binding, we can write, from constraint (10.49),

$$p_X = f(x - c_X) \tag{10.57}$$

and, making use of both (10.49) and (10.56),

$$p_N = f(-c_N) + f(x - c_X) - f(-c_X). \tag{10.58}$$

Substituting for p_X and p_N from (10.57) and (10.58) into the maximization problem (10.54), reduces the problem to

$$\underset{c_N, c_X}{Maximize}\ \theta[B - (1+t)c_X - tf(x - c_X) - f(x - c_X)] \\ + [1-\theta][B - (1+t)c_N - tf(-c_N) - tf(x - c_X) + tf(-c_X) - f(-c_N)]. \tag{10.59}$$

The first-order condition with respect to c_N is

$$(1-\theta)[-(1+t) + tf'(-c_N) + f'(-c_N)] = 0, \tag{10.60}$$

which simplifies to

$$f'(-c_N) = 1. \tag{10.61}$$

Thus, as in Procurement I, $f_N'(e - \overline{c}) = 1$. The normal firm chooses the efficient effort level e^* in equilibrium, and c_N takes the same value as it did in Procurement I. Equation (10.58) can be rewritten as

$$p_N = p^* + f(x - c_X) - f(-c_X). \tag{10.62}$$

Because $f(x - c_X) > f(-c_X)$, equation (10.62) shows that $p_N > p^*$. Incomplete information increases the price for the normal firm, which earns more than its reservation utility in the game with incomplete information. Since the expensive firm will earn exactly zero, this means that the government is on average providing its supplier with an above-market rate of return, not because of corruption or political influence, but because that is the way to induce normal suppliers to reveal that they do not have expensive problems. This should be kept in mind as an alternative to chapter 5's product quality model and chapter 8's efficiency wage model for why above-average rates of return persist.

2 The first-order condition with respect to c_X is

$$\theta[-(1+t) + tf'(x - c_X) + f'(x - c_X)] + [1-\theta][tf'(x - c_X) - tf'(-c_X)] = 0. \tag{10.63}$$

This can be rewritten as

$$f'(x - c_X) = 1 - \left(\frac{1-\theta}{\theta(1+t)}\right)[tf'(x - c_X) - tf'(-c_X)]. \tag{10.64}$$

Since the right hand side of equation (10.64) is less than one, the expensive firm has a lower level of f' than the normal firm, and if f' is lower and $f'' > 0$, effort must be less than the optimum, e^*. Perhaps this explains the expression "good enough for government work" – though our model can apply to any organization that is trying to buy goods instead of making them internally. Since, however, the expensive firm's participation constraint (10.49), is satisfied as an equality, it must also be true that $p_X < p^*$. The expensive firm's price is lower than under full information, although since its effort is lower its payoff stays the same.

I have not yet said whether the expensive firm's incentive compatibility constraint was binding. It is not: the expensive firm is not near being tempted to pick the normal firm's contract. This is a bit subtle. Setting the left-hand side of the incentive compatibility constraint (10.55) equal to zero because the participation constraint is binding for the expensive firm, substituting in for p_N from equation (10.58) and rearranging yields

$$f(x - c_N) - f(-c_N) \geq f(x - c_X) - f(-c_X). \tag{10.65}$$

As illustrated in figure 10.8, inequality (10.65) is true as a strict inequality, because f is convex ($f'' > 0$) and so the increment in f's value starting frome the lower base $-c_X$ is

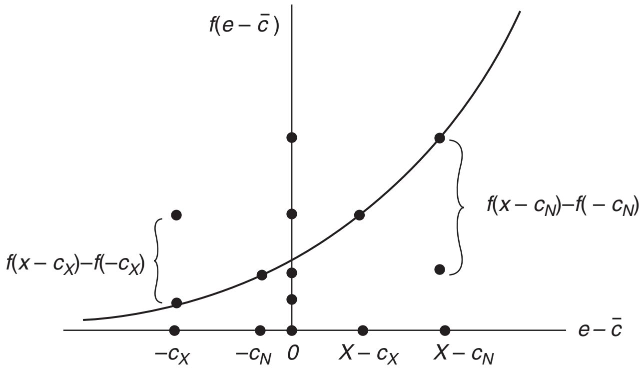

Figure 10.8 Why the expensive firm's incentive compatibility constraint is nonbinding.

smaller than starting from $-c_N$. Thus, the expensive firm's incentive compatibility constraint is nonbinding.

To summarize, the government's optimal contract will induce the normal firm to exert the first-best efficient effort level and achieve the first-best cost level, but will yield that firm a positive profit. The contract will induce the expensive firm to exert something less than the first-best effort level and result in a cost level higher than the first-best, but its profit will be zero.

There is a trade-off between the government's two objectives of inducing the correct amount of effort and minimizing the subsidy to the firm. Even under complete information, the government cannot provide a subsidy of zero, or the firms will refuse to build the missile. Under incomplete information, not only must the subsidies be positive but the normal firm earns **informational rents**; the government offers a contract that pays the normal firm more than under complete information to prevent it from mimicking an expensive firm and choosing an inefficiently low effort. The expensive firm, however, does choose an inefficiently low effort, because if it were assigned greater effort it would have to be paid a greater subsidy, which would tempt the normal firm to imitate it. In equilibrium, the government has compromised by having some probability of an inefficiently high subsidy ex post, and some probability of inefficiently low effort.

In the last version of the game, the firm's type is not known to either player until after the contract is agreed upon. The firm, however, learns its type before it must choose its effort level.

Procurement III: Moral Hazard with Hidden Information

THE ORDER OF PLAY

1. The government offers a contract $\{w(m) = c(m) + p(m), c(m)\}$ which pays the firm its observed cost c and a profit p if it announces its type to be m and incurs cost $c(m)$, and pays the firm zero otherwise.
2. The firm accepts or rejects the contract.
3. Nature assigns the firm expensive problems with the project, which add costs of x, with probability θ. A firm is thus "normal," with type N and $s = 0$, or "expensive," with type X and $s = x$. Only the firm observes its type.
4. If the firm accepted, it announces its type to be m and chooses effort level e, unobserved by the government.
5. If the firm accepted, it finishes the missile at a cost of $c = \overline{c} + x - e$ or $c = \overline{c} - e$ which is observed by the government, plus an additional cost $f(e - \overline{c})$ that the government does not observe. The government reimburses c and pays $p(c)$.

The Firm's payoff $(p - f)$, as in Procurement I.

The government's payoff function is exactly the same as in Procurement II, equation (10.54)

$$\underset{c_N, c_X, p_N, p_X}{Maximize}\ \theta\left[B - (1+t)c_X - tp_X - f(x - c_X)\right] + [1-\theta]\left[B - (1+t)c_N - tp_N - f(-c_N)\right]. \quad (10.54)$$

The incentive compatibility constraints are exactly the same as in Procurement II, equations (10.55) and (10.56). For the expensive firm the constraint is

$$\begin{aligned} &\pi_X(X) \geq \pi_X(N), \\ &p_X - f(x - c_X) \geq p_N - f(x - c_N), \end{aligned} \tag{10.55}$$

and for the normal firm it is

$$\begin{aligned} &\pi_N(N) \geq \pi_N(X), \\ &p_N - f(-c_N) \geq p_X - f(-c_X) \end{aligned} \tag{10.56}$$

The difference from Procurement II is that now there is just one participation constraint, not two, because the firm does not know its type at the time it agrees to the contract:

$$\theta[p_X - f(x - c_X)] + [1 - \theta][p_N - f(-c_N)] \geq 0. \tag{10.66}$$

Thus, the maximization problem is less constrained in Procurement III. The two participation constraints of Procurement II jointly imply constraint (10.66) is satisfied, but the reverse is not true. Now, the payoff of the expensive type can be negative, so long as the payoff of the normal type is positive enough to make the overall payoff non-negative.

The participation constraint, inequality (10.66), will be binding, because the government wants to keep the deadweight loss from taxation low, and it can now make the expensive firm's payoff negative enough to compensate for the normal firm's positive payoff. So long as the overall expected payoff is zero, the firm will still agree to the contract. There is no informational rent ex ante, unlike in Procurement II.

The normal firm's incentive compatibility constraint will still be binding. That firm does not want to admit that it can reduce costs easily, so it has a strong incentive to imitate the expensive firm, and must be bribed not to. The government will not want to pay a bigger bribe than necessary. If the pair (c_N, p_N) were strictly more attractive for the normal firm than (c_X, p_X), the government could reduce the price p_N. Constraint (10.56) is therefore satisfied as an equality.

Knowing that constraints (10.56) and (10.66) are binding, we can write from constraint (10.66),

$$p_X = f(x - c_X) - \frac{[1 - \theta][p_N - f(-c_N)]}{\theta}. \tag{10.67}$$

Substituting from (10.67) for p_X into (10.56), we get

$$p_N - f(-c_N) = f(x - c_X) - \frac{[1 - \theta][p_N - f(-c_N)]}{\theta} - f(-c_X). \tag{10.68}$$

This can be solved for p_N to yield

$$p_N = \theta[f(x - c_X) - f(-c_X)] + f(-c_N), \tag{10.69}$$

which when substituted into (10.67) yields

$$p_X = [1 - \theta][f(x - c_X) - f(-c_X)]. \tag{10.70}$$

Substituting for p_N and p_X from (10.69) and (10.70) into the maximization problem reduces the problem to

$$\underset{c_N, c_X}{Maximize}\ \theta\{B - (1+t)c_X - t[1-\theta][f(x-c_X) - f(-c_X)] - f(x-c_X)\}$$
$$+ [1-\theta][B - (1+t)c_N - t\{\theta[f(x-c_X) - f(-c_X)] + f(-c_N)\} - f(-c_N)]. \quad (10.71)$$

The first-order condition with respect to c_N is

$$(1-\theta)[-(1+t) + tf'(-c_N) + f'(-c_N)] = 0, \quad (10.72)$$

just as under adverse selection, which simplifies to

$$f'(-c_N) = 1. \quad (10.73)$$

Thus, the crucial efficiency condition (10.53) is satisfied: $f'(e - c_0) = 1$. The normal firm chooses the efficient effort level e^* in equilibrium, and c_N takes the same value as it did in Procurement I and II.

Now that we know that the effort level is e^*, we can say something about the normal firm's payoff. Since $p^* = f(-c_N)$, the optimal price equation (10.69) can be rewritten as

$$p_N = \theta[f(x-c_X) - f(-c_X)] + p^*, \quad (10.74)$$

Because $f(x-c_X) > f(-c_X)$, equation (10.74) shows that $p_N > p^*$, even though $e_N = e^*$. The normal firm earns a positive payoff even though information was symmetric at the start of the game. It earns an informational rent because partway through the game it learns its type and the government does not, and the government needs to pay something to induce it to admit to its low-cost type. The expensive firm must earn less than its reservation utility so that the overall participation constraint will be satisfied as an equality.

The government could have made the expensive firm's payoff negative in either of two ways: (1) a lower p_X than in Procurement II, or (2) a lower c_X than in Procurement II (and thus higher e_X). Option (2) is what we just derived. It is beter than (1) because it raises the expensive firm's effort to closer to the efficient level. This increases the amount of surplus, and the government will get the entire surplus since there is now no informational rent. Hence, the expensive firm's effort is higher in Procurement III than in Procurement II.

Now that c_X is higher for the same p_X, the expensive cost-price pair is less attractive than in Procurement II. As a result, $\pi_N(X)$ is lower, and the government can make the normal firm's cost-price pair less attractive too. This is we found that though c_N is unchanged from Procurement II, p_N has fallen.

To summarize, the government's optimal contract in Procurement III will induce the normal firm to exert the first-best efficient effort level and achieve the first-best cost level, but will yield that firm a positive payoff, though smaller than in Procurement II. The contract will induce the expensive firm to exert less than the first-best effort level, though more than in Procurement II, and result in a cost level higher than the first-best and a negative payoff. Overall, the firm's expected payoff will be zero.

This is what one might expect of moral hazard with hidden information as compared to adverse selection. Starting with symmetric information results in less gaming at the time of contract formation, so the principal's maximization problem has three constraints to satisfy instead of four, and this results in effort choices closer to the first-best. Effort choices still do not always reach the first-best, however, because information about player types does become asymmetric midway through the game, and the contract has to be designed to induce the informed player, the firm, to disclose its information.

A practical implication is that the parties ought to agree on a contract as early as possible in the procurement process, before one of them acquires an informational advantage. Of course, this does not apply if one of them *starts* with an informational advantage; then, delay before agreement might actually help, by giving the uninformed party time to learn the facts it needs.

A little reflection will provide a host of additional ways to alter the Procurement Game. What if the firm discovers its costs only after accepting the contract? What if two firms bid against each other for the contract? What if the firm can bribe the government? What if the firm and the government bargain over the gains from the project instead of the government being able to make a take-it-or-leave-it contract offer? What if the game is repeated, so the government can use the information it acquires in the second period? If it is repeated, can the government commit to long-term contracts? Can it commit not to renegotiate? See Laffont & Tirole (1993) if these questions interest you. If they merely confuse you, an aphorism by Doug Larson that McAfee (2002, p. 202) quotes is an apt summary of the Procurement Game: "Accomplishing the impossible means only that the boss will add it to your regular duties."

Notes

N10.1 Production Game VIII, cross checking, unravelling, and the Revelation Principle

- The revelation principle was developed in Myerson (1979), though the idea can be traced back to Gibbard (1973). It was named in Myerson (1981). Myerson's game theory book is, as one might expect, a good place to look for further details (Myerson [1991, pp. 258–63, 294–99]). See also the books by Fudenberg & Tirole (1991a) and Laffont & Tirole (1993), Baron's chapter in the 1989 *Handbook of Industrial Organization*, the 2002 and 2005 books on the principal–agent problem by Laffont & Martimort and Bolton & Dewatripont, and Stole's 2001 manuscript.
- Levmore (1982) discusses hidden knowledge problems in tort damages, corporate freezeouts, and property taxes in a law review article. A legal application of the idea of the Sender–Receiver Game, pointing out the usefulness of a mediator as a mechanism is Brown & Ayres (1994).
- Postcontractual adverse selection is common in public policy. Should the doctors who prescribe drugs also be allowed to sell them? The question trades off the likelihood of overprescription against the potentially lower cost and greater convenience of doctor-dispensed drugs. See "Doctors as Druggists: Good Rx for Consumers?" *The Wall Street Journal*, June 25, 1987, p. 24.
- A hidden knowledge game requires that the state of the world matter to one of the players' payoffs, but not necessarily in the same way as in Production Game VII. The Salesman Game of section 10.2 effectively uses the utility function $U(e, w, \theta)$ for the agent and $V(q - w)$ for the principal. The state of the world matters because the agent's disutility of effort varies across states. In other problems, his utility of money might vary across states.
- Eric Maskin came up with the idea of cross-checking in a 1977 MIT working paper, "Nash Implementation and Welfare Optimality" The working paper became known as a classic, but was

not published until 22 years later, in a 1999 issue of the *Review of Economic Studies* with several other articles on mechanism design.

N10.3 An example of postcontractual private knowledge: the Salesman Game

- Sometimes students know more about their class rankings than the professor does. One professor of labor economics used a mechanism of the following kind for grading class discussion. Each student *i* reports a number evaluating other students in the class. Student *i*'s grade is an increasing function of the evaluations given *i* by other students and of the correlation between *i*'s evaluations and the other students'. There are many Nash equilibria, but telling the truth is a focal point.
- In dynamic games of moral hazard with hidden knowledge the **ratchet effect** is important: the agent takes into account that his information-revealing choice of contract this period will affect the principal's offerings next period. A principal might allow high prices to a public utility in the first period to discover that its costs are lower than expected, but in the next period the prices would be reduced. The contract is ratcheted irreversibly to be more severe. Hence, the company might not choose a contract which reveals its costs in the first period. This is modelled in Freixas, Guesnerie, & Tirole (1985).

  Baron (1989) notes that the principal might purposely design the equilibrium to be pooling in the first period so self selection does not occur. Having learned nothing, he can offer a more effective separating contract in the second period.

N10.5 Price discrimination

- The names for price discrimination in part 2, chapter 17, section 5 of Pigou (1920) are: (1) first-degree (perfect price discrimination), (2) second-degree (interquantity price discrimination), and (3) third-degree (interbuyer price discrimination). These arbitrary names have plagued generations of students of industrial organization, in parallel with the appalling Type I and Type II errors of statistics (better named as False Negatives and False Positives). I invented the terms **interbuyer price discrimination** and **interquantity price discrimination** for this edition, with the excuse that I think their meaning will be clear to anyone who already knows the concepts under their Pigouvian names.

  McAfee (2002, p. 261) has a new taxonomy that may catch on: direct versus indirect price discrimination. Direct price discrimination is based on observing characteristics of the customer and charging him a price for a given unit that depends on those characteristics, an offer not open to everyone. Indirect price discrimination makes offers open to everyone, but offers which will separate the customers, whether by quantity desired, quality desired, attention paid to newspaper coupons, or other unobservable characteristics.
- A narrower category of nonlinear pricing is the **quantity discount**, in which the price per unit declines with the quantity bought. Sellers are often constrained to this, since if the price per unit rises with the quantity bought, some means must be used to prevent a canny consumer from buying two batches of small quantities instead of one batch of a large quantity.
- Wilson's 1993 book, *Nonlinear Pricing*, is a good reference on price discrimination.
- In Varian's Nonlinear Pricing Game the probabilities of types for each player are not independent, unlike in most games. This does not make the game more complicated, though. If the assumption were "Nature assigns each buyer a utility function u or v with independent probabilities of 0.5 for each type," then there would be not just two possible states of the world in this game – uv and vu for Smith and Jones's types – but four – uv, vu, uu, and vv. How would the equilibrium change?
- The careful reader will think, "How can we say that Buyer V always gets higher utility than Buyer U for given x? Utility cannot be compared across individuals, and we could rescale Buyer V's utility function to make him always have lower utility without altering the essentials of the utility

function." My reply is that more generally we could set up the utility functions as $v(x) + y$ and $u(x) + y$, with y denoting spending on all other goods (as Varian does in his book). Then to say that V always gets higher utility for a given x means that he always has a higher relative value than U does for good x relative to money. Rescaling to give V the utility function $.001v(x) + .001y$ would not alter that.

- The notation I used in Varian's Nonlinear Pricing Game is optimized for reading. If you wish to write this on the board or do the derivations for practice, use abbreviations like u_1 for $u_1(x_1)$, a for $v(x_1)$, and b for $v(x_2)$ to save writing. The tradeoff between brevity and transparency in notation is common, and must be made in light of whether you are writing on a blackboard or on a computer, for just yourself or for the generations.

N10.6 Rate-of-return regulation and government procurement

- I changed the notation from Laffont and Tirole and from my own previous edition. Rather than assign each type of firm a cost parameter β for a cost of $c = \beta - e$, I now assign each type of firm an ability parameter a, for a cost of $c = c_0 - a - e$. This will allow the desirable type of firm to be the one with the *High* value of the type parameter, as in most models.
- In practice, whether procurement is by the government or by other large organizations it must balance a multitude of concerns, not just incentives for the supplier. Three of the most important concerns are giving performance incentives to the government's own agents, (the ones who arrange the procurement), preventing corruption of those agents with kickbacks (a bribe awarded an agent in consideration of getting a contract), and transaction costs of various sorts. On this last, see Bajari & Tadelis (2001). Much insight can be had from auction theory as well, since auctions are frequently used for procurement. Paul Klemperer's 2004 *Auctions: Theory and Practice* (which is relatively nontechnical), and Paul Milgrom's 1999 *Auction Theory for Privatization* and 2004 *Putting Auction Theory to Work* are full of practical advice based on sound theory.

Problems

10.1: Unravelling (hard)

An elderly prospector owns a gold mine worth an amount θ drawn from the uniform distribution $U[0, 100]$ which nobody knows, including himself. He will certainly sell the mine, since he is too old to work it and it has no value to him if he does not sell it. The several prospective buyers are all risk-neutral. The prospector can, if he desires, dig deeper into the hill and collect a sample of gold ore that will reveal the value of θ. If he shows the ore to the buyers, however, he must show genuine ore, since an unwritten Law of the West says that fraud is punished by hanging offenders from joshua trees as food for buzzards.

(a) For how much can he sell the mine if he is clearly too feeble to have dug into the hill and examined the ore? What is the price in this situation if, in fact, the true value is $\theta = 70$?

(b) For how much can he sell the mine if he can dig the test tunnel at zero cost? Will he show the ore? What is the price in this situation if, in fact, the true value is $\theta = 70$?

(c) For how much can he sell the mine if, after digging the tunnel at zero cost and discovering θ, it costs him an additional 10 to verify the results for the buyers? What is his expected payoff?

(d) Suppose that with probability 0.5 digging the test tunnel costs 5 for the prospector, but with probability 0.5 it costs him 120. Keep in mind that the 0–100 value of the mine is net of the buyer's digging cost. Denote the equilibrium price that buyers will pay for the mine after the prospector approaches them without showing ore by P. What is the buyer's posterior belief about the probability it costs 120 to dig the tunnel, as a function of P? Denote this belief by $B(P)$. (Assume, as usual, that all these parameters are common knowledge, although only the prospector learns whether the cost is actually 0 or 120.)

(e) What is the prospector's expected payoff in the conditions of part (d) if (i) the tunnel costs him 120, or (ii) the tunnel costs him 5?

(f) What is the prospector's ex ante expected payoff in the conditions of part (d) – that is, what is his expected payoff viewed from before he knows the cost of digging a tunnel? How does that compare with his expected payoff when he can dig the tunnel at cost 5 with probability 1?

Table 10.1 The right to silence game payoffs

		Sally's Job		
		Job 1	*Job 2*	*Manager*
Sally knows	Task 1 is efficient (0.5)	2, 5	1, −2	3, 3
	Task 2 is efficient (0.5)	1, −2	2, 5	3, 3

Payoffs to: (Sally, Rayco).

10.2: Task assignment (medium)

Table 10.1 shows the payoffs in the following game. Sally has been hired by Rayco to do either Job 1, to do Job 2, or to be a Manager. Rayco believes that Tasks 1 and 2 have equal probabilities of being the efficient ones for Sally to perform. Sally knows which task is efficient, but what she would like best is a job as Manager that gives her the freedom to choose rather than have the job designed for the task. The CEO of Rayco asks Sally which task is efficient. She can either reply "task 1," "task 2," or be silent. Her statement, if she makes one, is an example of "cheap talk," because it has no direct effect on anybody's payoff. See Farrell & Rabin (1996).

(a) If Sally did not have the option of speaking, what would happen?

(b) There exist perfect Bayesian equilibria in which it does not matter how Sally replies. Find one of these in which Sally speaks at least some of the time, and explain why it is an equilibrium. You may assume that Sally is not morally or otherwise bound to speak the truth.

(c) There exists a perverse variety of equilibrium in which Sally always tells the truth and never is silent. Find an example of this equilibrium, and explain why neither player would have incentive to deviate to out-of-equilibrium behavior.

10.3: Agency law (easy)

Mr. Smith is thinking of buying a custom-designed machine from either Mr. Jones or Mr. Brown. This machine costs $5,000 to build, and it is useless to anyone but Smith. It is common knowledge that with 90 percent probability the machine will be worth $10,000 to Smith at the time of delivery, one year from today, and with 10 percent probability it will only be worth $2,000. Smith owns assets of $1,000. At the time of contracting, Jones and Brown believe there is a 20 percent chance that Smith is actually acting as an "undisclosed agent" for Anderson, who has assets of $50,000.

Find the price be under the following two legal regimes: (1) An undisclosed principal is not responsible for the debts of his agent; and (2) even an undisclosed principal is responsible for the debts of his agent. Also, explain (as part [3]) which rule a moral hazard model like this would tend to support.

10.4: Incentive compatibility and price discrimination (medium)

Two consumers have utility functions $u_1(x_1, y_1) = a_1 log(1 + x_1) + y_1 - 15$ and $u_2(x_2, y_2) = a_2 log(1 + x_2) + y_2 - 15$ where $2 < a_1 < a_2 < 12$. The price of the y-good is 1 and each consumer has an initial wealth of 15. A monopolist supplies the x-good. He has a constant marginal cost of 1.2 up to his capacity constraint of 10. He will offer at most two price-quantity packages, (r_1, x_1) and (r_2, x_2), where r_i is the total cost of purchasing x_i units. He cannot identify which consumer is which, but he can prevent resale.

(a) Write down the monopolist's profit maximization problem. You should have four constraints plus a capacity constraint.
(b) Which constraints will be binding at the optimal solution?
(c) Substitute the binding constraints into the objective function. What is the resulting expression? What are the first-order conditions for profit maximization? What are the profit-maximizing values of x_1 and x_2?

10.5: The Groves Mechanism (easy)

A new computer costing 10 million dollars would benefit existing Divisions 1, 2, and 3 of a company with 100 divisions. Each divisional manager knows the benefit to his division (variables v_i, $i = 1, \ldots, 3$), but nobody else does, including the company CEO. Managers maximize the welfare of their own divisions. What dominant strategy mechanism might the CEO use to induce the managers to tell the truth when they report their valuations? Explain why this mechanism will induce truthful reporting, and denote the reports by x_i, $i = 1, \ldots, 3$. (You may assume that any budget transfers to and from the divisions in this mechanism are permanent – that the divisions will not get anything back later if the CEO collects more payments than he gives, for example.)

10.6: The two-part tariff (easy) (Varian 14.10, modified)

One way to price discriminate is to charge a lump sum fee L to have the right to purchase a good, and then charge a per-unit charge p for consumption of the good after that. The standard example is an amusement park where the firm charges an entry fee and a charge for the rides inside the park. Such a pricing policy is known as a **two-part tariff**. Suppose that all consumers have identical utility functions given by $u(x)$ and that the cost of production is cx. If the monopolist sets a two-part tariff, will it produce the socially efficient level of output, too little, or too much?

10.7: Selling cars (medium)

A car dealer must pay \$10,000 to the manufacturer for each car he adds to his inventory. He faces three buyers. From the point of view of the dealer, Smith's valuation is uniformly distributed between \$11,000 and \$21,000, Jones's is between \$9,000 and \$11,000, and Brown's is between \$4,000 and \$12,000. The dealer's policy is to make a separate take-it-or-leave-it offer to each customer, and he is smart enough to avoid making different offers to customers who could resell to each other. Use the notation that the maximum valuation is $\overline{V}$ and the range of valuations is R.

(a) What will the offers be?
(b) Who is most likely to buy a car? How does this compare with the outcome with perfect price discrimination under full information? How does it compare with the outcome when the dealer charges \$10,000 to each customer?

(c) What happens to the equilibrium prices if with probability 0.25 each buyer has a valuation of \$0, but the probability distribution remains otherwise the same? What happens to the equilibrium expected profit?

(d) What happens to the equilibrium price the seller offers to seller Jones if with probability 0.25 Jones has a valuation of \$30,000, but with probability 0.75 his valuation is uniformly distributed between \$9,000 and \$11,000 as before? Show the relation between price and profit on a rough graph.

Regulatory Ratcheting: A Classroom Game for Chapter 10

Electricity demand facing each of several firms is perfectly inelastic at 1 gigawatt per firm. A firm will supply either 0 or 1 gigawatt. The price is chosen by the regulator. The regulator cares about two things: (1) getting electrical service, and (2) getting it at the lowest price possible. The utilities like profit and dislike effort. Throughout the game, utility i has "cost reduction" parameter x_i, which it knows but the regulator does not. This parameter is big if the utility can reduce its costs with just a little effort. Each year, the following events happen.

1 The regulator offers price P_i to firm i.
2 Firm i accepts or rejects.
3 If Firm i accepts, it secretly chooses its effort level e_i,
4 Nature secretly and randomly chooses the economywide shock u (uniform from 1 to 6) and Firm i's shock u_i (uniform from 1 to 6) and announces Firm i's cost, c_i. That cost equals

$$c_i = 20 + u + u_i - x_i e_i. \tag{10.75}$$

5 Firm i earns a period payoff of 0 if it rejects the contract. If it accepts, its payoff is

$$\pi_i = p_i - c_i - e_i^2 \tag{10.76}$$

The regulator earns a period payoff of 0 from firm i if its contract is rejected. Otherwise, its payoff from that firm is

$$\pi_{regulator}(p_i) = 50 - p_i \tag{10.77}$$

All variables take integer values.

The game repeats for as many years as the class has time for, with each firm keeping the same value of x throughout.