than a_2 at these wages? (Can you change the numbers so that this is so?) What would $C(a_2)$ be in that case?

■ 12. Prove proposition 5.

■ 13. Return to the finite action, finite outcome formulation of section 16.3 in the specialization where a is effort and the signals s are levels of gross profits. Assume that assumptions 1, 2, and 4 all hold and that u is strictly concave. An alternative to assumption 5 is: There are two probability distributions $(\pi_1^1, \ldots, \pi_M^1)$ and $(\pi_1^2, \ldots, \pi_M^2)$ such that for every action a_n there is a number $\alpha_n \in [0,1]$ with $\pi_{nm} = \alpha_n \pi_m^1 + (1-\alpha_n)\pi_m^2$. (Assume that α_n is increasing in n. What conditions must hold between π^1 and π^2 so that assumption 4 is guaranteed?) Show that this alternative with assumptions 1, 2, and 4, is sufficient to show that the optimal wage-incentive scheme is nondecreasing in gross profits. (Hint: Do *not* try to prove that the only binding relative incentive constraints will be those of index lower than the effort level that is being implemented.)

chapter seventeen

Adverse selection and market signaling

Imagine an economy in which the currency consists of gold coins. The holder of a coin is able to shave a bit of gold from it in a way that is undetectable without careful measurement; the gold so obtained can then be used to produce new coins. Imagine that some of the coins have been shaved in this fashion, while others have not. Then someone taking a coin in trade for goods will assess positive probability that the coin being given her has been shaved, and thus less will be given for it than if it was certain not to be shaved. The holder of an unshaved coin will therefore withhold the coin from trade; only shaved coins will circulate. This unhappy situation is known as Gresham's law — bad money drives out good.

17.1. Akerlof's model of lemons

Gresham's law has had more recent expression in Akerlof (1970). (See problem 12 after you finish this section, however.) In Akerlof's context, Gresham's law is rephrased as "Bad used cars drive out good." It works as follows.

Suppose there are two types of used cars: peaches and lemons. A peach, if it is known to be a peach, is worth \$3,000 to a buyer and \$2,500 to a seller. (We will assume the supply of cars is fixed and the supply of possible buyers is infinite, so that the equilibrium price in the peach market will be \$3,000.) A lemon, on the other hand, is worth \$2,000 to a buyer and \$1,000 to a seller. There are twice as many lemons as peaches.

If buyers and sellers both had the ability to look at a car and see whether it was a peach or a lemon, there would be no problem: Peaches would sell for \$3,000 and lemons for \$2,000.

Or if neither buyer nor seller knew whether a particular car was a peach or a lemon, we would have no problem (at least, assuming risk neutrality, which we will to avoid complications): A seller, thinking she has a peach with probability 1/3 and a lemon with probability 2/3, has a car that (in expectation) is worth $1,500. A buyer, thinking that the car might be a peach with probability 1/3 and a lemon with probability 2/3, thinks that the car is worth on average $2,333.33. Assuming an inelastic supply of cars and perfectly elastic demand, the market clears at $2,333.33.

Unhappily, it isn't like this with used cars. The seller, having lived with the car for quite a while, knows whether it is a peach or a lemon. Buyers typically can't tell. If we make the extreme assumption that buyers can't tell at all, then the peach market breaks down. To see this, begin by assuming that cars are offered for sale at any price above $1,000. All the lemons will be offered for sale. But only if the price is above $2,500 will any peaches appear on the market. Hence at prices below $2,500 and above $1,000, rational buyers will assume that the car must be a lemon. Why else would the seller be selling? Given this, the buyers conclude that the car is worth only $2,000. And at prices above $2,500, the car has a 2/3 chance of being a lemon, hence is worth $2,333.33. *There is no demand at prices above $2,000*, because (a) above $2,333.33, there is no demand whatsoever — no buyer is willing to pay that much — and (b) below $2,500 there is only demand starting at $2,000, since buyers assume that they must be getting a lemon. So we get as equilibrium: Only lemons are put on the market, at a price of $2,000. Further gains from trade are theoretically possible (between the owners of peaches and buyers), but these gains cannot in fact be realized, because buyers can't be sure that they aren't getting a lemon.

Note that if there were two peaches to every lemon, the story wouldn't be so grim. Then, as long as cars reach the market in these proportions, a buyer is willing to pay $2,666.67. And that is enough to induce owners of peaches to sell; we get the market clearing at this price, with all the cars for sale. (What does the demand curve look like in this case? What does the supply curve look like? Are there other market clearing prices? See problem 1.) Owners of peaches are not pleased about those lemon owners; without them, peach owners would be getting an extra $333.33 for their peaches. But at least peaches can be sold.

This is a highly stylized example of Akerlof's market for lemons. It illustrates the problem of *adverse selection*. Assume a particular good comes in many different qualities. If in a transaction one side but not the other knows the quality in advance, the other side must worry that it will get an adverse selection out of the entire population. The classic example of this is in life/health insurance. If premiums are set at actuarially fair rates

for the population as a whole, insurance may be a bad deal for healthy people, who then will refuse to buy. Only the sick and dying will sign up. And premium rates must then be set to reflect this.

The problem noted above becomes worse the greater the number of qualities of cars and the smaller the "valuation gaps," the differences between what a car is worth to a buyer and a seller, assuming they have the same information. Imagine, for example, that the quality spectrum of used cars runs from real peaches, worth $2,900 to sellers and $3,000 to buyers, down to real lemons, worth $1,900 to sellers and $2,000 to buyers. Between the two extremes are cars of every quality level, always worth $100 more to buyers than to sellers. Suppose that the distribution of quality levels is uniform between these two extremes. To be very specific, suppose there are 10,001 cars, one of which is worth $1,900.00 to its owner and $2,000.00 to buyers, a second worth $1,900.10 to its owner and $2,000.10 to buyers, and so on. What will be the equilibrium then? (Continue to assume inelastic supply and elastic demand at every level of quality.)

Let us draw the supply curve first. At price $p = \$1,900$ there is one car offered for sale; at $p = \$1,901$ there are 11 cars offered, and so on. At $p = \$2,900$, all 10,001 cars are offered for sale, and that is all that is offered at any higher price. If we smooth out supply over the small discrete bumps, we get the supply curve shown in figure 17.1(a).

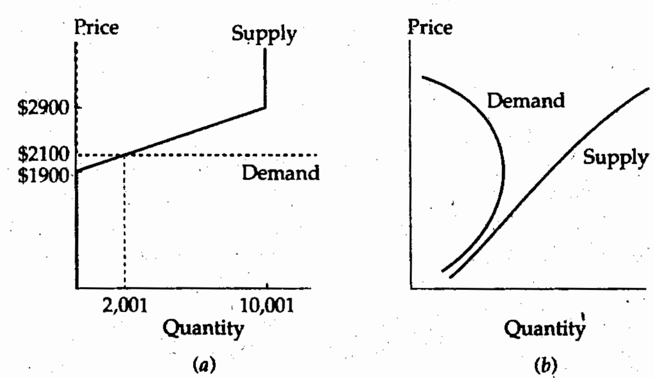

Figure 17.1. Supply and demand with adverse selection.
In (a), we have supply and demand for the second version of the lemons market. In (b), we depict the sad situation of upward sloping demand and a market that shuts down completely.

As for demand, at a price p between \$3,000 and \$2,000, buyers assume that only sellers who value their own cars at p or less are willing to sell. Hence a car being sold at price p has a quality level that makes it worth between \$2,000 and $\$(p + 100)$ to buyers, with each value in this range equally likely. Therefore, the average car being sold is worth $\$(2000 + p + 100)/2$. If p exceeds \$2,100, this average value is something less than p, *and there is no demand*. If p is less than \$2,100, this average value exceeds p, and there is infinite demand. At $p = \$2,100$ (plus or minus a penny) the average car offered for sale is worth \$2,100, and buyers are indifferent between buying and not. So we get the demand curve shown in figure 17.1(a) — no demand at prices exceeding \$2,100, perfectly elastic demand at $p = \$2,100$. The market equilibrium is where supply and demand cross, or $p = \$2,100$, at which only 2,001 of the 10,001 cars change hands. (And if the \$100 difference in valuations between buyers and sellers is \$50 instead, then the equilibrium price is $p = \$2,000$, and only 1,001 of the cars change hands.)

Let us do this one more time (just to pound it into your brain), at one level greater generality. We imagine that some item (a durable good, or some service) comes in one of N quality levels, $q^1, q^2, \ldots, q^N$, arranged in ascending order. The supply of quality n depends on the price p and is given by an upward sloping supply function $s^n(p)$. Demand depends on price p and the average quality in the market, which by the assumptions we've made can be computed as

$$q^{\text{avg}}(p) = \frac{\sum_n q^n s^n(p)}{\sum_n s^n(p)}.$$

Let $D(p, q^{\text{avg}})$ be this demand function, which we will assume is decreasing in p and increasing in q^{avg}. Hence market demand at a price p is obtained, at least if buyers anticipate where supply will come from, as $D(p, q^{\text{avg}}(p))$. To find the slope of market demand we compute

$$\frac{dD}{dp} = \frac{\partial D}{\partial p} + \frac{\partial D}{\partial q} \times \frac{dq^{\text{avg}}(p)}{dp}.$$

The first term on the right-hand side is negative. The second is a positive term, times $dq^{\text{avg}}(p)/dp$, which may be positive; average quality supplied may be increasing in price. Hence it is possible if the second term is large enough that demand has positive slope for some levels of price. And, therefore, it is even possible that we get the sort of picture in figure 17.1(b), where the only intersection between demand and supply is at zero; the market shuts down, even though gains from trade are possible.

17.2. Signaling quality

In spite of the lemons problem, used car markets and life/health insurance markets and all manner of markets subject to adverse selection do function. Why?[1] Often it is because the side to the transaction that has the superior information will do something that indicates the quality of the good being sold. With used cars, the seller may offer a partial warranty or may get the car checked by an independent mechanic.[2] In insurance, medical checkups are sometimes required. Another ploy in the realm of life insurance concerns "golden age" insurance, which is often marketed with the come-on that no one will be turned down. These policies often contain an important bit of fine print: "Benefits are greatly reduced for the first two years." That is, the insurance company is betting that if the buyer knows that he is going to die, it will probably happen quickly. The key to such signals is that the sellers of higher quality, or buyers in the case of insurance, to distinguish themselves from sellers of lower quality are willing to take actions the sellers of lower quality do not find worthwhile.

One can think of adverse selection as a special case of moral hazard and market signals as a special case of incentive schemes. We want the individual to self-identify as having a used car that is a lemon or a peach; as being gravely ill or not; and so on. But there is moral hazard in such self-identification; the seller of a used car cannot be trusted to represent honestly the quality of the car for sale unless provided with some incentive for doing so. The incentive could be relatively direct; the seller will be tossed in jail if he has egregiously misrepresented the quality of the car sold. Or it could be more indirect, as in the sort of market signals just described; the individual could be asked to give a six-month warranty if he says that the car is a peach. The point is that such an "incentive

[1] We saw one possible answer in chapter 14 — reputation. But in this chapter, we consider cases where the informed party isn't able to make use of a reputation, say because this is a one-time transaction. We could also use reputation indirectly, through recourse to a market intermediary such as a reputable used car dealer. But that takes us to topics we will discuss in chapter 20.

[2] Since the buyer can't be sure how independent this mechanic is, in the used car market in the United States it is often the buyer who will provide the mechanic and the money for this. But the signal is still being provided by the seller, insofar as she allows the buyer's mechanic to investigate the car. By way of contrast, in other countries "inspection companies" are so well known and credible that it is customary for the seller to provide a report from one of them. Remember this example when we discuss trilateral governance in chapter 20.

contract" (I will pay you \$X if you don't give me a warranty and \$Y > \$X if you do) gives the party with private information the incentive to self-identify honestly, and the moral hazard problem of self-identification is defeated.

The two classic models of this in the literature are Spence's (1974) job market signaling model and Rothschild and Stiglitz's (1976) model of an insurance market. While cast in different settings, the two seem at first to be about the same problem. But their analyses come to different conclusions. We will use the Spence setting to illustrate their two analyses (and further variations):[3]

We imagine a population of workers, some of high innate ability, some of low. We will use t to denote the level of ability, with $t = 2$ for high ability and $t = 1$ for low. The numbers of high- and low-ability workers are equal. There are also firms that will (eventually) hire the workers; these firms operate in a competitive labor market and so will be making zero profits (when all the dust settles) out of their workers.

The key ingredient comes next: Before going to work, workers seek education that enhances their productivity. Specifically, each worker chooses a level of education e from some set, say $[0, 16]$, for the number of years in school.[4] A worker of type t with education level e is worth precisely te to a firm. Note that for every level of education, a more able worker is worth twice as much as a less able worker. Firms, when they hire a worker, are unable to tell whether the worker is able or not. But they do get to see the worker's c.v., and they thereby learn how many years the worker went to school. Hence they can make wage offers contingent on the number of years that the worker went to school.

What do the workers want out of all this? They want higher wages, to be sure. But they also dislike education. And the less able dislike education even more than do the more able. Specifically, workers seek to maximize a utility function of wages and education, which we will write as $u_t(w, e)$, where w is the wage rate, e the education level, and t the worker's type. We assume that u_t is strictly increasing in w, strictly decreasing in e, and strictly concave. And we make the following very important assumption: Take any point (e, w) and consider the two indifference curves of high- and low-ability workers through that point. (Refer to figure 17.2, noting that the monotonicity and concavity assumptions we made concerning the functions u_t give us strictly convex indifference

[3] In the problems, you will be given some leads on how these two stories are told in the case of insurance markets.

[4] We allow any number between 0 and 16 to keep the pictures relatively simple. A treatment where education levels came only in discrete numbers would not be very different.

curves and the direction of increasing preference as depicted.) *We assume that the indifference curve of the low-ability worker is always more steeply sloped* (as in figure 17.2). That is, to compensate a worker for a given increase in education requires a greater increase in wages for a low-ability worker than for a high-ability worker. Both sorts of worker dislike education, but low-ability workers dislike education relatively more (measured in terms of wage compensation). This assumption is crucial for what follows because it implies that a more able worker finds it relatively cheaper (in terms of utility) to obtain a higher level of education, which then can be used to distinguish more able workers from those of lower ability. The assumption is known informally as the *single-crossing property*.[5] Utility functions of the form $u_t(w, e) = f(w) - k_t g(e)$, for f an increasing and concave function, g an increasing and strictly convex function, and the k_t positive constants with $k_2 < k_1$ constitute an example.

What would happen if the ability level of an employee were observable? Then a high-ability worker with education level e would be paid $2e$, and a low-ability worker with education level e' would be paid e' As in figure 17.3, high- and low-ability workers will pick education levels that maximize their utility, given these wages.

But we are interested here in cases where ability level is not directly observable. Consider the following two stories: *job market signaling* and then *worker self-selection facing a menu of contracts*.

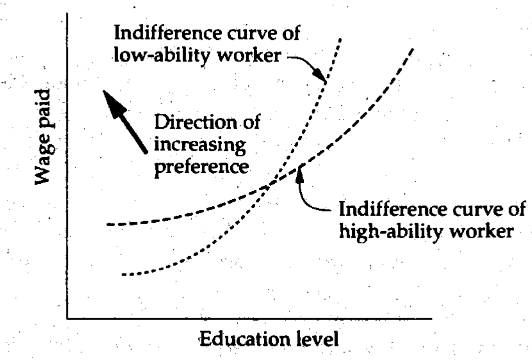

Figure 17.2. Worker indifference curves in wages-education level space.

[5] It is usually formulated with a bit more precision than we have done here, and you will be given the chance to provide a more precise formulation in the problems.

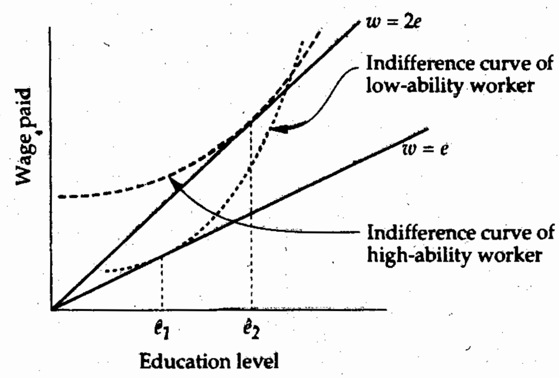

Figure 17.3. The full information equilibrium.
If ability is observable, high-ability workers will select education level $\hat{e}_2$ and be paid $2\hat{e}_2$, and low-ability workers will select $\hat{e}_1$ and be paid $\hat{e}_1$.

Spence's story: Job market signaling

Suppose that workers move first in the following sense. With no guarantees except that they expect the market mechanism to work when the time comes, workers choose how many years to go to school. They do so anticipating some wage function $w(e)$ that gives wages for every level of education. After they spend their time in school, they present themselves to a competitive labor market, and the firms in that market bid for their services.

Formally, an equilibrium in the sense of Spence consists of (a) an anticipated wage function $w(\cdot)$ that gives the wages $w(e)$ that workers anticipate will be paid to anyone obtaining this education level e, for every e, and (b) probability distributions π_t on the set of education levels for each type of worker $t = 1, 2$ such that the following two conditions are met:

(1) For each type t of worker and education level e, $\pi_t(e) > 0$ only if $u_t(w(e), e)$ achieves the maximum value of $u_t(w(e'), e')$ over all e'.

(2) For each education level e such that $\pi_1(e) + \pi_2(e) > 0$,

$$w(e) = \frac{.5\pi_1(e)}{.5\pi_1(e) + .5\pi_2(e)} e + \frac{.5\pi_2(e)}{.5\pi_1(e) + .5\pi_2(e)} 2e. \qquad (\clubsuit)$$

This has the following interpretation: $\pi_t(e)$ gives the proportion of type t workers who select education level e in equilibrium.[a] Then condition (1) is a *self-selection condition*: Based on the wages that workers anticipate, they only select education levels that with the corresponding wage levels maximize their utility. And condition (2) says that at education levels picked in equilibrium the wages paid are appropriate in a competitive labor market for a worker offering that education level. This comes about because the two fractions on the right-hand side of (♣) are the conditional probabilities that the worker is of low and high ability, respectively, obtained by Bayes rule.[6] Hence the right-hand side of (♣) is the conditional expected value of a worker presenting education level e, and (it is assumed) competition among firms pushes the worker's wage to that level.

In the definition of an equilibrium just given, we allow workers of a given type to choose more than one level of education. But we will restrict attention here to equilibria in which all the workers of a given type choose the same level of education:

*In the first type of equilibrium, workers of the two types are **separated**; this is called a **separating equilibrium**. All the workers of type $t = 1$ choose an education level e_1, and all the workers of type $t = 2$ choose an education level $e_2 \neq e_1$. Firms, seeing education level e_1, know that the worker is of type 1, and pay a wage e_1. And if they see education level e_2, they offer a wage of $2e_2$, because they correctly assume that worker is of type 2.*

*In the second type of equilibrium, workers all **pool** at a single education level; this is called a **pooling equilibrium**. All the workers choose some education level e^*, at which point wages are $1.5e^*$.*

The term *screening equilibrium* is often used interchangeably with separating equilibrium; cf. p. 478. In addition to these pooling and separating equilibria, there are other types, which you will explore in problem 2.

[a] If we wanted to be fancy, we would permit general probability distributions, replacing (1) with a statement about e in the support of the distribution and turning (2) into the appropriate statement about conditional probabilities. But this generality would be wasted. If you start with a general formulation, you can prove in this setting, with strictly concave u_t that no type would ever select more than a finite number of education levels; see problem 2.

[6] To be pedantic: $.5\pi_1(e)$ is the joint probability that a worker is of low ability and he chooses education level e obtained as the product of the marginal probability that a worker is of low ability, .5, and the conditional probability that a worker chooses education level e given he has low ability, $\pi_1(e)$. Hence $.5\pi_1(e) + .5\pi_2(e)$ is the marginal probability that a worker chooses education level e. And, by the rules of conditional probability, $.5\pi_1(e)/[.5\pi_1(e) + .5\pi_2(e)]$ is the conditional probability that a worker is of low ability given he chooses education level e.

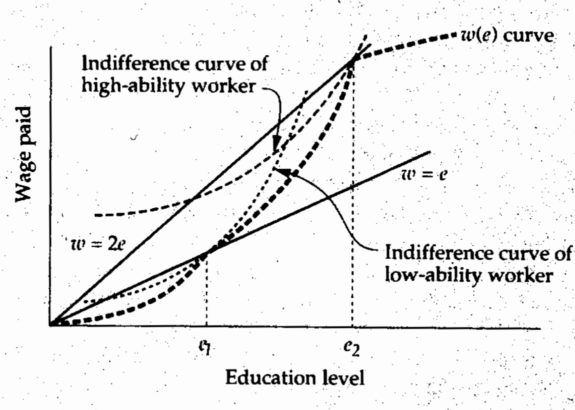

Figure 17.4. A separating equilibrium.
In this picture and in all others to follow, the heavy dashed curve indicates the workers' conjectures about what wages they will receive at every given level of education.

Return to the description of a separating equilibrium. For this to be an equilibrium, we need that workers, anticipating all this, are content to be separated in this fashion. That is, a worker of type 1 would rather choose the wage-education pair ($w = e_1, e = e_1$) than ($w = 2e_2, e = e_2$), and vice versa for workers of type 2. You'll see just such a situation in figure 17.4.

In this figure, note carefully the heavy dashed curve. This represents the function $w(e)$, workers' conjectures about the wages they will receive as a function of the level of education they select. Note that this function must lie everywhere at or below the indifference curves of the workers at points that the workers select; this is the self-selection condition. Condition (2) for an equilibrium pins down the $w(e)$ curve at the two levels of education selected, as shown.

This does not restrict the curve $w(e)$ at points e that are not selected in equilibrium. What restrictions might we wish to place on $w(e)$ at such points? For now, consider the following three possibilities:

(a) No restrictions whatsoever are placed on $w(e)$ at points e that are not selected in equilibrium. (We require throughout that $w(e)$ is nonnegative for every value of e; that is, the institution of slavery is not part of our economy.)

(b) We require that $e \leq w(e) \leq 2e$ for all e.

(c) We require (b) and, in addition, that $w(e)/e$ is nondecreasing.

What might motivate us to assume (b) or (c)? We reason as follows. A worker who chooses education level e is worth either e or $2e$ to a firm. The worker cannot possibly be worth more than $2e$ or less than e. Now if firms are unsure of the worker's level of ability, firms cannot be certain which of e or $2e$ is correct. And so, depending on their assessments for the chances that the worker is worth e or $2e$, they might be willing to bid anywhere between these two levels. But they would never pay more than $2e$, and the forces of competition between firms mean that they could never get away with paying less than e. Workers anticipate all this, so they anticipate that $e \leq w(e) \leq 2e$. As for (c), take this a step further and imagine that workers conjecture that firms will assess that a worker who chooses education level e will be of low ability with probability $\alpha(e)$. Assuming firms are risk neutral (which we do) and that they share assessments, competition amongst them would lead them to bid $w(e) = \alpha(e)e + (1-\alpha(e))2e$. (Note that this is just another way of giving our justification for [b].) And now (c) can be translated as: $\alpha(e)$ is nonincreasing. Or firms' conjectures as to the ability of a worker do not give a higher probability for a lower-ability worker if the worker obtains more education.

At the risk of (re)stating the obvious, let us rephrase the equilibrium condition (2) in these terms. *We require that firms' assessments given by $\alpha(e)$ are confirmed at education levels that workers do select in equilibrium.* Or, in symbols,

$$\alpha(e) = \frac{.5\pi_1(e)}{.5\pi_1(e) + .5\pi_2(e)}$$

If we imagine a stable population with new workers coming along and presenting themselves for employment each year, and if we imagine that the equilibrium distribution of workers' abilities and education levels is stable, then we would be saying that firms, based on their experience, know that distribution. What we do not restrict, or restrict only minimally, are the conjectures of firms at education level choices that are not chosen and, therefore, at which the firms have no experience.

You may be thinking that we shouldn't speak directly about the firms' conjectures concerning the ability of a worker who chooses a nonequilibrium level of education. The function $w(e)$ represents the workers' conjectures about the wages they will be paid. Hence, at least, we should speak of

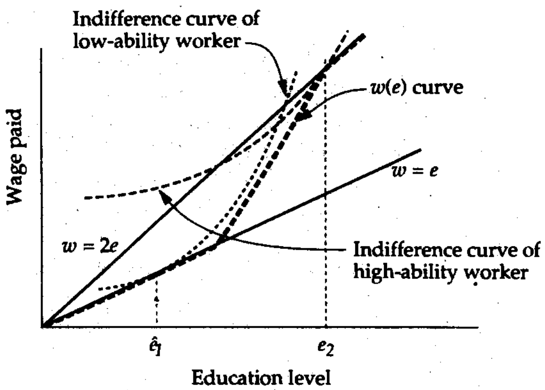

Figure 17.5. Another separating equilibrium.

the workers' conjectures concerning the firms' assessments about the ability of a worker who chooses a nonequilibrium level of education. And then, by having a single $w(e)$ function, we have implicitly assumed that all the workers have the same conjectures about the firms' assessments, and this conjecture holds that all the firms will have the same assessments. Are these assumptions of coincident conjectures and assessments necessary? If you are very careful, you should be able to see that much of what we will say does not depend on this, although we will hold to this implicit assumption to keep the exposition simple. You may also wonder why we have not introduced game-theoretic terminology for all of this. We will do so in the next section, but there are reasons to put this off for a while.

Note that the function $w(e)$ in figure 17.4 does not satisfy (b) or (c), although it does satisfy one somewhat obvious condition, namely that it is nondecreasing. In figure 17.5 we give a second separating equilibrium in which $w(e)$ does satisfy (b).

Note in figure 17.5 the point chosen by the low-ability workers, denoted by $\hat{e}_1$. This, as drawn, is the point along the ray $w = e$ that low ability workers like most; notice that their indifference curve through this point is tangent to $w = e$. In contrast, in figure 17.4, low-ability workers get less utility than they would get from $(w = \hat{e}_1, e = \hat{e}_1)$. We use $\hat{e}_1$

throughout this chapter to denote this particular point of tangency, and we have the following result:

Proposition 1. *In any separating equilibrium for which the function $w(e)$ satisfies (b), low-ability workers choose precisely $\hat{e}_1$ and get the corresponding wages. In any equilibrium (separating or not) for which the function $w(e)$ satisfies (b), low-ability workers get at least the utility that they get from $(w = \hat{e}_1, e = \hat{e}_1)$.*

You should have no problems at all seeing why this is true. Or, rather, you should have no problems with the first part of this proposition. The second part may be a tiny bit harder.

Pictures of pooling equilibria are relatively easy to draw. In a pooling equilibrium, recall, we assume that all the workers choose a single education level e^*. As before, firms' beliefs, seeing e^*, must be confirmed. Since we have supposed that the numbers of high- and low-ability workers in the population are equal, and since (in this equilibrium) all workers choose e^*, we see that in terms of the notation above $\alpha(e^*) = .5$, and $w(e^*) = 1.5e^*$. An equilibrium of this sort is given in figure 17.6. Note that the function $w(e)$ in this case satisfies (c). Note also that the function $w(e)$ has a kink at the pooling point. This is necessary; without a kink, we couldn't have $w(e)$ underneath both types' indifference curves and touching those curves at the pooling point.

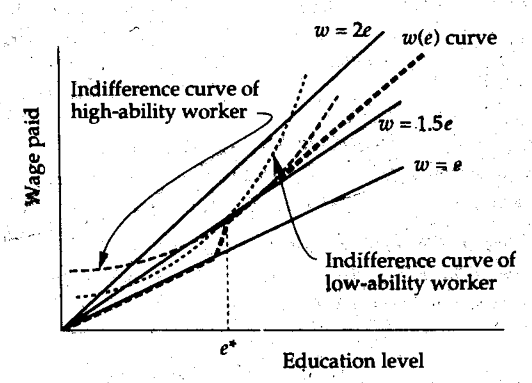

Figure 17.6. A pooling equilibrium.

As you can see (and imagine), there are a lot of equilibria here. Is there any reason to suspect that one is more plausible than another? We will get back to this question, but before doing so we look at another way to tell the basic story of market signaling.

The story of Rothschild and Stiglitz: Worker self-selection from a menu of contracts

In the story just told, workers choose education levels in anticipation of offers from the firm, and (we assume) those anticipations are correct, at least for the education levels actually selected. Let's turn this around and suppose that the firms move first. Specifically, suppose that firms offer to workers a number of "contracts" of the form (w, e) before workers go off to school. Workers consider the menu of contracts on offer, sign the one they like best, and then go off to school, content in the knowledge of what wage they will get once school is done (assuming they complete the number of years of schooling for which they have contracted). This story may seem a bit lame in the context of education level choices. If you find it so, think instead of insurance markets, the context with which Rothschild and Stiglitz deal explicitly. Insurance companies are willing to insure one's life in any of a number of ways, with different premiums, death benefits, exclusions, and so on. A customer shops for the policy best suited to himself, given both his preferences and, what is important here, the knowledge he has about his prospects for a long and healthy life.

An equilibrium in the sense of Rothschild and Stiglitz consists of (a) a menu of contracts, a set of pairs $\{(w^1, e^1), (w^2, e^2), \ldots, (w^k, e^k)\}$ for some finite integer k[b] and (b) a selection rule by which workers are "assigned" to contracts or, for each type t, a probability distribution π_t over $\{1, 2, \ldots, k\}$, that satisfy three conditions:

(1) Each type of worker is assigned only to contracts that are best for that worker among all the contracts in the menu. In symbols, $\pi_t(j) > 0$ only if $u_t(w^j, e^j)$ achieves the maximum of $u_t(w^{j'}, e^{j'})$ for $j' = 1, \ldots, k$.

(2) Each contract in the menu to which workers are assigned at least breaks even on average. (Otherwise, firms offering that contract would withdraw the contract.) In symbols, for each $j = 1, \ldots, k$, if $\pi_1(j) + \pi_2(j) > 0$, then

$$w^j \leq \frac{.5\pi_1(j)}{.5\pi_1(j) + .5\pi_2(j)} e^j + \frac{.5\pi_2(j)}{.5\pi_1(j) + .5\pi_2(j)} 2e^j.$$

[b] We assume that the menu of contracts offered is finite mostly for expositional convenience.

(3) No contract can be created that if offered in addition to those in the menu would make strictly positive profits for the firm offering it, assuming that workers choose among contracts in a manner consistent with rule (1) above.

We won't try to write (3) out formally — it makes a good exercise if you are fascinated by symbol manipulation.[c]

This change in formulation has dramatic effects. To begin the analysis, we assert something common to the Spence formulation:

Proposition 2. *In an equilibrium, any contract that is taken up by workers must earn precisely zero expected profits per worker.*

The proof of this is a little tricky.[7] A natural argument to try runs as follows: Suppose that (w', e') is offered, is taken by some workers, and earns an expected profit of size ϵ per worker who took it. Then have some firm offer $(w' + \epsilon/2, e')$. This new contract will attract all the workers who previously were attracted by (w', e') and will still return a profit of $\epsilon/2$, contradicting condition (3) of an equilibrium. The trouble with this argument is that the new contract may attract others besides those who previously took (w', e'), and those others may render this contract unprofitable.

Only a sketch of the correct proof will be given, with details left for the reader to fill in. First, the argument just given works if the contract (w', e') attracts only low-ability workers. Then if $(w' + \epsilon/2, e')$ manages to attract any high-ability workers, it is even more profitable. So the hard case is where (w', e') attracts some high-ability workers. At this point, wait to read through the proof of proposition 3, which shows how to break a pooling equilibrium. The key to that argument is that it is possible to devise a contract $(w' + \epsilon/2, e' + \delta)$, for $\delta > 0$ that is more attractive to high-ability workers than is (w', e') but is less attractive to low-ability workers. Hence if a firm offers this contract, it will attract all the high-ability workers who previously chose (w', e'), leaving the low-ability workers at (w', e') (or somewhere else). If (w', e') made profits ϵ per worker, then this new contract makes profits exceeding $\epsilon/2$ per worker.

[c] And if you do try to write this out formally, you will run up against the following question: Should we insist that the contract earn nonpositive profits for every assignment consistent with (1), or just that it make nonpositive profits for some assignment consistent with (1)? It won't matter to the theory which you choose, as long as you are careful about the arguments that follow.

[7] In Spence's formulation, it is true by fiat.

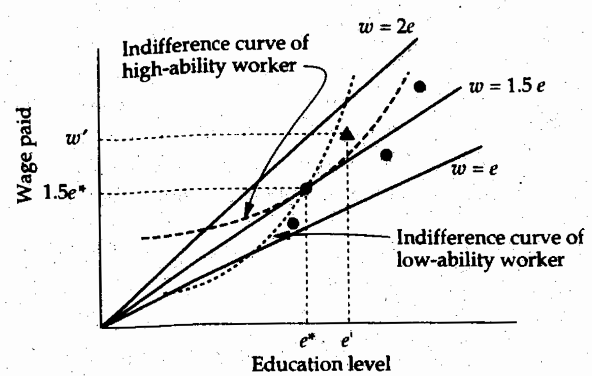

Figure 17.7. Destroying a pooling equilibrium.

Now we give a result that distinguishes this model from that of Spence:

Proposition 3. *It is impossible that an equilibrium is a pooling equilibrium.*

The argument runs as follows: Consider figure 17.7, and the pooling equilibrium depicted. Suppose the firms are offering a menu of contracts that causes all the workers to choose $(1.5e^*, e^*)$ as shown. In this figure, the filled-in dots are the menu of contracts that we suppose are offered. We have added a few contracts to the menu besides $(1.5e^*, e^*)$, although they are irrelevant, since (in our pooling equilibrium) the contract $(1.5e^*, e^*)$ is best for all types of worker.

From this position, any one of the firms can offer the contract (w', e') that is shown as a filled-in triangle. Saying in words what the picture shows: This is a contract that has slightly higher wages and education levels than $(1.5e^*, e^*)$, where the increased wages more than compensate a high-ability worker but don't compensate a low-ability worker relative to the pooling contract. Since this contract is added to the contracts already in the menu, it will only attract the high-ability workers; all the low-ability workers prefer the e^* offer. But if all the high-ability workers flock to (w', e'), and none of the low-ability workers do so, each worker who chooses (w', e') will be worth $2e'$ to the firm that makes this offer. The firm makes a profit, and we can't have an equilibrium.

The same sort of argument can be used to establish

Proposition 4. *It is impossible, in equilibrium, that any contract (w, e) is taken by positive fractions of the high- and low-ability workers both. Or, in other words, the only possible equilibria are separating equilibria.*

The details are left to you; a picture very much like figure 17.7 is the key.

Next we establish

Proposition 5. *There is a single candidate for a separating equilibrium. In this candidate equilibrium, low-ability workers choose the contract $(\hat{e}_1, \hat{e}_1)$, where $\hat{e}_1$ is defined as before, and ...* (to be continued)

Why must this be so? Because if low-ability workers are separated and are choosing any other contract, a firm could add to the menu of contracts a contract $(w = \hat{e}_1 - \delta, e = \hat{e}_1)$ for $\delta > 0$ small enough so that low-ability workers prefer this to the contract they are choosing in the supposed equilibrium. But then for any $\delta > 0$, this contract must be profitable.

Fixing $\hat{e}_1$ at this value, define the education level $\hat{f}_1$ to be that level of education such that low-ability workers are indifferent between $(w = \hat{e}_1, e = \hat{e}_1)$ and $(w = 2\hat{f}_1, e = \hat{f}_1)$. Figure 17.8 shows this point. And let $\hat{f}_2$ solve the problem: Maximize $u_2(2e, e)$ subject to $e \geq \hat{f}_1$. That is, $\hat{f}_2$ is the

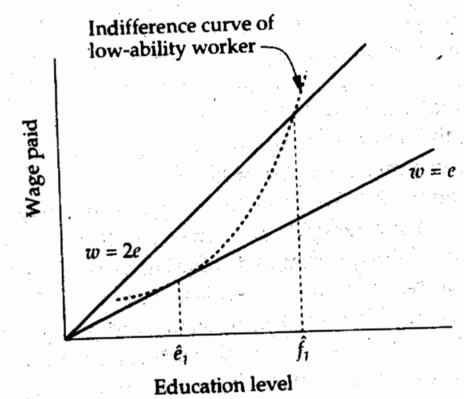

Figure 17.8. Determination of $\hat{f}_1$.

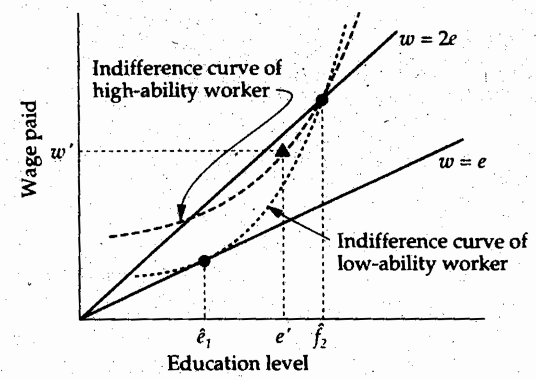

Figure 17.9. A separating equilibrium.

education level that high-ability workers would choose if they could have any point along the ray $w = 2e$ for $e \geq \hat{f}_1$.

Proposition 5, continued. *. . . and (in the single candidate separating equilibrium) high-ability workers get* $(w = 2\hat{f}_2, e = \hat{f}_2)$.

Why? Because if they are separated at any other contract, some firm could come along and offer a contract $(w = 2\hat{f}_2 - \delta, e = \hat{f}_2)$ that for δ sufficiently small would be more attractive to high-ability workers than the contract they are taking at the supposed equilibrium. Since $\hat{f}_2 \geq \hat{f}_1$, this contract is less appealing than $(w = \hat{e}_1, e = \hat{e}_1)$ for the low-ability workers. Hence this contract will attract precisely the high-ability workers. And thus for any $\delta > 0$, this contract is strictly profitable.

Consider figure 17.9. We have there depicted the proposed separating equilibrium in a case where $\hat{f}_2 = \hat{f}_1$. It may be helpful to say why no firm would try to break this equilibrium by offering a contract in a position such as (w', e') that is shown, with w' a bit less than $2e'$ and e' a bit less than $\hat{f}_1$. This contract, added to the menu, would certainly attract high-ability workers. And every high-ability worker attracted would be profitable at this wage. But it would *also* attract all the low-ability workers. And in this population, if you attract high- and low-ability workers in proportion to the population, you have to pay less than 1.5 of the education level they pick to make a profit.

Now consider figure 17.10. This is just like figure 17.9, except that the indifference curve of the high-ability workers through the point $(w = 2\hat{f}_2, e = \hat{f}_2)$ dips below the line $w = 1.5e$. In this case, a firm could offer a contract such as (w', e') as shown, below the line $w = 1.5e$, but still above the high-ability worker's indifference curve. This would break the equilibrium, because even though it attracts all the workers, high-ability and low-, it is still profitable.

In figure 17.10, then, there can be no equilibrium at all. Any sort of pooling is inconsistent with equilibrium. And the only possible separating equilibrium can also be broken. In contrast, although we won't go through all the details, in situations such as figure 17.9 (or situations where $\hat{f}_2 > \hat{f}_1$), the candidate separating equilibrium is an equilibrium. Summarizing all this:

Proposition 6. *In the formulation of Rothschild and Stiglitz, there is at most one equilibrium. In the candidate equilibrium, low-ability workers choose* $(w = \hat{e}_1, e = \hat{e}_1)$ *and high-ability workers choose the education level* $\hat{f}_2$ *such that* $(w = 2\hat{f}_2, e = \hat{f}_2)$ *is their most preferred point along the ray* $w = 2e$ *for* $e \geq \hat{f}_1$. *If the indifference curve for high-ability workers through the point* $(w = 2\hat{f}_2, e = \hat{f}_2)$ *dips below the pooling line* $w = 1.5e$, *then there is no equilibrium at all. If this indifference curve stays above (or just touches) the pooling line, then this single candidate equilibrium is an equilibrium.*

Quite a difference from Spence's model, where there were many possible equilibria!

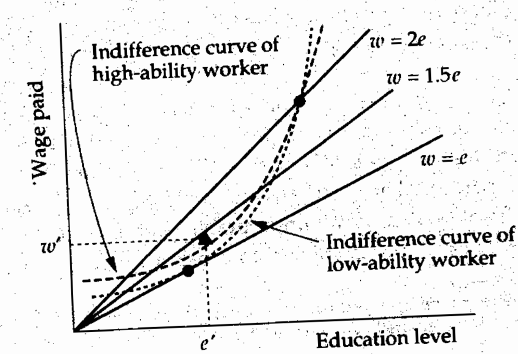

Figure 17.10. No equilibrium at all.